EPJ manuscript No.
(will be inserted by the editor)

QCD, Strings and Emergent Space

Sumit R. Das[1,a]

Department of Physics and Astronomy, University of Kentucky, Lexington, KY 40506, USA.

Abstract. This is an account of how attempts to derive String Theory from QCD led to contemporary ideas about emergent space and holography.

1 Introduction

Soon after Nambu passed away, his long-time colleague Peter Freund paid a beautiful tribute: "He would pull one rabbit out of the hat, and another, and then suddenly the rabbits would arrange themselves in a pattern and start dancing in a way youd never seen before. Where he got the idea, you could never imagine [1]". Nothing could describe the unfathomable originality of Nambu's work better. In this contribution I will talk about one of those rabbits which is perhaps less known than the others, but sparked a development which eventually led to a major revision of the way we think of gravity and space-time - something he probably did not anticipate [1].

2 QCD and Strings

The story begins with QCD and its relationship to String Theory - both of which Nambu pioneered. String Theory arose as a model for hadrons [2]. Soon after the discovery of asymptotic freedom, it was realized by several authors that these strings were confined electric flux tubes in the QCD vacuum - an idea developed by Nambu, 't Hooft and Mandelstam. This, however, raised several questions.

The first question relates to the origin of a string coupling constant. Usually, when we talk about particles, we always have a small number in the problem such that when this number is taken to zero, the particles are weakly coupled. For example in QED an electron is a rather complicated object with a cloud around it because of vacuum polarization. However, in the limit of a vanishing fine structure constant, the cloud disappears and one has a point particle. Strings in QCD are similarly fat objects with a width of the order of Λ_{QCD}. The problem, however, is that there is no dimensionless parameter in massless QCD so that we can tune it to a small value get a thin string. The Yang-Mills coupling which appears in the lagrangian is dimensionally transmuted into a scale.

[a] e-mail: das@pa.uky.edu

[1] A couple of disclaimers. First, this is not a historical account. Accordingly, the developments reviewed below are not presented in chronological order. Secondly, for the most part I have cited review articles or books rather than the original papers.

In a seminal paper,'t Hooft provided a surprising answer to this question. Instead of considering the usual $SU(3)$ gauge group of strong interactions, he considered a $SU(N)$ gauge group [3,4]. He then showed that in the limit of $N \to \infty$ and $g_{YM} \to 0$ with the 't Hooft coupling $\lambda = g_{YM}^2 N$ held fixed, Feynman diagrams can be thought of as discretizations of a two dimensional surface with an overall factor of N^χ, where χ is the Euler characteristic of the surface. However, this is precisely the topological expansion of string amplitudes in terms of worldsheets. So the idea was that these Feynman diagrams are discretizations of the worldsheet of strings and the coupling of this string theory is $1/N$. This sounds weird. How could a $N = \infty$ theory provide a reasonable approximation to a $N = 3$ theory ? However as 't Hooft argued, many phenomena in strong interactions can be understood in the context of an expansion in powers of $1/N$ - phenomena which are otherwise very difficult to understand. An important consequence of this is that the $N \to \infty$ limit is a classical limit. In this limit expectation values of products of gauge invariant operators factorize.

The second question is: what is the quantity in the Yang-Mills theory which behaves as a String Field ? What is the equation which this object obeys ? This question was addressed by Nambu [5]. Nambu proposed that the string field is the Wilson loop for a closed curve C

$$U(C) = P\exp[i \int_C A_\mu dx^\mu] \tag{1}$$

He showed that the equation for this object can be obtained by considering a small deformation of the loop C, leading to

$$\frac{\delta}{\delta \sigma_{\mu t}} U(C) = iU(C')F_{\mu t}(z)U(C'') \tag{2}$$

Here z is the point on the loop where a deformation is made and the loop C is split up into C' andC''. μ denotes a normal direction at a point z while t is the tangential direction. $\sigma_{\mu t}$ is the infinitesimal area element in (μt) plane caused by the deformation and $F_{\mu t}$ is the gauge field strength. Equation (2) can be thought of an equation in loop space, with the left hand side defining a loop space derivative. [5] showed that when the gauge field obeys some conditions, this leads to Virasoro equations for a string.

