5

A duality toolbox

5.1 Gauge/gravity duality

In the previous chapter we outlined the string theory reasoning behind the equivalence (4.27) between $\mathcal{N} = 4$ $SU(N_c)$ SYM theory and type IIB string theory on $\text{AdS}_5 \times \text{S}^5$. $\mathcal{N} = 4$ SYM theory is the unique maximally supersymmetric gauge theory in $(3 + 1)$ dimensions, whose field content includes a gauge field, six real scalars, and four Weyl fermions, all in the adjoint representation of the gauge group. The metric of $\text{AdS}_5 \times \text{S}^5$ is given by

$$ds^2 = ds^2_{\text{AdS}_5} + R^2 d\Omega_5^2 \,, \tag{5.1}$$

with

$$ds^2_{\text{AdS}_5} = \frac{r^2}{R^2}\eta_{\mu\nu}dx^\mu dx^\nu + \frac{R^2}{r^2}dr^2 \,, \qquad r \in (0, \infty) \,. \tag{5.2}$$

In the above equation $x^\mu = (t, \vec{x})$, $\eta_{\mu\nu}$ is the Minkowski metric in four spacetime dimensions, and $d\Omega_5^2$ is the metric on a unit five-sphere. The metric (5.2) covers the so-called "Poincaré patch" of a global AdS spacetime, and it is sometimes convenient to rewrite (5.2) using a new radial coordinate $z = R^2/r \in (0, \infty)$, in terms of which we have

$$ds^2_{\text{AdS}_5} = g_{MN}dx^M dx^N = \frac{R^2}{z^2}\left(\eta_{\mu\nu}dx^\mu dx^\nu + dz^2\right) \,, \quad x^M = (z, x^\mu) \,, \tag{5.3}$$

as used earlier in (4.10).

In (5.3), each constant-z slice of AdS_5 is isometric to four-dimensional Minkowski spacetime with x^μ identified as the coordinates of the gauge theory (see also fig. 4.1). As $z \to 0$ we approach the "boundary" of AdS_5. This is a boundary in the conformal sense of the word but not in the topological sense, since the prefactor R^2/z^2 in (5.3) approaches infinity there. Although this concept can be given a precise mathematical meaning, we will not need these details here. As motivated in

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

Section 4.1.1 it is natural to imagine that the Yang–Mills theory lives at the boundary of AdS_5. For this reason, below we will often refer to it as the boundary theory. As $z \to \infty$, we approach the so-called Poincaré horizon, at which the prefactor R^2/z^2 and the determinant of the metric go to zero.

5.1.1 UV/IR connection and renormalization group flow

Owing to the warp factor R^2/z^2 in front of the Minkowski metric in (5.3), energy and length scales along Minkowski directions in AdS_5 are related to those in the gauge theory by a z-dependent rescaling. More explicitly, consider an object with energy $E_{\rm YM}$ and size $d_{\rm YM}$ in the gauge theory. These are the energy and the size of the object measured in units of the coordinates t and $\vec{x}$. From (5.3) we see that the corresponding proper energy E and proper size d of this object in the bulk are

$$d = \frac{R}{z}\, d_{\rm YM}\,, \qquad E = \frac{z}{R}\, E_{\rm YM}\,, \tag{5.4}$$

where the second relation follows from the fact that the energy is conjugate to time, and so it scales with the opposite scale factor than d. We thus see that physical processes in the bulk with identical proper energies but occurring at different radial positions correspond to different gauge theory processes with energies that scale as $E_{\rm YM} \sim 1/z$. In other words, a gauge theory process with a characteristic energy $E_{\rm YM}$ is associated with a bulk process localized at $z \sim 1/E_{\rm YM}$ [594, 768, 671]. This relation between the radial direction z in the bulk and the energy scale of the boundary theory makes concrete the heuristic discussion of Section 4.1.1 that led us to identify the evolution of the bulk metric along the z-direction with the renormalization group flow of the gauge theory. In particular the high energy (UV) limit $E_{\rm YM} \to \infty$ corresponds to $z \to 0$, i.e. to the near-boundary region, while the low energy (IR) limit $E_{\rm YM} \to 0$ corresponds to $z \to \infty$, i.e. to the near-horizon region.

In a conformal theory, there exist excitations of arbitrarily low energies. This is reflected in the bulk in the fact that the geometry extends all the way to $z \to \infty$. As we will see in Section 5.2.2, for a confining theory with a mass gap m, the geometry ends smoothly at a finite value $z_0 \sim 1/m$. Similarly, at a finite temperature T, which provides an effective IR cut-off, the spacetime will be cut-off by an event horizon at a finite $z_0 \sim 1/T$ (see Section 5.2.1).

There is a large literature on what is often referred to as the "holographic renormalization group", namely mapping the radial evolution in the bulk gravity theory to the renormalization group flow equations of its dual boundary theory. For examples, see Refs. [33, 47, 379, 326, 112, 359, 310]. The basic goal is to relate the Einstein equations that describe how the bulk geometry changes as a function of position in the radial direction to the renormalization group equations that describe how the boundary quantum field theory changes as a function of energy scale, given that the boundary energy scale E is associated with a radial position $z(E) \sim 1/E$. Indeed, there have been recent efforts to develop precise parallels between the Wilsonian procedure of integrating out high energy degrees of freedom and integrating out a part of the bulk geometry near the boundary [438, 348]. One identifies the boundary theory Wilsonian effective action obtained by integrating out modes with energies larger than E with a bulk theory effective action obtained by integrating over all the bulk fields including the metric in the region of the bulk geometry that lies between the boundary at $z = 0$ and $z(E)$. The result of doing this partial path integral is an effective action defined on the $z = z(E)$ slice which governs the dynamics of the remaining bulk fields in the unintegrated part of the geometry and which can be mapped onto the boundary theory Wilsonian effective action. There has also been progress toward deriving the bulk gravity theory from the Wilsonian renormalization group flow of a boundary theory [572, 329, 574, 573].

5.1.2 Strong coupling from gravity

$\mathcal{N} = 4$ SYM theory is a scale-invariant theory characterized by two parameters: the Yang–Mills coupling g and the number of colors N_c. The theory on the right-hand side of (4.27) is a quantum gravity theory in a maximally symmetric spacetime which is characterized by the Newton's constant G and the string scale ℓ_s in units of the curvature radius R. The relations between these parameters are given by (4.29). Recalling that $G \sim \ell_p^8$, with ℓ_p the Planck length, these relations imply

$$\frac{\ell_p^8}{R^8} \propto \frac{1}{N_c^2}, \qquad \frac{\ell_s^2}{R^2} \propto \frac{1}{\sqrt{\lambda}}, \tag{5.5}$$

where $\lambda = g^2 N_c$ is the 't Hooft coupling and we have omitted only purely numerical factors.

The full IIB string theory on $\mathrm{AdS}_5 \times S^5$ is rather complicated and right now a systematic treatment of it is not available. However, as we will explain momentarily, in the limit

$$\frac{\ell_p^8}{R^8} \ll 1, \qquad \frac{\ell_s^2}{R^2} \ll 1 \tag{5.6}$$

the theory dramatically simplifies and can be approximated by classical supergravity, which is essentially Einstein's general relativity coupled to various matter fields. An immediate consequence of the relations (5.5) is that the limit (5.6) corresponds to

$$N_c \gg 1, \qquad \lambda \gg 1. \tag{5.7}$$

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

Equation (4.27) then implies that the planar, strongly coupled limit of the SYM theory can be described using just classical supergravity.

Let us return to why string theory simplifies in the limit (5.6). Consider first the requirement $\ell_s^2 \ll R^2$. This can be equivalently rewritten as $m_s^2 \gg \mathcal{R}$ or as $T_{\rm str} \gg \mathcal{R}$, where $\mathcal{R} \sim 1/R^2$ is the typical curvature scale of the space where the string is propagating. The condition $m_s^2 \gg \mathcal{R}$ means that one can omit the contribution of all the massive states of *microscopic* strings in low energy processes. In other words, only the massless modes of microscopic strings, i.e. the supergravity modes, need to be kept in this limit. This is tantamount to treating these strings as pointlike particles and ignoring their extended nature, as one would expect from the fact that their typical size, ℓ_s, is much smaller than the typical radius of curvature of the space where they propagate, R. The so-called α'-expansion on the string side (with $\alpha' = \ell_s^2$), which incorporates stringy effects associated with the finite length of the string in a derivative expansion, corresponds on the gauge theory side to an expansion around infinite coupling in powers of $1/\sqrt{\lambda}$.

The extended nature of the string, however, cannot be ignored in all cases. As we will see in the context of the Wilson loop calculations of Section 5.4 and in many other examples in Chapter 8, the description of certain physical observables requires one to consider long, *macroscopic* strings whose typical size is much larger than R – for example, this happens when the string description of such observables involves non-trivial boundary conditions on the string. In this case the full content of the second condition in (5.6) is easily understood by rewriting it as $T_{\rm str} \gg \mathcal{R}$. This condition says that the tension of the string is very large compared to the typical curvature scale of the space where it is embedded, and therefore implies that fluctuations around the classical shape of the string can be neglected. These long strings can still break and reconnect, but in between such processes their dynamics is completely determined by the Nambu–Goto equations of motion. In these cases, the α'-expansion (that is, the expansion in powers of $1/\sqrt{\lambda}$) incorporates stringy effects associated with fluctuations of the string that are suppressed at $\lambda \to \infty$ by the tension of the string becoming infinite in this limit. From this viewpoint, the fact that the massive modes of microscopic strings can be omitted in this limit is just the statement that string fluctuations around a pointlike string can be neglected.

Consider now the requirement $\ell_p^8 \ll R^8$. Since the ratio ℓ_p^8/R^8 controls the strength of quantum gravitational fluctuations, in this regime we can ignore quantum fluctuations of the spacetime metric and talk about a fixed spacetime like $\mathrm{AdS}_5 \times S^5$. The quantum gravitational corrections can be incorporated in a power series in ℓ_p^8/R^8, which corresponds to the $1/N_c^2$ expansion in the gauge theory. Note from (4.28) that taking the $N_c \to \infty$ limit at fixed λ corresponds to taking the string coupling $g_s \to 0$, meaning that quantum corrections corresponding to loops of string breaking off or reconnecting are suppressed in this limit.

In summary, we conclude that the strong coupling limit in the gauge theory suppresses the stringy nature of the dual string theory, whereas the large-N_c limit suppresses its quantum nature. When both limits are taken simultaneously the full string theory reduces to a classical gravity theory with a finite number of fields.

Given that the S^5 factor in (5.1) is compact, it is often convenient to express a ten-dimensional field in terms of a tower of fields in AdS_5 by expanding it in terms of harmonics on S^5. For example, the expansion of a scalar field $\phi(x, \Omega)$ can be written schematically as

$$\phi(x, \Omega) = \sum_{\ell} \phi_\ell(x) Y_\ell(\Omega) , \tag{5.8}$$

where x and Ω denote coordinates in AdS_5 and S^5 respectively, and $Y_l(\Omega)$ denote the spherical harmonics on S^5. Thus, for many purposes (but not all) the original duality (4.27) can also be considered as the equivalence of $\mathcal{N} = 4$ SYM theory (at strong coupling) with a gravity theory in AdS_5 only. This perspective is very useful in two important aspects. First, it makes manifest that the duality (4.27) can be viewed as an explicit realization of the holographic principle mentioned in Section 4.1.1, with the bulk spacetime being AdS_5 and the boundary being four-dimensional Minkowski spacetime. Second, as we will mention in Section 5.2.3, this helps to give a unified treatment of many different examples of the gauge/gravity duality. In most of this book we will adopt this five-dimensional perspective and work only with fields in AdS_5.

After dimensional reduction on S^5, the supergravity action can be written as

$$S = \frac{1}{16\pi G_5} \int d^5x \left[\mathcal{L}_{\rm grav} + \mathcal{L}_{\rm matt} \right] , \tag{5.9}$$

where

$$\mathcal{L}_{\rm grav} = \sqrt{-g} \left(\mathcal{R} + \frac{12}{R^2} \right) \tag{5.10}$$

is the Einstein–Hilbert Lagrangian with a negative cosmological constant $\Lambda = -6/R^2$ and $\mathcal{L}_{\rm matt}$ is the Lagrangian for matter fields. In the general case, the latter would include the infinite towers $\phi_\ell(x)$ coming from the expansion on the S^5. The metric (5.3) is a maximally symmetric solution of the equations of motion derived from the action (5.9) with all matter fields set to zero.

The relation between the effective five-dimensional Newton's constant G_5 and its ten-dimensional counterpart G can be read off from the reduction of the Einstein–Hilbert term,

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

$$\frac{1}{16\pi G}\int d^5x d^5\Omega\,\sqrt{-g_{10}}\,\mathcal{R}_{10} = \frac{R^5\Omega_5}{16\pi G_5}\int d^5x\,\sqrt{-g_5}\,\mathcal{R}_5 + \cdots\,, \tag{5.11}$$

where $\Omega_5 = \pi^3$ is the volume of a unit S^5. This implies

$$G_5 = \frac{G}{\Omega_5 R^5} = \frac{G}{\pi^3 R^5}\,, \qquad \text{i.e.} \quad \frac{G_5}{R^3} = \frac{\pi}{2N_c^2}\,, \tag{5.12}$$

where in the last equation we made use of (4.29).

5.1.3 Symmetries

Let us now examine the symmetries on both sides of the correspondence. The $\mathcal{N} = 4$ SYM theory is invariant not only under dilatations but under $\mathrm{Conf}(1, 3) \times SO(6)$. The first factor is the conformal group of four-dimensional Minkowski space, which contains the Poincaré group, the dilatation symmetry generated by D, and four special conformal transformations whose generators we will denote by K_μ. The second factor is the R-symmetry of the theory under which the ϕ^i in (4.16) transform as a vector. In order to provide an analogy of the baryon number in QCD, we will often select a $U(1)$ subgroup within the R-symmetry group and define a conserved, Abelian R-charge from its associated Noether current. In addition, the theory is invariant under sixteen ordinary or "Poincaré" supersymmetries, the fermionic superpartners of the translation generators P_μ, as well as under 16 special conformal supersymmetries, the fermionic superpartners of the special conformal symmetry generators K_μ.

The string side of the correspondence is of course invariant under the group of diffeomorphisms, which are gauge transformations. The subgroup of these consisting of large gauge transformations that leave the asymptotic (i.e. near the boundary) form of the metric invariant is precisely $SO(2, 4) \times SO(6)$. The first factor, which is isomorphic to $\mathrm{Conf}(1, 3)$, corresponds to the isometry group of AdS_5, and the second factor corresponds to the isometry group of S^5. As usual, large gauge transformations must be thought of as global symmetries, so we see that the bosonic global symmetry groups on both sides of the correspondence agree. In more detail, the Poincaré group of four-dimensional Minkowski spacetime is realized inside $SO(2, 4)$ as transformations that act separately on each of the constant-z slices in (5.3) in an obvious manner. The dilation symmetry of Minkowski spacetime is realized in AdS_5 as the transformation $(t, \vec{x}) \to C(t, \vec{x})$, $z \to Cz$ (with C a positive constant), which indeed leaves the metric (5.3) invariant. The four special conformal transformations of Minkowski spacetime are realized in a slightly more involved way as isometries of AdS_5.

An analogous statement can be made for the fermionic symmetries. $\mathrm{AdS}_5 \times S^5$ is a maximally supersymmetric solution of type IIB string theory, and so it possesses

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

32 Killing spinors which generate fermionic isometries. These can be split into two groups that match those of the gauge theory.[1]

We therefore conclude that the global symmetries are the same on both sides of the duality. It is important to note, however, that on the gravity side the global symmetries arise as large gauge transformations. In this sense there is a correspondence between global symmetries in the gauge theory and gauge symmetries in the dual string theory. This is an important general feature of all known gauge/gravity dualities, to which we will return below after discussing the field/operator correspondence. It is also consistent with the general belief that the only conserved charges in a theory of quantum gravity are those associated with global symmetries that arise as large gauge transformations.

5.1.4 Matching the spectrum: the field/operator correspondence

We now consider the mapping between the spectra of the two theories. To motivate the main idea, we begin by recalling that the SYM coupling constant g^2 is identified (up to a constant) with the string coupling constant g_s. As discussed below (4.15), in string theory this is given by $g_s = e^{\Phi_\infty}$, where Φ_∞ is the value of the dilaton at infinity, in this case at the AdS boundary, ∂AdS. This suggests that deforming the gauge theory by changing the value of a coupling constant corresponds to changing the value of a bulk field at ∂AdS. More generally, one may imagine deforming the gauge theory action as

$$S \to S + \int d^4x \, \phi(x)\mathcal{O}(x)\,, \tag{5.13}$$

where $\mathcal{O}(x)$ is a local, gauge-invariant operator and $\phi(x)$ is a possibly point-dependent coupling, namely a source. If $\phi(x)$ is constant, then the deformation above corresponds to simply changing the coupling for the operator $\mathcal{O}(x)$. The example of g suggests that to each possible source $\phi(x)$ for each possible local, gauge-invariant operator $\mathcal{O}(x)$ there must correspond a dual bulk field $\Phi(x, z)$ (and vice versa) such that its value at the AdS boundary may be identified with the source, namely:

$$\phi(x) = \Phi|_{\partial \mathrm{AdS}}(x) \equiv \lim_{z\to 0} z^{\alpha_\Phi} \Phi(x, z)\,. \tag{5.14}$$

The power α_Φ in the last expression is chosen so that the limit is well-defined, and is thus determined by the boundary asymptotic behavior of $\Phi(x, z)$. The explicit asymptotic behavior of various types of fields, and hence the values of their α_Φ, will be discussed below and in the next subsection.

[1] In both boundary and bulk, bosonic and fermionic symmetries combine together to form a supergroup $SU(2, 2|4)$.

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

This one-to-one map between bulk fields in AdS and local, gauge-invariant operators in the gauge theory is known as the field/operator correspondence. The field and the operator must have the same quantum numbers under the global symmetries of the theory, but there is no completely general and systematic recipe to identify the field dual to a given operator. Fortunately, an additional restriction is known for a very important set of operators in any gauge theory: conserved currents associated to global symmetries, such as the $SO(6)$ symmetry in the case of the $\mathcal{N} = 4$ SYM theory. The source $a_\mu(x)$ coupling to a conserved current $J^\mu(x)$ as

$$\int d^4x\, a_\mu(x) J^\mu(x) \tag{5.15}$$

may be thought of as an external background gauge field, and we can view it as the boundary value of a dynamical gauge field $A_M(x, z)$ in AdS, i.e.

$$a_\mu(x) = \lim_{z\to 0} A_\mu(z, x)\,, \tag{5.16}$$

meaning that, in the notation of (5.14), a gauge field has $\alpha_A = 0$. The identification (5.16) is natural given that, as we discussed in Section 5.1.3, continuous global symmetries in the boundary theory should correspond to large gauge transformations in the bulk. This identification will be confirmed below by examining the asymptotic behavior of A_μ near the boundary, see (5.32) and the discussion around it.

An especially important set of conserved currents in any translationally invariant theory are those encapsulated in the energy–momentum tensor operator $T^{\mu\nu}(x)$. The source $h_{\mu\nu}(x)$ coupling to $T^{\mu\nu}(x)$ as

$$\int d^4x\, h_{\mu\nu}(x) T^{\mu\nu}(x) \tag{5.17}$$

can be interpreted as a deformation of the boundary spacetime metric. In the absence of any such boundary metric deformation, from (5.3) we see that the asymptotic AdS bulk metric $g_{\mu\nu}$ and the boundary Minkowski metric are related by

$$g_{\mu\nu}(z, x) \to \frac{R^2}{z^2}\,\eta_{\mu\nu}, \qquad z \to 0\,. \tag{5.18}$$

In the presence of a boundary metric deformation $h_{\mu\nu}$ it is thus natural to relate the full boundary metric $g^{(b)}_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ to the bulk metric as

$$g^{(b)}_{\mu\nu}(x) = \lim_{z\to 0} \frac{z^2}{R^2}\, g_{\mu\nu}(z, x)\,, \tag{5.19}$$

meaning that for the metric $\alpha_g = 2$. The relation (5.19) should also be valid for $h_{\mu\nu}$ which is not infinitesimal, i.e. for a general curved boundary metric. Given (5.17),

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

the identification (5.19) has important implications: the dual of a translationally invariant gauge theory, in which the energy–momentum tensor is conserved, must involve gravity.

5.1.5 Normalizable vs. non-normalizable modes and mass–dimension relation

Having motivated the field/operator correspondence, we now elaborate on two important aspects of this correspondence: how the conformal dimension of an operator is related to properties of the dual bulk field [392, 803], and how to interpret normalizable and non-normalizable modes of a bulk field in the boundary theory [113, 114].

For illustration we will consider a massive bulk scalar field Φ, dual to some scalar operator $\mathcal{O}$ in the boundary theory. Although our main interest is the case in which the boundary theory is four-dimensional, it is convenient to present the equations for a general boundary spacetime dimension d. For this reason we will work with a generalization of the AdS metric (5.3) in which $x^\mu = (t, \vec{x})$ denote coordinates of a d-dimensional Minkowski spacetime.

The bulk action for Φ can be written as

$$S = -\frac{1}{2}\int dz\, d^d x \sqrt{-g}\, \left[g^{MN}\partial_M \Phi \partial_N \Phi + m^2\Phi^2\right] + \cdots \,. \tag{5.20}$$

We have canonically normalized Φ, and the dots stand for terms of order higher than quadratic. We have omitted these terms because they are proportional to positive powers of Newton's constant, and are therefore suppressed by positive powers of $1/N_c$.

Since the bulk spacetime is translationally invariant along the x^μ-directions, it is convenient to introduce a Fourier decomposition in these directions by writing[2]

$$\Phi(z, x^\mu) = \int \frac{d^d k}{(2\pi)^d}\, e^{ik\cdot x}\, \Phi(z, k^\mu)\,, \tag{5.21}$$

where $k \cdot x \equiv \eta_{\mu\nu} k^\mu x^\mu$ and $k^\mu \equiv (\omega, \vec{k})$, with ω and $\vec{k}$ the energy and the spatial momentum, respectively. In terms of these Fourier modes the equation of motion for Φ derived from the action (5.20) is

$$z^{d+1}\partial_z(z^{1-d}\partial_z \Phi) - k^2 z^2 \Phi - m^2 R^2 \Phi = 0\,, \qquad k^2 = -\omega^2 + \vec{k}^2\,. \tag{5.22}$$

Near the boundary $z \to 0$, the above equation can be readily solved perturbatively in z to obtain the asymptotic behavior:

$$\Phi(z, k) \approx A(k)\, \left(z^{d-\Delta} + \cdots\right) + B(k)\, \left(z^{\Delta} + \cdots\right) \qquad \text{as } z \to 0\,, \tag{5.23}$$

[2] For notational simplicity we will use the same symbol to denote a function and its Fourier transform, distinguishing them only through their arguments.

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

where

$$\Delta = \frac{d}{2} + \nu\,, \qquad \nu = \sqrt{m^2R^2 + \frac{d^2}{4}}\,. \tag{5.24}$$

In (5.23), "$+\cdots$" denotes subleading terms in each of the two linearly independent solutions. In subsequent equations we shall continue to suppress subleading terms, displaying only the leading term for each linearly independent solution and not even writing the "$+\cdots$". Since k enters (5.22) as a parameter, the integration "constants" A and B in general depend on k.

Fourier transforming (5.23) back into coordinate space, we then find

$$\Phi(z,x) \approx A(x)\,z^{d-\Delta} + B(x)\,z^{\Delta} \qquad \text{as } z \to 0\,. \tag{5.25}$$

The exponents in (5.25) are real provided

$$m^2R^2 \geq -\frac{d^2}{4}\,. \tag{5.26}$$

In fact, one can show that the theory is stable for any m^2 in the range (5.26), whereas for $m^2R^2 < -d^2/4$ there exist modes that grow exponentially in time and the theory is unstable [194, 195, 617]. In other words, in AdS space a field with a negative mass-squared does not lead to an instability provided the mass-squared is not "too negative". Equation (5.26) is often called the Breitenlohner–Freedman (BF) bound.

In the stable region (5.26) one must still distinguish between the finite interval $-d^2/4 \leq m^2R^2 < -d^2/4+1$ and the rest of the region, $m^2R^2 \geq -d^2/4+1$. In the first case both terms in (5.25) are normalizable with respect to the inner product

$$(\Phi_1,\Phi_2) = -i\int_{\Sigma_t} dz d\vec{x}\,\sqrt{-g}\,g^{tt}(\Phi_1^*\partial_t\Phi_2 - \Phi_2\partial_t\Phi_1^*)\,, \tag{5.27}$$

where Σ_t is a constant-t slice. We will comment on this case at the end of this section.