It turns out that a precise definition of the loop space derivative is rather subtle. Soon after [5] appeared, several groups realized that this can be done if one works on a lattice. In that case one can derive precise Dyson Schwinger equations relating Wilson loops and its correlators in the 't Hooft large N limit. Factorization in this limit implies that the expectation value of the Wilson loop operator itself obeys a closed equation [6]. These papers led to a vigorous effort to derive the string theory which comes from QCD. However, even today this program has not worked very well - mostly hindered by ambiguties in renormalizing these equations properly.

3 Solvable Matrix Models

Inspired by 't Hooft's Large N limit, many authors explored solvable models of large N matrices. Among them a paper by Brezin, Itzykson, Parisi and Zuber (BIPZ) played a major role in future developments. The idea was to consider a carricature of Yang-Mills theory by ignoring part or whole of the space-time dependence. When the entire space-time dependence is ignored, one has a path integral over matrices. One of the models which BIPZ considered is that of a single hermitian matrix M with an action of the form

$$\int dM \, \exp[-S] \tag{3}$$

where S is an action which is invariant under the unitary transformations of M and dM stands for the invariant measure on Hermitian matrices. For example, one could have

$$S = \frac{1}{2g^2}\,\mathrm{Tr}\left[M^2 + V(M)\right] \tag{4}$$

This kind of random matrix model has a long history starting with the work of Wigner and Dyson where such models were used to model energy levels of complex nuclei.

The ‘t Hooft large N limit would entail $g \to 0, N \to \infty$ with $\lambda = g^2 N$ held fixed. Thus the overall coefficient can be written as $\frac{N}{\lambda}$. In this large-N limit one can then proceed by evaluating the integral by saddle point method. A key result was that for the particular model (4) there is a critical coupling corresponding to a finte radius of convergence of a perturbative expansion in g.

Another class of models which was solved in BIPZ concerns the quantum mechanics of a single hermitian matrix with a hamiltonian of the form

$$H = \mathrm{Tr}\left[-\frac{\partial}{\partial M}\frac{\partial}{\partial M} + V(M)\right] \tag{5}$$

This class of models will play a key role in the following.

These works brought Matrix Models as the center of the discussion with numerous works exploring various large-N Matrix Models and their phase structure [7]. In particular, models of unitary matrices were shown to display novel phase transitions now known as Gross-Witten-Wadia transitions.

4 Manufacturing space-time

The quest for obtaining the correct loop equations led to a rather unexpected consequence. In a remarkable paper, Eguchi and Kawai argued that in the leading order of the large N limit the Dyson Schwinger equations obeyed by Wilson loops in some d dimensional Euclidean lattice gauge theory are identical to those obtained from a single plaquette model. The order in which the link matrices appear in the “Wilson loop” in the single plaquette model somehow encodes space. Therefore a large-N gauge theory in some d dimensional euclidean space can be expressed entirely in terms of an integral over matrices. Working from the matrix model side this means that at large N the dynamics can be faithfully *interpreted* as that of a quantum field theory. The initial idea required modification - one final form of this proposal is the “Twisted Eguchi-Kawai Model”, which is in fact quite general. Starting from a field theory of e.g. a $N \times N$ hermitian scalar field one can arrive at this matrix model with no space by the substitution

$$\phi(x) \to D(x)\Phi D^\dagger(x) \tag{6}$$

where Φ is a single $N \times N$ hermitian matrix and e.g. in 4 euclidean dimensions, and

$$D(x) = \prod_{\mu=1}^{4} (\Gamma_\mu)^{x_\mu} \tag{7}$$

Γ_μ are $N \times N$ matrices obeying the ‘t Hooft algebra

$$\Gamma_\mu \Gamma_\nu = \exp[\frac{2\pi i}{N} n_{\mu\nu}]\Gamma_\nu \Gamma_\mu \tag{8}$$

and $n_{\mu\nu}$ is a 4×4 antisymmetric matrix with integer entries. The substitution (6) in the action yields the action of the “reduced model”. One can then use the reduced model to calculate correlators of Φ, and (6) reproduces the correlators of the field theory in the large N limit. There are various versions of reduced models, which are e.g. reviewed in [8]. The problem with all those, however, is that the $1/N$ corrections do not correspond to the $1/N$ corrections of the corresponding field theory. Neverthless this was an indication that space-time can be manufactured from internal degrees of freedom.