For the moment let us assume that $m^2R^2 \geq -d^2/4+1$. In this case the first term in (5.25) is non-normalizable and the second term, which is normalizable, does not affect the leading boundary behavior. As motivated in the previous section, the boundary value of a bulk field Φ should be identified with the source for the corresponding boundary operator $\mathcal{O}$. Since in (5.25) the boundary behavior of Φ is controlled by $A(x)$, the presence of such a non-normalizable term should correspond to a deformation of the boundary theory of the form

$$S_{\text{bdry}} \to S_{\text{bdry}} + \int d^dx\,\phi(x)\mathcal{O}(x)\,, \qquad \text{with } \phi(x) = A(x)\,. \tag{5.28}$$

In other words, *the non-normalizable term determines the boundary theory Lagrangian*. In particular, we see that in order to obtain a finite source $\phi(x)$ for

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

a scalar operator $\mathcal{O}(x)$ which is dual to an AdS scalar field $\Phi(x, z)$ of mass m, with m related to Δ through (5.24), we need to make the identification

$$\phi(x) = \Phi|_{\partial \mathrm{AdS}}(x) \equiv \lim_{z \to 0} z^{\Delta - d} \Phi(z, x) , \tag{5.29}$$

which is (5.14) with the choice $\alpha_\Phi = \Delta - d$.

In contrast, the normalizable modes are elements of the bulk Hilbert space. More explicitly, in the canonical quantization one expands Φ in terms of a basis of normalizable solutions of (5.22), from which one can then build the Fock space and compute the bulk Green's functions, etc. The equivalence between the bulk and boundary theories implies that their respective Hilbert spaces should be identified. Thus we conclude that *normalizable modes should be identified with states of the boundary theory*. This identification gives an important tool for finding the spectrum of low energy excitations of a strongly coupled gauge theory. In the particular example at hand, one can readily see from (5.22) that, for a given $\vec{k}$, there is a continuous spectrum of ω, consistent with the fact that the boundary theory is scale invariant.

Furthermore, as will be discussed in Section 5.3 (and in Appendix C), the coefficient $B(x)$ of the normalizable term in (5.25) can be identified with the expectation value of $\mathcal{O}$ in the presence of the source $\phi(x) = A(x)$, namely

$$\langle \mathcal{O}(x) \rangle_\phi = 2\nu B(x) . \tag{5.30}$$

In the particular case of a purely normalizable solution, i.e. one with $A(x) = 0$, this equation yields the expectation value of the operator in the undeformed theory.

Equations (5.25), (5.28) and (5.30) imply that Δ, introduced in (5.24), should be identified as the conformal dimension of the boundary operator $\mathcal{O}$ dual to Φ [803]. Indeed, recall that a scale transformation of the boundary coordinates $x^\mu \to C x^\mu$ corresponds to the isometry $x^\mu \to C x^\mu, z \to C z$ in the bulk. Since Φ is a scalar field, under such an isometry it transforms as $\Phi'(Cz, Cx^\mu) = \Phi(z, x^\mu)$, which implies that the corresponding functions in the asymptotic form (5.25) must transform as $A'(Cx^\mu) = C^{\Delta - d} A(x^\mu)$ and $B'(Cx^\mu) = C^{-\Delta} B(x^\mu)$. This means that $A(x)$ and $B(x)$ have mass scaling dimensions $d - \Delta$ and Δ, respectively. Eqs. (5.28) and (5.30) are then consistent with each other and imply that $\mathcal{O}(x)$ has mass scaling dimension Δ.

The mass–dimension relations (5.24), the near-boundary behavior (5.25), and the identification (5.29) are modified for fields of nonzero spin. For example, for a massive vector field whose bulk action is given by the Maxwell action plus a mass term, one finds,

$$\Delta = \frac{d}{2} + \sqrt{\frac{(d-2)^2}{4} + m^2 R^2} . \tag{5.31}$$

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

A gauge field A_M has $m^2 = 0$, which means $\Delta = d-1$ as expected for a conserved boundary current, see Section 5.1.4. By an analysis similar to that discussed above for the case of a scalar field, it can be shown that, near the boundary the vector field, A_M has the asymptotic behavior

$$A_\mu = a_\mu + b_\mu z^{d-2}, \qquad \text{as } z \to 0 \tag{5.32}$$

confirming the identification (5.16).

For the metric (a massless spin-two field), analysis of the Einstein equations leads to $\Delta = d$, as expected for the stress–energy tensor. An intuitive way to understand this is to note that just as the transverse traceless part of the graviton behaves like a massless scalar field in Minkowski space, in AdS space the transverse traceless part of a metric fluctuation behaves like a minimally coupled massless scalar, as we discuss further in Section 6.2.2. More explicitly, when a transverse traceless metric perturbation is written with one upper and one lower index, it satisfies the same equation that a massless scalar field does. We then note from (5.19) that any metric perturbation with one upper and one lower index tends to a finite limit upon approaching the boundary, just like a massless scalar field. Consequently, the component of the boundary theory stress tensor that is dual to the transverse traceless part of the bulk metric has scaling dimension d. By covariance this means that all components of the boundary theory stress tensor scale in this way. Note that upon applying the scaling argument below Eq. (5.30) to Eq. (5.19), one finds that $g^{(b)}_{\mu\nu}$ and thus $h_{\mu\nu}$ does not scale under a scaling transformation, which then gives the correct scaling dimension for $T^{\mu\nu}$. This provides a quick consistency check of (5.19).

Before closing this section, let us return to the range $-d^2/4 \le m^2 < -d^2/4+1$. We shall be brief because this is not a case that arises in later sections. Since in this case both terms in (5.25) are normalizable, either one can be used to build the Fock space of physical states of the theory [194, 195]. This gives rise to two different boundary CFTs in which the dimensions of the operator $\mathcal{O}(x)$ are Δ or $d-\Delta$, respectively [538]. It was later realized [805, 142] that even more general choices are possible in which the modes used to build the physical states have both A and B, nonzero. These choices correspond to different quantizations from the bulk viewpoint, and to deformations by double-trace operators from the gauge theory viewpoint.

5.2 Generalizations

5.2.1 Nonzero temperature and nonzero chemical potential

As discussed in Section 4.3, the same string theory reasoning giving rise to the equivalence (4.27) can be generalized to nonzero temperature by replacing the pure AdS metric (5.2) by that of a black brane in AdS_5 [804], Eq. (4.30), which we repeat here for convenience:

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

$$ds^2 = \frac{r^2}{R^2}\left(-f dt^2 + d\vec{x}^2\right) + \frac{R^2}{r^2 f} dr^2, \quad f(r) = 1 - \frac{r_0^4}{r^4}. \tag{5.33}$$

Equivalently, in terms of the z-coordinate, we replace Eq. (5.2) by Eq. (4.32), i.e.

$$ds^2 = \frac{R^2}{z^2}\left(-f dt^2 + d\vec{x}^2\right) + \frac{R^2}{z^2 f} dz^2, \quad f(z) = 1 - \frac{z^4}{z_0^4}. \tag{5.34}$$

The metrics above have an event horizon at $r = r_0$ and $z = z_0$, respectively, and the regions outside the horizon correspond to $r \in (r_0, \infty)$ and $z \in (0, z_0)$. This generalization can also be directly deduced from (4.27) as the black brane (5.33)–(5.34) is the only metric on the gravity side that satisfies the following properties: (i) it is asymptotically AdS_5; (ii) it is translationally invariant along all the boundary directions and rotationally-invariant along the boundary spatial directions; (iii) it has a temperature and satisfies all laws of thermodynamics. It is therefore natural to identify the temperature and other thermodynamical properties of (5.33)–(5.34) with those of the SYM theory at nonzero temperature.

We mention in passing that there is also a nice connection between the black brane geometry (5.33)–(5.34) and the thermal-field formulation of finite-temperature field theory in terms of real time. Indeed, the fully extended spacetime of the black brane has two boundaries. Each of them supports an identical copy of the boundary field theory which can be identified with one of the two copies of the field theory in the Schwinger–Keldysh formulation. The thermal state can also be considered as a specific entangled state of the two field theories. For more details see [593, 451].

The Hawking temperature of the black brane can be calculated via the standard method [376] (see Appendix B for details) of demanding that the Euclidean continuation of the metric (5.34) obtained by the replacement $t \to -i t_E$,

$$ds_E^2 = \frac{R^2}{z^2}\left(f dt_E^2 + dx_1^2 + dx_2^2 + dx_3^2\right) + \frac{R^2}{z^2 f} dz^2, \tag{5.35}$$

be regular at $z = z_0$. This requires that t_E be periodically identified with a period β given by

$$\beta = \frac{1}{T} = \pi z_0. \tag{5.36}$$

The temperature T is identified with the temperature of the boundary SYM theory, since t_E corresponds precisely to the Euclidean time coordinate of the boundary theory. We emphasize here that while the Lorentzian spacetime (5.34) can be extended beyond the horizon $z = z_0$, the Euclidean metric (5.35) exists only for $z \in (0, z_0)$ as the spacetime ends at $z = z_0$, and ends smoothly once the choice (5.36) is made.

For a boundary theory with a $U(1)$ global symmetry, like $\mathcal{N} = 4$ SYM theory, one can furthermore turn on a chemical potential μ for the corresponding $U(1)$ charge. From the discussion of Section 5.1.4, this requires that the bulk gauge field A_μ which is dual to a boundary current J_μ satisfies the boundary condition

$$\lim_{z\to 0} A_t = \mu \, . \tag{5.37}$$

The above condition along with the requirement that the field A_μ should be regular at the horizon implies that there should be a radial electric field in the bulk, i.e. the black hole is now charged. We will not write the metric of a charged black hole explicitly, as we will not use it in this book. For more details and its applications, see e.g. [261, 262, 393, 228, 302, 303, 426, 483]. Similarly, in the case of theories with fundamental flavor introduced as probe D-branes, a baryon number chemical potential corresponds to an electric field on the branes [529, 461, 539, 606, 643, 516, 809, 141, 309, 718, 644, 375, 515].

5.2.2 A confining theory

Although our main interest is the deconfined phase of QCD, in this section we will briefly describe a simple example of a duality for which the field theory possesses a confining phase [804]. For simplicity we have chosen a model in which the field theory is three-dimensional, but all the essential features of this model extend to the string duals of more realistic confining theories in four dimensions.

We start by considering $\mathcal{N} = 4$ SYM theory at finite temperature. In the Euclidean description the system lives on $\mathbb{R}^3 \times S^1$. The circle direction corresponds to the Euclidean time, which is periodically identified with period $\beta = 1/T$. As is well known, at length scales much larger than β one can effectively think of this theory as the Euclidean version of pure three-dimensional Yang–Mills theory. The reasoning is that at these scales one can perform a Kaluza–Klein reduction along the circle. Since the fermions of the $\mathcal{N} = 4$ theory obey antiperiodic boundary conditions around the circle, their zero-mode is projected out, which means that all fermionic modes acquire a tree-level mass of order $1/\beta$. The scalars of the $\mathcal{N} = 4$ theory are periodic around the circle, but they acquire masses at the quantum level. The only fields that cannot acquire masses are the gauge bosons of the $\mathcal{N} = 4$ theory, since masses for them are forbidden by gauge invariance. Thus, at long distances the theory reduces to a pure Yang–Mills theory in three dimensions, which is confining and has a map gap. The Lorentzian version of the theory is simply obtained by analytically continuing one of the $\mathbb{R}^3$ directions into the Lorentzian time. Thus, in this construction the "finite temperature" of the original four-dimensional theory is a purely theoretical device. The effective Lorentzian theory in three dimensions is at zero temperature.

In order to obtain the gravity description of this theory we just need to implement the above procedure on the gravity side. We start with the Lorentzian metric (5.1)–(5.2) dual to $\mathcal{N} = 4$ SYM at zero temperature. Then we introduce a nonzero temperature by going to Euclidean signature via $t \to -it_{\mathrm{E}}$ and periodically identifying the Euclidean time. This results in the metric (5.35). Finally, we analytically continue one of the $\mathbb{R}^3$ directions, say x_3, back into the new Lorentzian time: $x_3 \to it$. The final result is the metric

$$ds^2 = \frac{R^2}{z^2}\left(-dt^2 + dx_1^2 + dx_2^2 + f\,dt_{\mathrm{E}}^2\right) + \frac{R^2}{z^2 f}dz^2\,. \tag{5.38}$$

In this metric the directions t, x_1, x_2 correspond to the directions in which the effective three-dimensional Yang–Mills theory lives. The direction t_{E} is now a compact spatial direction. Note that since the original metric (5.35) smoothly ends at $z = z_0$, so does (5.38). This leads to a dramatic difference between the gauge theory dual to (5.38) and the original $\mathcal{N} = 4$ theory: the fact that the radial direction smoothly closes off at $z = z_0$ introduces a mass scale in the boundary theory. To see this, note that the warp factor R^2/z^2 has a lower bound. Thus, when applying the discussion of Section 5.1.1 to (5.38), E_{YM} in Eq. (5.4) will have a lower limit of order $M \sim 1/z_0$, implying that the theory develops a mass gap of this order. This can also be explicitly verified by solving the equation of motion of a classical bulk field (which is dual to some boundary theory operator) in the metric (5.38): for any fixed $\vec{k}$ one finds a discrete spectrum of normalizable modes with a mass gap of order M. (Note that since the size of the circle parametrized by t_{E} is proportional to $1/z_0$, the mass gap is in fact comparable to the energies of Kaluza–Klein excitations on the circle.) As explained in Section 5.1.5, these normalizable modes can be identified with the glueball states of the boundary theory.