5 Two dimensional Gravity and Dynamical Triangulations

String theory arose as a model of hadrons. However the mathematical formulation of String theory seemed to require that the dimensionality of space-time should be 26 for bosonic strings and 10 for fermionic strings. This motivated a lot of effort to try to find ways to find consistent string theories in lower dimensions. In 1980 Polyakov published two papers which provided a new perspective. Polyakov showed that the Feynman path integral representation of the propagator of a single string may be written (for a bosonic string) as

$$\int \mathcal{D}X^\mu(\xi^a)\mathcal{D}h_{ab}(\xi^a) \exp\left[-\frac{1}{4\pi\alpha'}\int d\xi^a \sqrt{h}h^{ab}\partial_a X^\mu \partial_b X^\nu \eta_{\mu\nu}\right] \qquad (9)$$

Here $\xi^a, a = 1, 2$ are two coordinates on the worldsheet of the string, $X^\mu(\xi), \mu = 1 \cdots d$ denotes the embedding of the worldsheet in d dimensional space-time, and h_{ab} denotes the intrinsic metric on the worldsheet. At the classical level one can use the equations of motion to express h_{ab} in terms of the X^μ - the resulting action is the Nambu-Goto action. The worldsheet of the string is therefore a theory of two dimensional gravity coupled to "matter" fields X^μ. Gravity in two dimensions has only one physical mode. In a conformal gauge $h_{ab} = e^\phi \delta_{ab}$ and the field ϕ is the Liouville field. In this formalism the role of critical dimensions of string theory became transparent.

In his original paper, ‘t Hooft had shown that the Feynman diagrams of Yang-Mills theory form a discretization of a two dimensional surface, which can be thought of a discretization of a string worldsheet. It was also clear that this applies to a field theory of any set of matrix valued fields. Consider the dual of this lattice discretization. If the field theory in question is e.g. a M^3 field theory, the dual lattice is made of triangles which may be taken to be equilateral. If six such triangles meet at a vertex of this dual lattice, the surface has vanishing curvature at this point. By the same token, more than six or less than six triangles meeting at a point leads to a defect with an associated curvature. Thus the sum over Feynman diagrams becomes a sum over all possible triangulations of the surface. In other words one has a dynamically triangulated worldsheet. As ‘t Hooft had realized, the string coupling for this theory should be $1/N$.

In the mid-1980's it was realized by Kazakov and Migdal, and by Ambjorn, Druhus and Frohlich that if there is a critical point of the theory of matrices where the average number of tiles in a Feynman diagram diverges, one should be able to take a continuum limit so that the surface generated by these diagrams becomes smooth. Furthermore, the sum over triangulations becomes an integral over various metrics over the two dimensional surface - the integral over h_{ab} in (9). If the matrix field theory involves a matrix $M(x^\mu)$, each Feynman diagram involves an integral over the x^μ at the vertices of the diagram. This is the integral over X^μ in (9). In particular, if one considers the Feynman diagrams of a Matrix Model like (4), the only integral that remains is the integral over this metric - pure two dimensional gravity. For a field theory in d dimensions, one has d scalar fields coupled to two dimensional gravity. In other words one has a field theory on a random surface.

A major breakthrough in this line of development came from another seminal paper by Knizhnik, Polyakov and Zamolodchikov (KPZ) in 1988, followed by two other papers by David and by Distler and Kawai. In these papers, one understood clearly how Liouville theory coupled to matter can be quantized in a consistent fashion, and how to understand critical behavior on random surfaces. In particular KPZ derived a formula which expresses the critical exponents of a conformal field theory coupled to Liouville in terms of the critical exponents of the theory withut dynamical gravity and the central charge. Soon after this paper, Kazakov showed that one can choose the potential $V(M)$ in Matrix Models of the type (4) to describe multicritical points with exponents of conformal field theories with central charges $c < 1$ coupled to dynamical two dimensional gravity predicted by KPZ. These works led to a very fruitful period of research in the theory of noncritical strings [9]. However, it turned out to be hard to find any consistent bosonic string theory with $d > 1$: these theories always had a tachyon.

6 Strings and Quantum Gravity

While this activity centered on understanding String Theory starting from QCD was going on, a parallel and seemingly disconnected program was pursued by a handful of physicists. This concerned developing String Theory, more particularly supersymmetric string theory as a consistent theory of quantum gravity. The story began with two independent papers in 1974 : one by Yoneya (then a graduate student at Hokkaido) and the other by Scherk and Schwarz. The spectrum of String Theory always contained a massless spin-2 mode: understanding this mode as a hadronic excitation has always been a challenge. Yoneya and Scherk and Schwarz took the revolutionary step of proposing that this spin-2 particle is in fact the graviton - the quantum of gravitational field. Thus String Theory should be considered as a theory of quantum gravity. Remarkably, this idea did not have too many followers for about a decade. Only a handful of physicists seriously pursued this idea - most notably Green and Schwarz [10].

In the mid 1980's there was a resurgence of interest Kaluza Klein theories. However, very soon Witten realized that the usual Kaluza Klein theories cannot describe chiral fermions after compactification. This was the time when some physicists turned to Superstring Theory as a candidate for a unified theory of all fundamental forces with gravity.

There was, however, one problem. If the theory is quantum mechanically consistent, there should be no anomalies. In a landmark paper in the summer of 1984, Green and Schwarz indeed proved that when the gauge groups of superstring theories are chosen appropriately, all anomalies cancel. It then became clear that superstring theory is a serious candidate for a theory which can incorporate quantum gravity. In the next few years, a large number of physicists switched to research in String Theory. The development of two dimensional conformal field theory by Belavin, Polyakov and Zamolodchikov led to introduction of new powerful tools, leading to rapid progress [11].

I was among the numerous converts to String Theory. I was a postdoc at Fermilab and would drive down to Chicago often to discuss with Nambu. I recall asking Nambu about his thoughts about these new developments. He certainly thought that these developments were interesting and important, but did not display much enthusiasm. I later realized that this was quite characteristic of Nambu. He was never really attached emotionally to some of his own creations.

7 Non-critical Strings and Emergent Space

If the large N expansion of field theory of matrices gives rise to some kind of String Theory, and consistent string theories describe gravity it would be natural to guess that there are matrix field theories which describe quantum gravity. However explicit realizations of this idea took a while to materialize, and needed one key insight - the emergence of space dimensions in this story. We already saw that Eguchi-Kawai models provided a way of manufacturing space-time dimensions from large N matrix models. The kind of emergence we will now discuss will be somewhat different, and needed some insights from the theory of noncritical strings.

The first ingredient involves the physical understanding of the Polyakov path integral (9). This is usually regarded as the propagator for a non-critical string in d dimensions, since the worldsheet fields X^μ are naturally thought of as the coordinates of a point on the string in d dimensions. However the path integral also involves an integration over the worldsheet metric - what does that signify ? This became clear soon after the work of Knizhnik, Polyakov and Zamolodchikov, David and of Distler and Kawai - especially in the latter two papers. The metric has one degree of freedom, the Liouville mode. Thus the path integral for a bosonic string is actually over $(d+1)$ scalar fields,

$$\int \mathcal{D}\phi \mathcal{D}X^\mu \, \exp[-\frac{1}{2}\int d^2\xi \, \sqrt{\hat{h}} \left[\hat{h}^{ab} \left(\partial_a\phi\partial_b\phi + \partial_a X^\mu \partial_b X_\mu\right) + Q\hat{R}\phi + \mu e^\phi\right] \qquad (10)$$

where $\hat{h}$ denotes a fiducial metric, $Q = \sqrt{\frac{25-d}{3}}$ and $\hat{R}$ is the Ricci scalar for $\hat{h}$. μ is a worldsheet cosmological constant. In 1988 Satchi Naik, Spenta Wadia and I realized that this means that the Liouville mode ϕ should be regarded as an additional dimension in space-time. For general $d \neq 25$ the target space theory does not have translation invariance in the ϕ direction. In fact the term $Q\hat{R}\phi$ means that in the target space there is a dilaton which is linear in this direction. For $d = 25$, the field ϕ is at par with the other X^μ - there is a $(d+1)$ dimensional Poincare invariance. This is the 26 dimensional critical string. For $d < 25$ it is natural to identify ϕ as a space coordinate, and for $d > 25$ a time coordinate [12].

The second ingredient was the realization by several groups - Brezin and Kazakov, Douglas and Shenker and Gross and Migdal - that the matrix models whose Feynman diagrams are described by the Polyakov path integral with $d = 0$ can be in fact used to provide a non-pertubative definition of these string theories. The matrix model would be of the type (3), and one needs a double scling limit where the couplings contained in the potential $V(M)$ are tuned to specific critical values, while $N \to \infty$ with a combination of a power of N and the departure of the coupling from the critical value held fixed. In this limit, the Feynman diagrams which form discretizations of two dimensional surfaces of arbitrary genus contribute to the same order, which is why this is non-perturbative in the string coupling [13].