The fact that the gauge theory dual to the geometry (5.38) is a confining theory is further supported by several checks, including the following two. First, analysis of the expectation value of a Wilson loop reveals an area law, as will be discussed in Section 5.4. Second, the gravitational description can be used to establish that the theory described by (5.38) undergoes a deconfinement phase transition at a temperature $T_c \sim M$ set by the mass gap, above which the theory is again described by a geometry with a black hole horizon [804] (see [601, 672] for reviews).

The above construction resulted in an effective confining theory in three dimensions because we started with the theory on the worldvolume of D3-branes, which is a four-dimensional SYM theory. By starting instead with the near horizon solution of a large number of non-extremal D4-branes, which describes a SYM theory in five dimensions, the above procedure leads to the string dual of a Lorentzian confining theory that at long distance reduces to a four-dimensional pure Yang–Mills theory [804]. This has been used as the starting point of the Sakai–Sugimoto model for QCD [721, 722], which incorporates spontaneous chiral symmetry breaking and its restoration at high temperatures [30, 670]. For reviews on some of these topics see for example [601, 672].

5.2.3 Other generalizations

In addition to (4.27), many other examples of gauge/string dualities are known in different spacetime dimensions (see e.g. [29] and references therein). These include theories with fewer supersymmetries and theories which are not scale invariant, in particular confining theories [688, 537, 597] (see e.g. [27, 764] for reviews).

For a d-dimensional conformal theory, the dual geometry on the gravity side contains a factor of AdS_{d+1} and some other compact manifold.[3] When expanded in terms of the harmonics of the compact manifold, the duality again reduces to that between a d-dimensional conformal theory and a gravity theory in AdS_{d+1}. In particular, in the classical gravity limit, this reduces to Einstein gravity in AdS_{d+1} coupled to various matter fields with the precise spectrum of matter fields depending on the specific theory under consideration. For a nonconformal theory the dual geometry is in general more complicated. Some simple, early examples were discussed in [487]. If a theory has a mass gap, the dual bulk geometry either closes off at some finite value of z_0 as in the example of Section 5.2.2 or ends in some IR singularity that can be reached in a finite proper distance.

All known examples of gauge/string dual pairs share the following common features with (4.27): (i) the field theory is described by elementary bosons and fermions coupled to non-Abelian gauge fields whose gauge group is specified by some N_c; (ii) the string description reduces to classical (super)gravity in the large-N_c, strong coupling limit of the field theory. In this book we will use (4.27) as our prime example for illustration purposes, but the discussion can be immediately applied to other examples including nonconformal ones.

5.3 Correlation functions of local operators

In this section we will explain how to calculate correlation functions of local gauge-invariant operators of the boundary theory in terms of the dual gravity description. We will mostly focus on one-point and two-point functions, in the latter case in particular on real time retarded correlators which are important for determining linear response, transport coefficients, and spectral functions. We will, however, begin by describing the general prescription for computing n-point Euclidean correlation functions.

[3] Not necessarily in a direct product; the product may be warped.

5.3.1 General prescription for Euclidean correlators

In view of the field/operator correspondence discussed in Sections 5.1.4 and 5.1.5, it is natural to postulate that the Euclidean partition functions of the boundary and bulk theories must agree, namely that [392, 803]

$$Z_{\rm CFT}\left[\phi(x)\right] = Z_{\rm string}\left[\Phi|_{\partial {\rm AdS}}(x)\right] . \tag{5.39}$$

Both sides of this equation require explanation. The left-hand side of (5.39) is the most general partition function in the CFT, including a source for each gauge-invariant operator in the theory, namely

$$Z_{\rm CFT}\left[\phi(x)\right] \equiv \langle e^{\int \phi \mathcal{O}} \rangle . \tag{5.40}$$

Here one should think of $\phi(x)$ in Eq. (5.39) as succinctly indicating the collection of all such sources. The expectation value $\langle \cdots \rangle$ can be in the vacuum or a thermal state. Since AdS has a boundary, to define the string theory partition function on the right-hand side of (5.39) one needs to specify a boundary condition for each bulk field. The collection of all such boundary conditions is indicated by $\Phi|_{\partial {\rm AdS}}(x)$ in Eq. (5.39). Having parsed both sides of it, the equality in (5.39) makes sense because both sides of the equation are functionals of the same variables upon the identification of ϕ and $\Phi|_{\partial {\rm AdS}}(x)$ in (5.14).

The right-hand side of (5.39) is in general not easy to compute, but it simplifies dramatically in the classical gravity limit (5.6), where it can be obtained using the saddle point approximation as

$$Z_{\rm string}[\phi] \simeq \exp\left(S^{\rm (ren)}[\Phi_c^{(E)}]\right) , \tag{5.41}$$

where we have absorbed a conventional minus sign into the definition of the Euclidean action which avoids having some additional minus signs in various equations below and in the analytic continuation to Lorentzian signature. In Eq. (5.41), $S^{\rm (ren)}[\Phi_c^{(E)}]$ is the renormalized on-shell classical supergravity action [450, 111, 635, 338, 558, 314, 158], namely the classical action evaluated on a Euclidean solution $\Phi_c^{(E)}$ of the classical equations of motion determined by the boundary identification with ϕ, i.e. the Euclidean version of (5.14), and by the requirement that the solution be regular everywhere in the interior of the spacetime. The on-shell action needs to be renormalized because it typically suffers from IR divergences due to the integration region near the boundary of AdS [803]. These divergences are dual to UV divergences in the gauge theory, consistent with the UV/IR correspondence. The procedure to remove these divergences on the gravity side is well understood and is referred to as "holographic renormalization". (It is no more similar to the holographic renormalization group that we mentioned in Section 5.1.1 than the renormalization group is to traditional renormalization.) Although holographic renormalization is an important ingredient of the gauge/string duality, it is also somewhat technical. In Appendix C we briefly review it in the context of a two-point function calculation. The interested reader may consult the literature cited above, as well as the review [742], for details.

From (5.39) and (5.41) we thus find that in the large-N_c and large-λ limit, the boundary theory free energy is given by

$$\log Z_{\text{CFT}}[\phi(x)] = S^{(\text{ren})}[\Phi_c^{(E)}] \,. \tag{5.42}$$

Corrections to Eq. (5.41) can be included as an expansion in α' and g_s, which correspond to $1/\sqrt{\lambda}$ and $1/N_c$ corrections in the gauge theory, respectively. Note that since the classical action (5.9) on the gravity side is proportional to $1/G_5$, from Eq. (5.12) we see that $S^{(\text{ren})}[\Phi_c^{(E)}] \sim N_c^2$, as one would expect for the generating functional of an $SU(N_c)$ SYM theory in the large-N_c limit. From (5.42), in the large-N_c and large-λ limit, connected correlation functions of the gauge theory are given simply by functional derivatives of the on-shell, classical gravity action:

$$\langle \mathcal{O}(x_1) \ldots \mathcal{O}(x_n) \rangle = \left. \frac{\delta^n S^{(\text{ren})}[\Phi_c^{(E)}]}{\delta\phi(x_1) \ldots \delta\phi(x_n)} \right|_{\phi=0} . \tag{5.43}$$

This concludes our general discussion of n-point functions. In Appendix C we give an explicit computation of the Euclidean two-point function for a scalar operator in a CFT. For some early work on the evaluation of higher-point functions see Refs. [358, 581, 260].

5.3.2 One-point functions

Here we describe how to compute the one-point function (i.e. expectation value) of an operator in a general time-dependent state which may not have a Euclidean analytic continuation. We first consider a generic scalar operator, and then turn more specifically to the stress tensor and a conserved current.

From (5.43), the Euclidean one point function of a scalar operator $\mathcal{O}$ in the presence of the source ϕ is given by

$$\langle \mathcal{O}(x) \rangle_\phi = \frac{\delta S^{(\text{ren})}[\Phi_c^{(E)}]}{\delta\phi(x)} = \lim_{z\to 0} z^{d-\Delta} \frac{\delta S^{(\text{ren})}[\Phi_c^{(E)}]}{\delta\Phi_c^{(E)}(z,x)} \,, \tag{5.44}$$

where in the second equality we have used (5.29). In classical mechanics, it is well known that the variation of the action with respect to the boundary value of a field results in the canonical momentum Π conjugate to the field, where the boundary in that case is usually a constant-time surface. (See, e.g., Ref. [566].) In the present case the boundary is a constant-z surface, but it is still useful to proceed by analogy with classical mechanics and to think of the derivative in the last term in (5.44) as

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

the renormalized canonical momentum conjugate to $\Phi_c^{(E)}$ evaluated on the classical solution:

$$\Pi_c^{(\text{ren})}(z,x) = \frac{\delta S^{(\text{ren})}[\Phi_c^{(E)}]}{\delta \Phi_c^{(E)}(z,x)}. \tag{5.45}$$

With this definition, Eq. (5.44) takes the form

$$\langle \mathcal{O}(x)\rangle_\phi = \lim_{z\to 0} z^{d-\Delta}\Pi_c^{(\text{ren})}(z,x) \tag{5.46}$$

which can further be shown to yield

$$\langle \mathcal{O}(x)\rangle_\phi = 2\nu B(x), \tag{5.47}$$

where we have used (5.25) and have identified $A(x)$ in (5.25) with $\phi(x)$. See Appendix C for a discussion. In the absence of a source, i.e. if $\phi(x) = A(x) = 0$, then (5.47) gives the expectation value of $\mathcal{O}$ in terms of the fall-off of a normalizable solution.

The prescription (5.46) or (5.47) requires only knowledge of the asymptotic boundary behavior of the bulk solution Φ_c and is thus much simpler to compute than (5.44). More importantly, the formulation (5.44) does not generalize to a generic time-dependent state which does not have a Euclidean analytic continuation, while the expressions (5.46) or (5.47) do have straightforward generalizations. Recall that a normalizable solution in the bulk is mapped to a state in the boundary. Evaluating (5.46) or (5.47) for such a bulk solution then gives the expectation value in the corresponding state on the boundary.

Let us now consider the one-point function of the stress–energy tensor which, upon making the identification (5.42), can be obtained from the expression

$$\langle T^{\mu\nu}(x)\rangle = \frac{2}{\sqrt{g^{(b)}(x)}}\frac{\delta S^{(\text{ren})}[g^{(b)}]}{\delta g^{(b)}_{\mu\nu}(x)} = \lim_{z\to 0}\frac{z^{d+2}}{R^{d+2}}\frac{2}{\sqrt{\det g_{\mu\nu}(x,z)}}\frac{\delta S^{(\text{ren})}[g]}{\delta g_{\mu\nu}(x,z)}, \tag{5.48}$$

where $g^{(b)}_{\mu\nu}$ is the metric for the boundary theory and where the various expressions should all be understood in Euclidean signature. The first equality follows from the standard field theory definition of the stress tensor and we have used (5.19) in the second equality. Note that in the last expression $\det g_{\mu\nu}$ is the determinant of $g_{\mu\nu}(x,z)$, which is the part of the bulk metric along boundary directions or, equivalently, the induced metric on a constant-z hypersurface.

As in the scalar case, the variation of the bulk on-shell action with respect to the boundary value of $g_{\mu\nu}$ is given by the canonical momentum $\Pi^{\mu\nu}$ conjugate to $g_{\mu\nu}$ evaluated on the classical solution:

$$\frac{\delta S^{(\text{ren})}}{\delta g_{\mu\nu}} = \Pi^{\mu\nu}_{(\text{ren})} = \frac{\sqrt{\det g_{\mu\nu}}}{16\pi G_N}(K^{\mu\nu} - g^{\mu\nu}K) + \frac{\delta S^{(\text{ct})}[g]}{\delta g_{\mu\nu}(x,z)} \tag{5.49}$$

with $K_{\mu\nu}$ the extrinsic curvature for a constant-z hypersurface. In (5.49), the first term is the standard canonical momentum in general relativity, while $S^{(\mathrm{ct})}[g]$ is the counterterm that must be added to the action in order to make the total action finite. $S^{(\mathrm{ct})}[g]$ is dimension-dependent for a general curved boundary metric $g^{(b)}$, but has a universal form when the boundary metric is flat, in which case one has [111]

$$S^{(\mathrm{ct})} = -\frac{1}{8\pi G_N}\frac{d-1}{R}\int_{z\to 0} d^d x\sqrt{\det g_{\mu\nu}}\ , \tag{5.50}$$

where the integral is over a constant-z slice. From (5.48)–(5.50) we thus find that, if the boundary theory has a flat metric,

$$\langle T^{\mu\nu}\rangle = \lim_{z\to 0}\frac{1}{8\pi G_N}\frac{R^{d+2}}{z^{d+2}}\left(K^{\mu\nu} - g^{\mu\nu}K - \frac{d-1}{R}g^{\mu\nu}\right)\ . \tag{5.51}$$

As discussed above in the scalar case, the expression (5.51) can be applied to a general bulk Lorentzian geometry to find the expectation value of the stress tensor in the corresponding dual state. In particular, it applies to non-equilibrium states. Equation (5.51) will play an important role in Chapter 7 and in Section 8.3.