This kind of double scaling limit was soon extended to Matrix Quantum Mechanics described by (4) by Brezin and Kazakov, Ginsparg and Zinn-Justin and Gross and Mikovich. The Feynman digrams of this theory would have vertices which are points in time - these should then lead to a Polyakov path integral of the type (9) where the index μ runs over one value, time. However, we just saw that this should actually describe a string theory not in one dimension, but in two dimensions. How does that happen in matrix quantum mechanics ?

In a work with Antal Jevicki in 1990 I considered this problem using the collective field formalism developed by Antal and Sakita in the 1980's. We are interested in the singlet sector of a hamiltonian of the form (5). In this sector the dynamical variables are the eigenvalues of the matrix $\lambda_i(t)$, and at large N this can be in turn rewritten in terms of the density of eigenvalues

$$\rho(x,t) = \frac{1}{N}\sum_{i=1}^{N}\delta(x-\lambda_i(t)) \tag{11}$$

called the collective field, and its canonically conjugate momentum. Now, $\rho(x,t)$ looks like a two dimensional field. x is like an additional dimension which comes from the internal matrix degrees of freedom. However, there is no reason why the collective hamiltonian $H[\rho(x),\pi_\rho(x)]$ would be local in this variable x. Nevertheless, it turns out that in the double scaling limit the hamiltonian is indeed local in the leading order of large-N expansion - it represents a self interacting massless scalar field in $1+1$ dimensions with a coupling which depends on x. In fact, in their original paper Brezin et.al. had shown that the eigenvalues become coordinates of fermions moving on a line. Sengupta and Wadia and Gross and Klebanov showed that the collective field is in fact the bosonization of the corresponding fermionic field. This is precisely what one expects if this matrix quantum mechanics indeed describes two dimensional string theory. This string theory is very simple: the only propagating degree of freedom is a massless scalar field. In fact, the collective field can related directly to this scalar field. In addition to this dynamical mode, there are infinite tower of modes which can provide backgrounds - pretty much like the longitudinal mode in electrodynamics. At low energies these modes consist of a metric and a dilaton. Dilaton-gravity system in $1+1$ dimensions cannot have wavelike solutions, but can have nontrivial solutions. This is a theory containing gravity, even though there are no gravitational waves. A series of subsequent papers established in great detail that the matrix model reproduces the S Matrix of the string theory and it was manifest that scattering of the massless scalar involve gravitational interactions [14].

This model is the earliest example of what we now call the holographic correspondence. A quantum field theory in $0+1$ dimensions - Matrix Quantum Mechanics - is dual to a $1+1$ dimensional theory which contains gravity (together with other fields). The large-N theory has no notion of space. The internal degrees of freedom of matrix quantum mehanics have metamorphosed into an emergent space dimension.

If one could do something like this for a theory of matrices which are themselves fields in some number of dimensions one would be able to directly construct higher dimensional string theories. Unfortunately such a direct approach does not work beyond the example we considered. Further progress came from a rather different area of physics - the physics of black holes.

8 Black Holes in String Theory [15]

Ever since Hawking discovered that black holes radiate almost thermally due to quantum effects, physicists have been trying to make sense of black hole thermodynamics and of the apparent puzzles posed by Hawking radiation. Normally, thermodynamics is a coarse-grained description of a system which has a microscopic structure. However, for black holes it has been traditionally difficult to come up with a microscopic structure - this seemed to be at odds with No Hair Theorems.

This situation changed soon after the “Second String Revolution” in 1995. Following Sen’s breakthrough work showing that $N=4$ Yang-Mills theory is self-dual, Witten showed that all the known string theories can be in fact dual to each other and therefore describe a single theory, whose proper description is in terms of “M-theory” - a yet unknown theory whose low energy limit is 11 dimensional supergravity.

Duality symmetries have been explored by practitioners of supergravity for quite a while, but were usually ignored by the string theory community. These symmetries were known to lead to higher dimensional extended objects - membranes, 3 branes, 5 branes and so on. Soon Polchinski showed that a class of such higher dimensional objects called D-branes have to be included in the spectrum of string theories.

The interesting thing about D-branes is that their excitations are described by open strings whose ends are stuck to the brane, but otherwise free to oscillate in space. And of course the low energy limit of open string theory is a gauge theory. Witten soon showed that this low energy gauge theory of a stack of N such coincident D-branes is in fact $SU(N)$ Yang-Mills theories. The Yang-Mills coupling is related to the string coupling by $g_s = g_{YM}^2$.

D-branes have a tension which is proportional to $\frac{1}{g_s}$, so that the mass of a stack of N D-branes will be proportional to $\frac{N}{g_s}$. Thus the gravitational field produced by such a stack will be proportional to $\frac{G_N N}{g_s}$ where G_N is Newton's gravitational constant. However G_N is proportional to g_s^2 so the field is proportional to $g_s N = g_{YM}^2 N$. Now consider the classical limit of string theory, i.e. $g_s \to 0$. For a fixed N this would mean that the gravitational field vanishes. However if we consider the limit $N \to 0$ together with $g_s \to 0$ with $g_s N = g_{YM}^2 N$ held fixed, we have a classical limit where the stack of D-branes produces a finite gravitational field. However this is precisely the ‘t Hooft limit of the Yang-Mills theory ! Furthermore, when the ‘t Hooft coupling $\lambda = g_{YM}^2 N$ is large one should be able to describe this stack of D branes in terms of its gravitational field.

In fact, such a large number of D branes would behave exactly like a black hole. This is pretty much like what happens in electrodynamics. At the microscopic level a charged object is made out of electrons and described by the theory of quantum electrodynamics. However this is not a useful way to describe the object when there are 10^{25} electrons. A more useful way is to describe this in terms of the electric field produced - the solution of Gauss' Law. Similarly while the microscopic open string theory, or Yang-Mills theory at low energies, is a useful description when there are few D branes, it is not very useful when we have a huge number of D-branes. Rather we should describe these in terms of its gravitational field - the black hole metric which is a solution of Einstein's equation.

We therefore arrive at examples of black holes whose microscopic structure is known - and described by Yang Mills theory. The idea that sufficiently excited states of strings describe black holes is in fact much older than the realization that D-branes describe black holes. The interest in black holes among high energy theorists exploded after the discovery that the two dimensional string theory we discussed in the previous section have black hole solutions. Mandal, Sengupta and Wadia found black hole solutions of the low energy equations of motion and at the same time Witten showed that there is an exact formulation of strings moving in such a black hole background. Very soon Callan, Giddings, Harvey and Strominger pioneered the use of two dimensional black holes to understand Hawking radiation. A lot of subtle issues were clarified in the next few years, but a sharp resolution of the black hole information paradox was not in sight. In particular, there was no microscopic understanding of black hole thermodynamics.

In 1993 Susskind suggested that one should really think of black holes as excited states of strings - if one knew how to count the number of such states Ω at some given energy one could use Boltzmann's formula $S = \log \Omega$ to have a microscopic understanding of Bekenstein-Hawking entropy. This is a daunting task, since one has to calculate this degeneracy at strong couplings. However, as suggested by Vafa, this should be possible for extremal black holes in supergravity where non-renormalization theorems would allow the extrapolation of a weak coupling formula to strong coupling.

The first such calculation was performed by Sen for black hole like objects which are described by strings wound around a compact direction with waves moving purely in one direction. Such objects were studied earlier by Dabholkar and Harvey. The extremal version of these did not have finite area horizons. However, they have a generalized notion of a horizon, called a stretched horizon. Rather surprisingly Sen showed that the degeneracy of states predicted by string theory is proportional to the area of the stretched horizon. The proportionality constant cannot be determined since the definition of a stretched horizon is not very precise. Nevertheless, this was the first indication that a count of states can indeed account for Bekenstein-Hawking entropy.

Extremal black holes do not radiate - they have zero temperature. To understand Hawking radiation one needs to go away from extremality by exciting the extremal black holes. I and Samir Mathur figured out how to describe these excitations for the kind of systems which Sen analyzed. However a precise understanding of Bekenstein-Hawking formula was not in sight.

This is where D-branes made a huge difference. In General Relativity there are extremal black holes which have large horizons - in fact the usual Reissner-Nordstrom black holes of Einstein-Maxwell theory are of this type. They have zero temperature, but non-zero entropy. In early 1996 Strominger and Vafa showed how to describe such Reissner-Norsdtrom black holes as configurations of D-branes. They were able to count the degeneracy of extremal states for given charges. As in Sen's calculation, this can be done at weak coupling and extrapolated to strong coupling. The result was stunning - the answer was in exact agreement with the Bekenstein Hawking formula - and in this case the horizon had a finite size and the Bekenstein Hawking formula was reliable. Very soon this agreement was extended to slightly non-extremal black holes by several authors.