Finally we briefly mention the prescription for extracting the expectation value of a conserved current j^μ in a boundary state dual to some given bulk state. Suppose the corresponding bulk gauge field A_M has the Maxwell action

$$S = -\frac{1}{4}\int dz\, d^d x\,\sqrt{-g}\,F_{MN}F^{MN}\ . \tag{5.52}$$

The canonical momentum conjugate to A_μ is $\Pi^\mu = -\sqrt{-g}F^{z\mu}$ and (5.46)–(5.47) then generalize to

$$\langle j^\mu\rangle = -\lim_{z\to 0}\sqrt{-g}F^{z\mu} = -(d-2)R^{d-3}\eta^{\mu\nu}b_\mu \tag{5.53}$$

where b_μ is the coefficient of the normalizable term in (5.32) and $\eta^{\mu\nu}$ is the boundary Minkowski metric.

5.3.3 Real time two-point functions

We now proceed to the prescription for calculating real time correlation functions *in equilibrium*. We will focus our discussion on *retarded* two-point functions because of their important role in characterizing linear response. Also, once the retarded function is known one can then use standard relations to obtain the other Green's functions. The calculation of two-point functions out of equilibrium does not yield closed form expressions like those we shall find in equilibrium below, and we shall not present it here. However, for an out-of-equilibrium formulation that that yields explicit expressions suitable for numerical evaluation, see Ref. [241].

We start with linear response in Euclidean signature. In momentum space the response, i.e the expectation value of an operator, is proportional to the corresponding source, and the constant of proportionality (for each momentum) is the two-point function of the operator:

$$\langle \mathcal{O}(\omega_E, \vec{k}) \rangle_\phi = G_E(\omega_E, \vec{k}) \phi(\omega_E, \vec{k}) \,. \tag{5.54}$$

Then, Eqs. (5.46)–(5.47) yield

$$G_E(\omega_E, \vec{k}) = \frac{\langle \mathcal{O}(\omega_E, \vec{k}) \rangle_\phi}{\phi(\omega_E, \vec{k})} = \lim_{z \to 0} z^{2(d-\Delta)} \frac{\Pi_c^{(\mathrm{ren})}}{\Phi_c^{(E)}} = 2\nu \frac{B(\omega_E, \vec{k})}{A(\omega_E, \vec{k})} \,, \tag{5.55}$$

where ω_E denotes Euclidean frequency. See Appendix C for further discussion.

If the Euclidean correlation functions G_E are known exactly, the retarded functions G_R can then be obtained via the analytic continuation

$$G_R(\omega, \vec{k}) = G_E(-i(\omega + i\epsilon), \vec{k}) \,. \tag{5.56}$$

In most examples of interest, however, the Euclidean correlation functions can only be found numerically and analytic continuation to Lorentzian signature becomes difficult. Thus, it is important to develop techniques to calculate real time correlation functions directly. Based on an educated guess that passed several consistency checks, a prescription for calculating retarded two-point functions in Lorentzian signature was first proposed by Son and Starinets in Ref. [747]. The authors of Ref. [451] later justified the prescription and extended it to n-point functions. Here we will follow the treatment given in Refs. [482, 481]. For illustration we consider the retarded two-point function for a scalar operator $\mathcal{O}$ at nonzero temperature, which can be obtained from the propagation of the dual scalar field Φ in the geometry of an AdS black brane. The action for Φ again takes the form (5.20) with g_{MN} now given by the black brane metric (5.34).

Before giving the prescription, we note that in Lorentzian signature one cannot directly apply the procedure summarized by Eq. (5.43) to obtain retarded functions. There are two immediate complications/difficulties. First, the Lorentzian black hole spacetime contains an event horizon and one also needs to impose appropriate boundary conditions there when solving the classical equation of motion for Φ. Second, since partition functions are defined in terms of path integrals, the resulting correlation functions should be time ordered.[4] As we now describe, both complications can be dealt with in a simple manner.

[4] While it is possible to obtain Feynman functions this way, the procedure is quite subtle, since Feynman functions require imposing different boundary conditions for positive- and negative-frequency modes at the horizon and the choices of positive-frequency modes are not unique in a black hole spacetime. The correct choice corresponds to specifying the so-called Hartle–Hawking vacuum. For details see Ref. [451]. In contrast, the retarded function does not depend on the choice of the bulk vacuum in the classical limit as the corresponding bulk retarded function is given by the commutator of the corresponding bulk field – see Eq. (5.66).

The idea is to analytically continue the Euclidean classical solution that we have denoted $\Phi_c^{(E)}(\omega_E, \vec{k})$, as well as Eq. (5.55), to Lorentzian signature according to (5.56). Clearly the analytic continuation of $\Phi_c^{(E)}(\omega_E, \vec{k})$,

$$\Phi_c(\omega, \vec{k}) = \Phi_c^{(E)}(-i(\omega + i\epsilon), \vec{k}) , \tag{5.57}$$

solves the Lorentzian equation of motion. In addition, this solution obeys the infalling boundary condition at the future event horizon of the black brane metric (5.34). This property is important as it ensures that the retarded correlator is causal and only propagates information forward in time. This is intuitive since we expect that, classically, information can fall into the black hole horizon but not come out, so the retarded correlator should have no outgoing component. Although it is intuitive, given its importance let us briefly verify that the infalling boundary condition is satisfied. The Lorentzian equation of motion in momentum space for Φ_c in the black brane metric (5.34) takes the form

$$z^5 \partial_z \Big[z^{-3} f(z) \partial_z \Phi \Big] + \frac{\omega^2 z^2}{f(z)} \Phi - \vec{k}^2 z^2 \Phi - m^2 R^2 \Phi = 0 , \tag{5.58}$$

where $\vec{k}^2 = \delta^{ij} k_i k_j$. The corresponding Euclidean equation is obtained by setting $\omega = i\omega_E$. Near the horizon $z \to z_0$, since $f \to 0$ the last two terms in (5.58) become negligible compared with the second term and can be dropped. The resulting equation (with only the first two terms of (5.58)) then takes the simple form

$$\begin{aligned} \text{Lorentzian :} \quad & \partial_\xi^2 \Phi + \omega^2 \Phi = 0 , \\ \text{Euclidean :} \quad & \partial_\xi^2 \Phi - \omega_E^2 \Phi = 0 , \end{aligned} \tag{5.59}$$

in terms of a new coordinate

$$\xi \equiv \int^z \frac{dz'}{f(z')} . \tag{5.60}$$

Since $\xi \to +\infty$ as $z \to z_0$, in order for the Euclidean solution to be regular at the horizon we must choose the solution with the decaying exponential, i.e. $\Phi_c^{(E)}(\omega_E, \xi) \sim e^{-\omega_E \xi}$. The prescription (5.57) then yields $\Phi_c(\omega, \xi) \sim e^{i\omega\xi}$. Going back to coordinate space we find that near the horizon

$$\Phi_c(t, \xi) \sim e^{-i\omega(t-\xi)} . \tag{5.61}$$

As anticipated, this describes a wave propagating towards the direction in which ξ increases, i.e. falling into the horizon. Had we chosen the opposite sign in the prescription (5.57) we would have obtained an outgoing wave, as appropriate for the advanced correlator which obeys an outgoing boundary condition at the past event horizon of the metric (5.34), and has no infalling component.

Given a Lorentzian solution satisfying the infalling boundary condition at the horizon which can be expanded near the boundary according to (5.23), as emphasized below (5.47), Eqs. (5.46)–(5.47) can then be applied directly to such a Lorentzian solution, yielding the Lorentzian counterpart of (5.55):

$$G_R(\omega, \vec{k}) = \lim_{z \to 0} z^{2(d-\Delta)} \frac{\Pi_c^{(\mathrm{ren})}}{\Phi_c(\omega, \vec{k})} = 2\nu \frac{B(\omega, \vec{k})}{A(\omega, \vec{k})} \,. \tag{5.62}$$

Incidentally, Eq. (5.62) shows that the retarded correlator possesses a pole precisely at those frequencies for which $A(\omega, \vec{k})$ vanishes. In other words, the poles of the retarded two-point function are in one-to-one correspondence with normalizable solutions of the equations of motion which are infalling at the horizon. Owing to the infalling boundary conditions at the horizon, such modes have a discrete spectrum and their frequencies have strictly negative imaginary parts. In the gravity literature such modes are referred to as quasinormal modes. In the field theory context, poles of retarded Green functions encode much of the physics of a system including the presence of hydrodynamic modes, the way in which out-of-equilibrium states relax toward equilibrium and the presence of quasiparticles, if any. We will return to this discussion at length in the context of strongly coupled $\mathcal{N} = 4$ SYM theory in Chapter 6 and in particular in Section 6.4.

For practical purposes, let us recapitulate here the main result of this section, namely the algorithmic procedure for computing the real time, finite-temperature retarded two-point function of a local, gauge-invariant operator $\mathcal{O}(x)$. This consists of the following steps.

(1) Identify the bulk mode $\Phi(x, z)$ dual to $\mathcal{O}(x)$.

(2) Find the Lorentzian-signature bulk effective action for Φ to quadratic order, and the corresponding linearized equation of motion in momentum space.

(3) Find a solution $\Phi_c(k, z)$ to this equation with the boundary conditions that the solution is infalling at the horizon and behaves as

$$\Phi_c(z, k) \approx A(k)\, z^{d-\Delta} + B(k)\, z^{\Delta} \tag{5.63}$$

near the boundary ($z \to 0$), where Δ is the dimension of $\mathcal{O}(x)$, d is the spacetime dimension of the boundary theory and $A(k)$ should be thought of as an arbitrary source for $\mathcal{O}(k)$. $B(k)$ is not an independent quantity but is determined by the boundary condition at the horizon and $A(k)$.

(4) The retarded Green's function for $\mathcal{O}$ is then given by

$$G_R(k) = 2\nu \frac{B(k)}{A(k)} \,, \tag{5.64}$$

where ν is defined in Eq. (5.24).

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

In Section 9.5.2 we will discuss in detail an example of a retarded correlator of two electromagnetic currents.

Before closing this section, we note that an alternative way to compute boundary correlation functions which works in both Euclidean and Lorentzian signature is [118]

$$\langle \mathcal{O}(x_1) \cdots \mathcal{O}(x_n) \rangle = \lim_{z_i \to 0} (2\nu z_1^{\Delta}) \cdots (2\nu z_n^{\Delta}) \langle \Phi(z_1, x_1) \cdots \Phi(z_n, x_n) \rangle \qquad (5.65)$$

where the correlator on the right-hand side is a correlation function in the bulk theory. In (5.65) it should be understood that whatever ordering one wants to consider, it should be same on both sides. For example, for the retarded two-point function G_R of $\mathcal{O}$

$$G_R(x_1 - x_2) = \lim_{z_1, z_2 \to 0} (2\nu z_1^{\Delta})(2\nu z_2^{\Delta}) \, \mathcal{G}_R(z_1, x_1; z_2, x_2) \, , \qquad (5.66)$$

where $\mathcal{G}_R$ denotes the retarded Green's function of the bulk field Φ.

5.4 Wilson loops

The expectation values of Wilson loops

$$W^r(\mathcal{C}) = \mathrm{Tr}\, \mathcal{P} \exp\left[i \int_{\mathcal{C}} dx^{\mu} \, A_{\mu}(x) \right] , \qquad (5.67)$$

are an important class of non-local observables in any gauge theory. Here, $\int_{\mathcal{C}}$ denotes a line integral along the closed path $\mathcal{C}$, $W^r(\mathcal{C})$ is the trace of an $SU(N)$-matrix in the representation r (one often considers fundamental or adjoint representations, i.e. $r = F, A$), the vector potential $A_{\mu}(x) = A^a_{\mu}(x) \, T^a$ can be expressed in terms of the generators T^a of the corresponding representation, and $\mathcal{P}$ denotes path ordering. The expectation values of Wilson loops contain information about the nonperturbative physics of non-Abelian gauge field theories and have applications to many physical phenomena such as confinement, thermal phase transitions, quark screening, etc. For many of these applications it is useful to think of the path $\mathcal{C}$ as that traversed by a quark. We will discuss some of these applications in Chapter 8. Here, we describe how to compute expectation values of Wilson loops in a strongly coupled gauge theory using its gravity description.

We again use $\mathcal{N} = 4$ SYM theory as an example. Now recall that the field content of this theory includes six scalar fields $\vec{\phi} = (\phi^1, \ldots \phi^6)$ in the adjoint representation of the gauge group. This means that in this theory one can write down the following generalization of (5.67) [595, 711]:

$$W(\mathcal{C}) = \frac{1}{N_c} \mathrm{Tr} P \exp\left[i \oint_{\mathcal{C}} ds \, \left(A_{\mu} \dot{x}^{\mu} + \vec{n} \cdot \vec{\phi} \, \sqrt{\dot{x}^2} \right) \right] , \qquad (5.68)$$

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

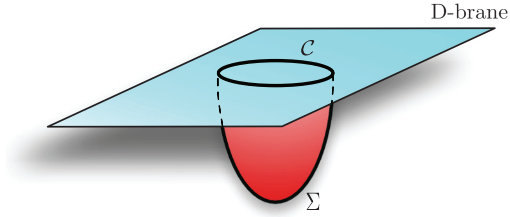

Figure 5.1 String worldsheet associated with a Wilson loop.

where $\vec{n}(s)$ is a unit vector in $\mathbb{R}^6$ that parametrizes a path in this space (or, more precisely, in S^5), just like $x^\mu(s)$ parametrizes a path in $\mathbb{R}^{(1,3)}$. The factor of $\sqrt{\dot{x}^2}$ is necessary to make $\vec{n} \cdot \vec{\phi}\sqrt{\dot{x}^2}$ a density under worldline reparametrizations. Note that the operators (5.67) and (5.68) are equivalent in the case of a light-like loop (as will be discussed in Section 8.5) for which $\dot{x}^2 = 0$.

An important difference between the operators (5.67) and (5.68) is that (5.67) breaks supersymmetry, whereas (5.68) is locally 1/2-supersymmetric, meaning that for a straight-line contour (that is time-like in Lorentzian signature) the operator is invariant under half of the supercharges of the $\mathcal{N} = 4$ theory.

We will now argue that the generalized operator (5.68) has a dual description in terms of a string worldsheet. For this purpose it is useful to think of the loop $\mathcal{C}$ as the path traversed by a quark. Although the $\mathcal{N} = 4$ SYM theory has no quarks, we will see below that these can be simply included by introducing in the gravity description open strings attached to a D-brane sitting at some radial position proportional to the quark mass. The endpoint of the open string on the D-brane is dual to the quark, so the boundary $\partial\Sigma$ of the string worldsheet Σ must coincide with the path $\mathcal{C}$ traversed by the quark – see Fig. 5.1. This suggests that we must identify the expectation value of the Wilson loop operator, which gives the partition function (or amplitude) of the quark traversing $\mathcal{C}$, with the partition function of the dual string worldsheet Σ [595, 711]:

$$\langle W(\mathcal{C})\rangle = Z_{\text{string}}[\partial\Sigma = \mathcal{C}]\,. \tag{5.69}$$

For simplicity, we will focus on the case of an infinitely heavy (non-dynamical) quark. This means that we imagine that we have pushed the D-brane all the way to the AdS boundary. Under these circumstances the boundary $\partial\Sigma = \mathcal{C}$ of the string worldsheet also lies within the boundary of AdS.

The key point to recall now is that the string endpoint couples both to the gauge field and to the scalar fields on the D-brane. This is intuitive since, after all, we obtained these fields as the massless modes of a quantized open string with endpoints attached to the D-brane. Physically, the coupling to the scalar fields is just a

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

reflection of the fact that a string ending on a D-brane "pulls" on it and deforms its shape, thus exciting the scalar fields which parametrize this shape. The direction orthogonal to the D-brane in which the string pulls is specified by $\vec{n}$. The coupling to the gauge field reflects the fact that the string endpoint behaves as a pointlike particle charged under this gauge field. We thus conclude that an open string ending on a D-brane with a fixed $\vec{n}$ excites both the gauge and the scalar fields, which suggests that the correct Wilson loop operator dual to the string worldsheet must include both types of fields and must therefore be given by (5.68).

The dual description of the operator (5.67) is the same as that of (5.68) except that the Dirichlet boundary conditions on the string worldsheet along the S^5 directions must be replaced by Neumann boundary conditions [42] (see also [330]). One immediate consequence is that, to leading order, the strong coupling results for the Wilson loop (5.68) with constant $\vec{n}$ and for the Wilson loop (5.62) are the same. However, the two results differ at the next order in the $1/\sqrt{\lambda}$ expansion, since in the case of (5.67) we would have to integrate over the point on the sphere where the string is sitting. More precisely, at the one-loop level in the α'-expansion one finds that the determinants for quadratic fluctuations are different in the two cases [331].

In the large-N_c, large-λ limit, the string partition function $Z_{\text{string}}[\partial\Sigma = \mathcal{C}]$ greatly simplifies and is given by the exponential of the classical string action, i.e.

$$Z_{\text{string}}[\partial\Sigma = \mathcal{C}] = e^{iS(\mathcal{C})} \quad \rightarrow \quad \langle W(\mathcal{C})\rangle = e^{iS(\mathcal{C})} \,. \tag{5.70}$$

The classical action $S(\mathcal{C})$ can in turn be obtained by extremizing the Nambu–Goto action for the string worldsheet with the boundary condition that the string worldsheet ends on the curve $\mathcal{C}$. More explicitly, parameterizing the two-dimensional world sheet by the coordinates $\sigma^\alpha = (\tau, \sigma)$, the location of the string world sheet in the five-dimensional spacetime with coordinates x^M is given by the Nambu–Goto action (4.13). The fact that the action is invariant under coordinate changes of σ^α will allow us to pick the most convenient worldsheet coordinates (τ, σ) for each occasion.

Note that the large-N_c and large-λ limits are both crucial for (5.70) to hold. Taking $N_c \rightarrow \infty$ at fixed λ corresponds to taking the string coupling to zero, meaning that we can ignore the possibility of loops of string breaking off from the string world sheet. Additionally taking $\lambda \rightarrow \infty$ corresponds to sending the string tension to infinity, which implies that we can neglect fluctuations of the string world sheet. Under these circumstances the string worldsheet "hanging down" from the contour $\mathcal{C}$ takes on its classical configuration, without fluctuating or splitting off loops.

As a simple example let us first consider a contour $\mathcal{C}$ given by a straight line along the time direction with length $\mathcal{T}$ which describes an isolated static quark at rest. On the field theory side we expect that the expectation value of the Wilson line should be given by

$$\langle W(\mathcal{C})\rangle = e^{-iM\mathcal{T}}\,, \tag{5.71}$$

where M is the mass of the quark. From the symmetry of the problem, the corresponding bulk string worldsheet should be that of a straight string connecting the boundary and the Poincaré horizon and translated along the time direction by $\mathcal{T}$. The action of such a string worldsheet is infinite since the proper distance from the boundary to the center of AdS is infinite. This is consistent with the fact that the external quark has an infinite mass. A finite answer can nevertheless be obtained if we introduce an IR regulator in the bulk, putting the boundary at $z = \epsilon$ instead of $z = 0$. From the IR/UV connection this corresponds to introducing a short-distance (UV) cut-off in the boundary theory. Choosing $\tau = t$ and $\sigma = z$ the string worldsheet is given by $x^i(\sigma, \tau) = \text{const.}$, and the induced metric on the worldsheet is then given by

$$ds^2 = \frac{R^2}{\sigma^2}(-d\tau^2 + d\sigma^2)\,. \tag{5.72}$$

Evaluating the Nambu–Goto action on this solution yields

$$S = S_0 \equiv -\frac{\mathcal{T}R^2}{2\pi\alpha'}\int_\epsilon^\infty \frac{dz}{z^2} = -\frac{\sqrt{\lambda}}{2\pi\epsilon}\mathcal{T}\,, \tag{5.73}$$

where we have used the fact that $R^2/\alpha' = \sqrt{\lambda}$. Using (5.70) and (5.71) we then find that

$$M = \frac{\sqrt{\lambda}}{2\pi\epsilon}\,. \tag{5.74}$$

5.4.1 Rectangular loop: vacuum

Now let us consider a rectangular loop sitting at a constant position on the S^5 [711, 595]. The long side of the loop extends along the time direction with length $\mathcal{T}$, and the short side extends along the x_1-direction with length L. We will assume that $\mathcal{T} \gg L$. Such a configuration can be though of as consisting of a static quark–antiquark pair separated by a distance L. Therefore we expect that the expectation value of the Wilson loop (with suitable renormalization) gives the potential energy between the pair, i.e. we expect that

$$\langle W(\mathcal{C})\rangle = e^{-iE_{\text{tot}}\mathcal{T}} = e^{-i(2M+V(L))\mathcal{T}} = e^{iS(\mathcal{C})}\,, \tag{5.75}$$

where E_{tot} is the total energy for the whole system and $V(L)$ is the potential energy between the pair. In the last equality we have used (5.70). We will now proceed to calculate $S(\mathcal{C})$ for a rectangular loop.

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

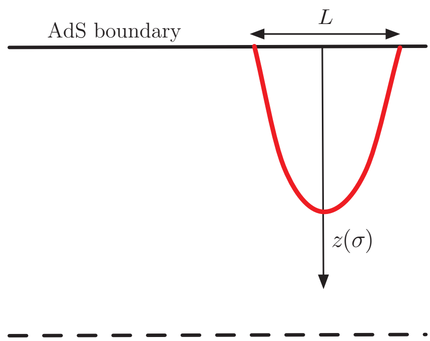

Figure 5.2 String (red) associated with a quark–antiquark pair.

It is convenient to choose the worldsheet coordinates to be

$$\tau = t, \qquad \sigma = x_1 \,. \tag{5.76}$$

Since $\mathcal{T} \gg L$, we can assume that the surface is translationally invariant along the τ direction, i.e. the extremal surface should have non-trivial dependence only on σ. Given the symmetries of the problem we can also set

$$x_3(\sigma) = \text{const.}\,, \qquad x_2(\sigma) = \text{const.} \tag{5.77}$$

Thus the only non-trivial function to solve for is $z = z(\sigma)$ (see Fig. 5.2), subject to the boundary condition

$$z\left(\pm\frac{L}{2}\right) = 0\,. \tag{5.78}$$

Using the form (5.3) of the spacetime metric and Eqs. (5.76)–(5.77), the induced metric on the worldsheet is given by

$$ds^2_{ws} = \frac{R^2}{z^2}\left(-d\tau^2 + (1 + z'^2)d\sigma^2\right)\,, \tag{5.79}$$

giving rise to the Nambu–Goto action

$$S_{\text{NG}} = -\frac{R^2\mathcal{T}}{2\pi\alpha'}\int_{-\frac{L}{2}}^{\frac{L}{2}} d\sigma\,\frac{1}{z^2}\sqrt{1+z'^2}\,, \tag{5.80}$$

where $z' = dz/d\sigma$. Since the action and the boundary condition are symmetric under $\sigma \to -\sigma$, $z(\sigma)$ should be an even function of σ. Introducing dimensionless coordinates via

$$\sigma = L\,\xi\,, \qquad z(\sigma) = L\,y(\xi) \tag{5.81}$$

we then have

$$S_{\rm NG} = -\frac{2R^2}{2\pi\alpha'}\frac{\mathcal{T}}{L}\, Q\,, \qquad \text{with } Q = \int_0^{\frac{1}{2}} \frac{d\xi}{y^2}\sqrt{1+y'^2}\,. \tag{5.82}$$

Note that Q is a numerical constant. As we will see momentarily, it is in fact divergent and therefore it should be defined more carefully. The equation of motion for y is given by

$$y'^2 = \frac{y_0^4 - y^4}{y^4} \tag{5.83}$$

with y_0 the turning point at which $y' = 0$, which by symmetry should happen at $\xi = 0$. Thus, y_0 can be determined by the condition

$$\frac{1}{2} = \int_0^{\frac{1}{2}} d\xi = \int_0^{y_0} \frac{dy}{y'} = \int_0^{y_0} dy\, \frac{y^2}{\sqrt{y_0^4 - y^4}} \quad \to \quad y_0 = \frac{\Gamma(\frac{1}{4})}{2\sqrt{\pi}\,\Gamma(\frac{3}{4})}\,. \tag{5.84}$$

It is then convenient to change integration variable in Q from ξ to y to get

$$Q = y_0^2 \int_0^{y_0} \frac{dy}{y^2\sqrt{y_0^4 - y^4}}\,. \tag{5.85}$$

This is manifestly divergent at $y = 0$, but the divergence can be interpreted as coming from the infinite rest masses of the quark and the antiquark. As in the discussion after (5.71), we can obtain a finite answer by introducing an IR cut-off in the bulk by putting the boundary at $z = \epsilon$, i.e. by replacing the lower integration limit in (5.85) by ϵ. The potential $V(L)$ between the quarks is then obtained by subtracting $2M\mathcal{T}$ from (5.82) (with M given by (5.74)) and then taking $\epsilon \to 0$ at the end of the calculation. One then finds the finite answer

$$V(L) = -\frac{4\pi^2}{\Gamma^4(\frac{1}{4})}\frac{\sqrt{\lambda}}{L}\,, \tag{5.86}$$

where again we used the fact that $R^2/\alpha' = \sqrt{\lambda}$ to translate from gravity to gauge theory variables. Note that the $1/L$ dependence is simply a consequence of conformal invariance. The non-analytic dependence on the coupling, i.e. the $\sqrt{\lambda}$ factor, could not be obtained at any finite order in perturbation theory. From the gravity viewpoint, however, it is a rather generic result, since it is due the fact that the tension of the string is proportional to $1/\alpha'$. The above result is valid at large λ. At small λ, the potential between a quark and an antiquark in an $\mathcal{N} = 4$ theory is given by [342]

$$E = -\frac{\pi\lambda}{L} \tag{5.87}$$

to lowest order in the weak coupling expansion.

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

It is remarkable that the calculation of a Wilson loop in a strongly interacting gauge theory has been simplified to a classical mechanics problem no more difficult than finding the catenary curve describing a string suspended from two points, hanging in a gravitational field – in this case the gravitational field of the AdS spacetime.

Note that given (5.86), the boundary short-distance cut-off ϵ in (5.74) can be interpreted as the size of the external quark. One might have expected (incorrectly) that a short distance cut-off on the size of the quark should be given by the Compton wavelength $1/M \sim \epsilon/\sqrt{\lambda}$, which is much smaller than ϵ. Note that the size of a quark should be defined by either its Compton wavelength or by the distance between a quark and an antiquark at which the potential is of the order of the quark mass, whichever is bigger. In a weakly coupled theory, the Compton wavelength is bigger, while in a strongly coupled theory with potential (5.86), the latter is bigger and is of order ϵ.

5.4.2 Rectangular loop: nonzero temperature

We now consider the expectation of the rectangular loop at nonzero temperature [712, 190]. In this case the bulk gravity geometry is given by that of the black brane (5.34). The set-up of the calculation is exactly the same as in Eqs. (5.76)–(5.78) for the vacuum. The induced worldsheet metric is now given by

$$ds^2_{ws} = \frac{R^2}{z^2}\left(-f(z)d\tau^2 + \left(1 + \frac{z'^2}{f}\right)d\sigma^2\right), \tag{5.88}$$

which yields the Nambu–Goto action

$$S_{\rm NG} = -\frac{R^2\mathcal{T}}{2\pi\alpha'}\int_{-\frac{L}{2}}^{\frac{L}{2}} d\sigma\,\frac{1}{z^2}\sqrt{f(z) + z'^2}\,. \tag{5.89}$$

The crucial difference between the equation of motion following from (5.89) and that following from (5.80) is that in the present case there exists a maximal value $L_s \sim 1/T$ beyond which nontrivial solutions cease to exist [712, 190] – see Fig. 5.3. Instead, the solution beyond this maximal separation consists of two disjoint vertical strings ending at the black hole horizon. The physical reason can be easily understood qualitatively from the figure. At some separation, the lowest point on the string touches the horizon. Surely at and beyond this separation the string can minimize its energy by splitting into two independent strings, each of which falls through the horizon. The precise value of L_s is defined as the quark–antiquark separation at which the free energy of the disconnected configuration becomes smaller than that of the connected configuration. This happens at a value of L for which the lowest point of the connected configuration is close to but still somewhat above the horizon. Once $L > L_s$, the quark–antiquark separation can

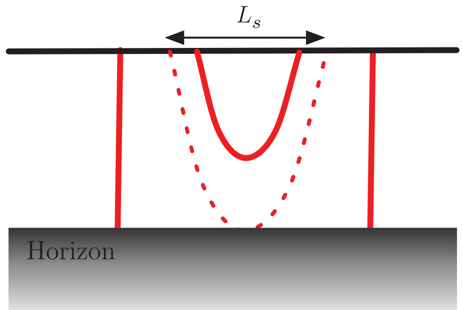

Figure 5.3 String (red) associated with a quark–antiquark pair in a plasma with temperature $T > 0$. The preferred configuration beyond a certain separation L_s consists of two independent strings.

then be increased further at no additional energy cost, so the potential becomes constant and the quark and the antiquark are perfectly screened from each other by the plasma between them. See, for example, Ref. [108] for a careful discussion of the corrections to this large-N_c, large-λ result.

5.4.3 Rectangular loop: a confining theory

For comparison, let us consider the expectation value of a rectangular loop in the $2+1$-dimensional confining theory [804] (for a review see [760]) whose metric is given by (5.38), which we reproduce here for convenience:

$$ds^2 = \frac{R^2}{z^2}\left(-dt^2 + dx_1^2 + dx_2^2 + f\, dt_{\mathrm{E}}\right) + \frac{R^2}{z^2 f} dz^2, \qquad f = 1 - \frac{z^4}{z_0^4}. \tag{5.90}$$

As discussed earlier, the crucial difference between (5.90) and AdS is that the spacetime (5.90) ends smoothly at a finite value $z = z_0$, which introduces a scale in the theory. The difference as compared to the finite-temperature case is that in the confining geometry the string has no place to end, so in order to minimize its energy it tends to drop down to z_0 and to run parallel there – see Fig. 5.4.

Again the set-up of the calculation is completely analogous to the cases above. The induced worldsheet metric is now given by

$$ds_{ws}^2 = \frac{R^2}{z^2}\left(-d\tau^2 + \left(1 + \frac{z'^2}{f}\right) d\sigma^2\right), \tag{5.91}$$

and the corresponding the Nambu–Goto action is

$$S_{\mathrm{NG}} = -\frac{R^2 \mathcal{T}}{2\pi\alpha'} \int_{-\frac{L}{2}}^{\frac{L}{2}} d\sigma \frac{1}{z^2}\sqrt{1 + \frac{z'^2}{f(z)}}\,. \tag{5.92}$$

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

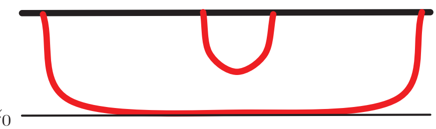

Figure 5.4 String (red) associated with a quark–antiquark pair in a confining theory.

When L is large, the string quickly drops to $z = z_0$ and runs parallel there. We thus find that the action can be approximated by (after subtracting the vertical parts which can be interpreted as being due to the static quark masses)

$$-S(\mathcal{C}) - 2M\mathcal{T} \approx \frac{R^2\mathcal{T}}{2\pi\alpha'}\frac{L}{z_0^2}\,, \tag{5.93}$$

which gives rise to a confining potential

$$V(L) = \sigma_s L\,, \qquad \sigma_s = \frac{\sqrt{\lambda}}{2\pi z_0^2}\,. \tag{5.94}$$

The constant σ_s can be interpreted as the effective string tension. As mentioned in Section 5.2.2 the mass gap for this theory is $M \sim 1/z_0$, so we find that $\sigma_s \sim \sqrt{\lambda}M^2$. Although we have described the calculation only for one example of a confining gauge theory, the qualitative features of Fig. 5.4 generalize. In a confining gauge theory with a dual gravity description, as a quark–antiquark pair are separated the string hanging beneath them sags down to some "depth" z_0 and then as the separation is further increased it sags no further. Further increasing the separation means adding more and more string at the same depth z_0, which costs an energy that increases linearly with separation. Clearly, any metric in which a suspended string behaves like this cannot be conformal; it has a length scale z_0 built into it in some way. This length scale z_0 in the gravitational description corresponds via the IR/UV correspondence to the mass gap $M \sim 1/z_0$ for the gauge theory and to the size of the "glueballs" in the gauge theory, which is of order z_0.

To summarize, we note that the qualitative behavior of the Wilson loop discussed in various examples above is only determined by gross features of the bulk geometry. The $1/L$ behavior (5.87) in the conformal vacuum follows directly from the scaling symmetry of the bulk geometry; the area law (5.93) in the confining case has to do with the fact that a string has no place to end in the bulk when the geometry smoothly closes off; and the screening behavior at finite temperature is a consequence of the fact that a string can fall through the black hole horizon. The difference between Figs. 5.2 and 5.4 highlights the fact that $\mathcal{N} = 4$ SYM theory is not a good model for the vacuum of a confining theory like QCD. However, as we will discuss in Section 8.7, the potential obtained from Fig. 5.3 is not a bad caricature of what happens in the deconfined phase of QCD. This is one of many ways of seeing that $\mathcal{N} = 4$ SYM at $T \neq 0$ is more similar to QCD above T_c than $\mathcal{N} = 4$ SYM at $T = 0$ is to QCD at $T = 0$. A heuristic way of thinking about this is to note that at low temperatures the putative horizon would be at a $z_{\rm hor} > z_0$, i.e. it is far below the bottom of Fig. 5.4, and therefore it plays no role while at large temperatures, the horizon is far above z_0 and it is z_0 that plays no role. At some intermediate temperature, the theory has undergone a phase transition from a confined phase described by Fig. 5.4 into a deconfined phase described by Fig. 5.3.[5] Unlike in QCD, this deconfinement phase transition is a first order phase transition in the large-N_c, strong coupling limit under consideration, and the theory in the deconfined phase loses all memory about the confinement scale z_0. Presumably corrections away from this limit, in particular finite-N_c corrections, could turn the transition into a higher order phase transition or even a crossover.

5.5 Introducing fundamental matter

All the matter degrees of freedom of $\mathcal{N} = 4$ SYM, the fermions and the scalars, transform in the adjoint representation of the gauge group. In QCD, however, the quarks transform in the fundamental representation. Moreover, most of what we know about QCD phenomenologically comes from the study of quarks and their bound states. Therefore, in order to construct holographic models more closely related to QCD, we must introduce degrees of freedom in the fundamental representation. It turns out that there is a rather simple way to do this in the limit in which the number of quark species, or flavors, is much smaller than the number of colors, i.e. when $N_f \ll N_c$. Indeed, in this limit the introduction of N_f flavors in the gauge theory corresponds to the introduction of N_f D-brane probes in the AdS geometry sourced by the D3-branes [28, 517, 513]. This is perfectly consistent with the well-known fact that the topological representation of the large-N_c expansion of a gauge theory with quarks involves Riemann surfaces with boundaries – see Section 4.1.2. In the string description, these surfaces correspond to the worldsheets of open strings whose endpoints must be attached to D-branes. In the context of the gauge/string duality, the intuitive idea is that closed strings living in AdS are dual to gauge-invariant operators constructed solely out of gauge fields and adjoint matter, e.g. $\mathcal{O} = \mathrm{Tr} F^2$, whereas open strings are dual to meson-like operators, e.g. $\mathcal{O} = \bar{q} q$. In particular, the two endpoints of an open string, which are forced to lie on the D-brane probes, are dual to a quark and an antiquark, respectively.

[5] The way we have described the transition is a crude way of thinking about the so-called Hawking–Page phase transition between a spacetime without and with a black hole [437, 803].

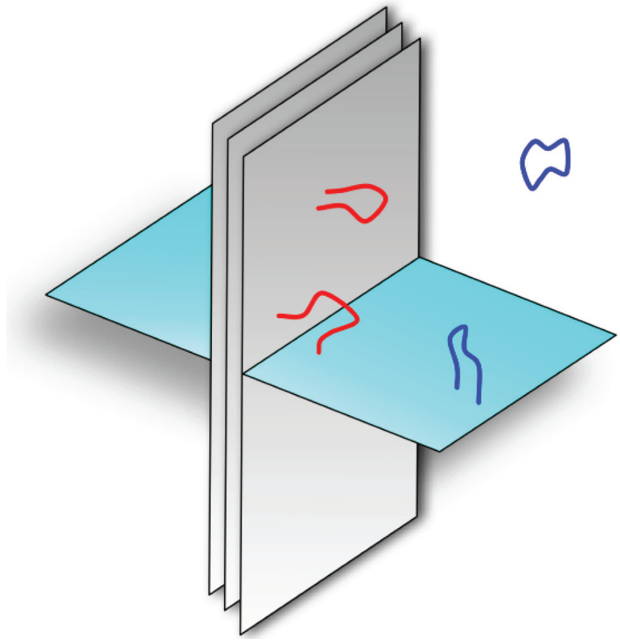

Figure 5.5 Excitations of the system in the open string description.

5.5.1 The decoupling limit with fundamental matter

The fact that the introduction of gauge theory quarks corresponds to the introduction of D-brane probes in the string description can be more "rigorously" motivated by repeating the arguments of Sections 4.2.2, 4.2.3 and 4.3 in the presence of N_f Dp-branes, as indicated in Fig. 5.5. We shall be more precise about the value of p and the precise orientation of the branes later; for the moment we simply assume $p > 3$.

As in Section 4.2.2, when $g_s N_c \ll 1$ the excitations of this system are accurately described by interacting closed and open strings living in flat space. In this case, however, the open string sector is richer. As before, open strings with both endpoints on the D3-branes give rise, at low energies, to the $\mathcal{N} = 4$ SYM multiplet in the adjoint of $SU(N_c)$. We see from Eq. (4.17) that the coupling constant for these degrees of freedom is dimensionless, and therefore these degrees of freedom remain interacting at low energies. The coupling constant for the open strings with both endpoints on the Dp-branes, instead, has dimensions of (length)$^{p-3}$. Therefore the effective dimensionless coupling constant at an energy E scales as $g_{\mathrm{Dp}} \propto E^{p-3}$. Since we assume that $p > 3$, this implies that, just like the closed strings, the p–p strings become noninteracting at low energies. Finally, consider the sector of open strings with one endpoint on the D3-branes and one endpoint on the Dp-branes. These degrees of freedom transform in the fundamental of the gauge group

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

on the D3-branes and in the fundamental of the gauge group on the Dp-branes, namely in the bifundamental of $SU(N_c) \times SU(N_f)$. Consistently, these 3–p strings interact with the 3–3 and the p–p strings with strengths given by the corresponding coupling constants on the D3-branes and on the Dp-branes. At low energies, therefore, only the interactions with the 3–3 strings survive. In addition, since the effective coupling on the Dp-branes vanishes, the corresponding gauge group $SU(N_f)$ becomes a global symmetry group. This is the origin of the flavor symmetry expected in the presence of N_f (equal mass) quark species in the gauge theory.

To summarize, when $g_s N_c \ll 1$ the low energy limit of the D3/Dp system yields two decoupled sectors. The first sector is free and consists of closed strings in ten-dimensional flat space and p-p strings propagating on the worldvolume of N_f Dp-branes. The second sector is interacting and consists of a four-dimensional $\mathcal{N} = 4$ SYM multiplet in the adjoint of $SU(N_c)$, coupled to the light degrees of freedom coming from the 3–p strings. We will be more precise about the exact nature of these degrees of freedom later, but for the moment we emphasize that they transform in the fundamental representation of the $SU(N_c)$ gauge group, and in the fundamental representation of a global, flavor symmetry group $SU(N_f)$.

Consider now the closed string description at $g_s N_c \gg 1$. In this case, as in Section 4.2.3, the D3-branes may be replaced by their backreaction on spacetime. If we assume that $g_s N_f \ll 1$, which is consistent with $N_f \ll N_c$, we may still neglect the backreaction of the Dp-branes. In other words, we may treat the Dp-branes as probes living in the geometry sourced by the D3-branes, with the Dp-branes not modifying this geometry. The excitations of the system in this limit consist of closed strings and open p–p strings that propagate in two different regions, the asymptotically flat region and the $\mathrm{AdS}_5 \times S^5$ throat – see Fig. 5.6. As in Section 4.2.3, these two regions decouple from each other in the low energy limit. Also as in Section 4.2.3, in this limit the strings in the asymptotically flat region become noninteracting, whereas those in the throat region remain interacting because of the gravitational redshift.

Comparing the two descriptions above, we see that the low energy limit at both small and large values of $g_s N_c$ contains a free sector of closed and open p–p strings. As in Section 4.3 we identify these free sectors, and we conjecture that the interacting sectors on each side provide dual descriptions of the same physics. In other words, we conjecture that the $\mathcal{N} = 4$ SYM coupled to N_f flavors of fundamental degrees of freedom is dual to type IIB closed strings in $\mathrm{AdS}_5 \times S^5$, coupled to open strings propagating on the worldvolume of N_f Dp-brane probes.

It is worth clarifying the following conceptual point before closing this section. It is sometimes stated that, in the 't Hooft limit in which $N_c \to \infty$ with N_f fixed, the dynamics is completely dominated by the gluons, and therefore that the quarks can be completely ignored. One may then wonder what the interest of introducing

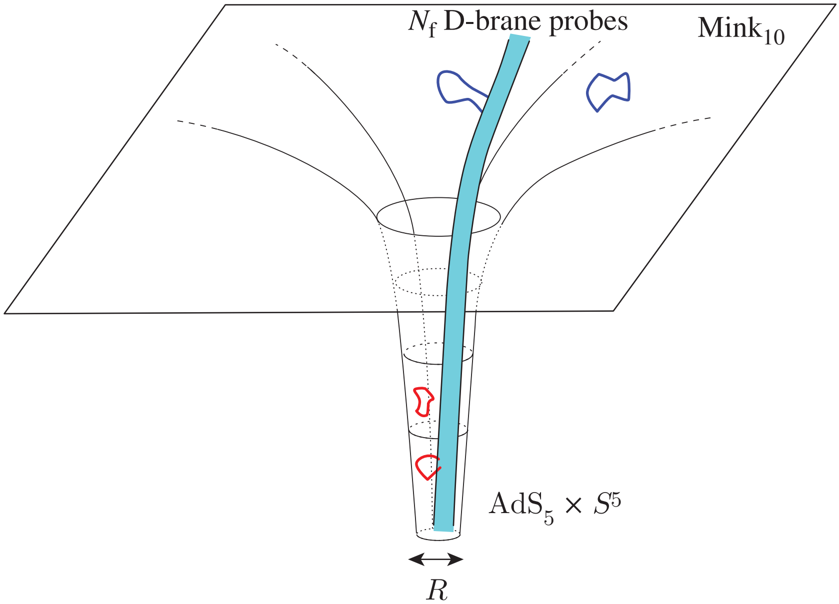

Figure 5.6 Excitations of the system in the second description.

fundamental degrees of freedom in a large-N_c theory may be. There are several answers to this. First of all, in the presence of fundamental matter, it is more convenient to think of the large-N_c limit *à la* Veneziano, in which N_f/N_c is kept small but finite. Any observable can then be expanded in powers of $1/N_c^2$ and N_f/N_c. As we will see, this is precisely the limit that is captured by the dual description in terms of N_f D-brane probes in AdS$_5$ × S^5. The leading D-brane contribution will give us the leading contribution of the fundamental matter, of relative order N_f/N_c. The Veneziano limit is richer than the 't Hooft limit, since setting $N_f/N_c = 0$ one recovers the 't Hooft limit. The second point is that, even in the 't Hooft limit, the quarks should not be regarded as irrelevant, but rather as valuable probes of the gluon-dominated dynamics. It is their very presence in the theory that allows one to ask questions about heavy quarks in the plasma, jet quenching, meson physics, photon emission, etc. The answers to these questions are of course dominated by the gluon dynamics, but without dynamical quarks in the theory such questions cannot even be posed. There is a completely analogous statement in the dual gravity description. To leading order the geometry is not modified by the presence of D-brane probes, but one needs to introduce these probes in order to pose questions about heavy quarks in the plasma, parton energy loss, mesons, photon production, etc. In this sense, the D-brane probes allow one to decode information already contained in the geometry.

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

5.5.2 Models with fundamental matter

Above, we motivated the inclusion of fundamental matter via the introduction of N_f "flavor" Dp-brane probes in the background sourced by N_c "color" D3-branes. However, we were deliberately vague about the value of p, about the relative orientation between the flavor and the color branes, and about the precise nature of the flavor degrees of freedom in the gauge theory. Here we will address these points. Since we assumed $p > 3$ in order to decouple the p–p strings, and since we wish to consider stable Dp-branes in type IIB string theory, we must have $p = 5$ or $p = 7$ – see Section 4.2.2. In other words, we must consider D5- and D7-brane probes.

Consider first adding flavor D5-branes. We will indicate the relative orientation between these and the color D3-branes by an array like, for example,

$$\begin{array}{lccccccccc} \text{D3:} & 1 & 2 & 3 & \_ & \_ & \_ & \_ & \_ & \_ \\ \text{D5:} & 1 & 2 & \_ & 4 & 5 & 6 & \_ & \_ & \_ \end{array} \,. \tag{5.95}$$

This indicates that the D3- and the D5-branes share the 12-directions. The 3-direction is transverse to the D5-branes, the 456-directions are transverse to the D3-branes, and the 789-directions are transverse to both sets of branes. This means that the two sets of branes can be separated along the 789-directions, and therefore they do not necessarily intersect, as indicated in Fig. 5.7. It turns out that the lightest states of a D3–D5 string have a minimum mass given by what one would have expected on classical grounds, namely $M = T_{\text{str}} L = L/2\pi\ell_s^2$, where T_{str} is the string tension (4.11) and L is the minimum distance between the D3- and the D5-branes.[6] These states can therefore be arbitrarily light, even massless, provided L is sufficiently small. Generic excited states, as usual, have an additional mass set by the string scale alone, m_s. The only exception are excitations in which the string moves rigidly with momentum $\vec{p}$ in the 12-directions, in which case the energy squared is just $M^2 + \vec{p}^{\,2}$. This is an important observation because it means that in the decoupling limit, in which one focuses on energies $E \ll m_s$, only a finite set of modes of the D3–D5 strings survive, and moreover these modes can only propagate along the directions common to both branes. From the viewpoint of the dual gauge theory, this translates into the statement that the degrees of freedom in the fundamental representation are localized on a defect – in the example at hand, on a plane that extends along the 12-directions and lies at a constant position in the 3-direction. As an additional example, the configuration

[6] In order to really establish this formula one must quantize the D3–D5 strings and compute the ground state energy. In the case at hand, the result coincides with the classical expectation. The underlying reason is that, because the configuration (5.95) preserves supersymmetry, corrections to the classical ground state energy coming from bosonic and fermionic quantum fluctuations cancel each other out exactly. For other brane configurations like (5.97) this does not happen.

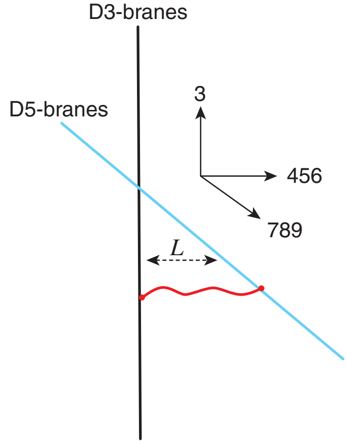

Figure 5.7 D3–D5 configuration (5.95) with a string (red) stretching between them. The 12-directions common to both branes are suppressed.

$$
\begin{array}{lccccccccc}
\text{D3:} & 1 & 2 & 3 & \_ & \_ & \_ & \_ & \_ & \_ \\
\text{D5:} & 1 & \_ & \_ & 4 & 5 & 6 & 7 & \_ & \_
\end{array}
\tag{5.96}
$$

corresponds to a dual gauge theory in which the fundamental matter is localized on a line – the 1-direction.

We thus conclude that, if we are interested in adding to the $\mathcal{N} = 4$ SYM theory fundamental matter degrees of freedom that propagate in 3+1 dimensions (just like the gluons and the adjoint matter), then we must orient the flavor D-branes so that they extend along the 123-directions. This condition leaves us with two possibilities:

$$
\begin{array}{lccccccccc}
\text{D3:} & 1 & 2 & 3 & \_ & \_ & \_ & \_ & \_ & \_ \\
\text{D5:} & 1 & 2 & 3 & 4 & 5 & \_ & \_ & \_ & \_
\end{array}
\tag{5.97}
$$

and

$$
\begin{array}{lccccccccc}
\text{D3:} & 1 & 2 & 3 & \_ & \_ & \_ & \_ & \_ & \_ \\
\text{D7:} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & \_ & \_
\end{array} \; .
\tag{5.98}
$$

So far we have not been specific about the precise nature of the fundamental matter degrees of freedom – for example, whether they are fermions or bosons, etc. This also depends on the relative orientation of the branes. It turns out that for the configuration (5.97), the ground state energy of the D3–D5 strings is (for sufficiently small L) negative, that is, the ground state is tachyonic, signaling an instability in the system. This conclusion is valid at weak string coupling, where the string spectrum can be calculated perturbatively. While it is possible that the

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press

instability is absent at strong coupling, we will not consider this configuration further in this book.

We are therefore left with the D3–D7 system (5.98). Quantization of the D3–D7 strings shows that the fundamental degrees of freedom in this case consist of N_f complex scalars and N_f Dirac fermions, all of them with equal masses given by

$$M_{\mathrm{q}} = \frac{L}{2\pi\alpha'} \,. \tag{5.99}$$

In a slight abuse of language, we will collectively refer to all these degrees of freedom as “quarks”. The fact that they all have exactly equal masses is a reflection of the fact that the addition of the N_f D7-branes preserves a fraction of the original supersymmetry of the SYM theory. More precisely, the original $\mathcal{N} = 4$ is broken down to $\mathcal{N} = 2$, under which the fundamental scalars and fermions transform as part of a single supermultiplet. In the rest of the book, especially in Chapter 9, we will focus our attention on this system as a model for gauge theories with fundamental matter.

https://doi.org/10.1017/9781009403504.005 Published online by Cambridge University Press