What about Hawking radiation ? Callan and Maldacena charted out the mechanism which would lead to such a radiation. In the microscopic theory of D branes these would be two open strings joining to form a closed string which could then leave the brane and shoot off to infinity. In the late 1970's Unruh and Page had calculated the rate of Hawking radiation from a black hole by performing a classical calculation of absorption of waves by a black hole and using equilibrium condition to infer the radiation rate. For Schwarzchild black holes they found that at low energies the absorption cross section is equal to the area. Dhar, Mandal and Wadia showed that this continues to be the case for the black holes in String Theory considered by Strominger and Vafa. And the results of Callan and Maldacena did indicate that a microscopic calculation would also lead to a result proportional to the area. However, it was not at all clear how a precise calculation can be done. In the summer of 1996 I and Mathur realized that the closed string modes which account for most of the low energy absorption or radiation have universal couplings to the open string modes on the D-branes, so that a precise calculation is possible. The result was surprising: the weak coupling microscopic calculation of Hawking radiation *exactly* agreed with the classical calculation ! These D-brane configurations were indeed black holes !

Soon afterward, Gubser, Klebanov and Tseytlin found that this agreement is not restricted to black holes. They looked at stack of 3-branes whose geomtery is horizonless and completely smooth. The microscopic theory of N such branes is a $3+1$ dimensional maximally supersymmetric $SU(N)$ gauge theory. They showed that a calculation in this theory precisely reproduces the classical calculation of absorption of waves in the three brane geometry.

In all these calculations it was clear that a theory on D-branes is able to describe gravitational physics in higher dimensions. For p-branes in 10 dimensional superstring theory, the microscopic theory is a $p+1$ dimensional large N gauge theory. However this seems to contain all the information about the gravitational theory which of course lives in the entire 10 dimeensional space. This looks very similar to Matrix Quantum Mechanics - a $0+1$ dimensional large N theory - describing a $1+1$ dimensional string theory as we had found earlier. Unlike this latter example, there is no analog of collective field theory which explicitly demonstrates this emergence of additional dimensions.

It was also realized that this kind of connection is a manifestation of open string closed string duality. From the early days of string theory it has been known that processes involving open strings can be alternatively described in terms of closed strings, though they were both thought to live in the same number of dimensions. In such dualities, however, the low energy limit of the open string theory is usually never equivalent to a low energy limit of the closed string theory - rather the open string low energy limit involves all the higher string modes. However, in these D-brane examples it appeared that the low energy theory on the D-branes - Yang Mills theory - can reproduce results in General Relativity which is the low energy of the higher dimensional closed string theory. How does this happen ?

9 The AdS/CFT correspondence [16]

In a landmark paper in 1997, Maldacena explained this puzzle: this insight led to our current understanding of emergent space. He argued that the key point is a large gravitational redshift in the horizon geometry. This means that if one takes a low energy limit with the energy defined as usual in the asymptotic region, one automatically retains all the higher closed string modes near the horizon. Thus in this low energy limit, there is a "duality" between the Yang-Mills theory describing the physics of D-branes and closed string theory in the near-horizon region. A stack of 3-branes in fact provides the simplest set-up. The near horizon geometry is $AdS_5 \times S^5$ while the low energy theory of the branes is $N = 4$ super-Yang-Mills in $3 + 1$ dimensions. In the 't Hooft large N limit of the Yang Mills theory there is an alternative description in terms of closed string theory in ten dimensional $AdS_5 \times S^5$. This is the simplest example of the AdS/CFT correspondence. Four of these dimensions were present as the space-time dimensions of the Yang-Mills theory. The remaining six dimensions were emergent - they arose of the internal degrees of freedom of the fields of the Yang Mills theory. This is pretty much like the earlier example of the two dimensional string: the $N = 4$ Yang Mills theory plays the role of matrix quantum mechanics. However, unlike the 2d string/ matris quantum mechanics duality there is as yet no explicit construction of the closed string theory from the Yang-Mills.

After many twists and turns, 't Hooft's idea found a concrete realization in this setting. Furthermore there is one additional aspect which makes this correspondence even more useful. The curvature of the $AdS_5 \times S^5$ in units of the string length is proportional to $(g_{YM}^2 N)^{-1/2}$, and unlike usual QCD the coupling g_{YM} of the $N = 4$ theory is a free parameter. Therefore, when the Yang-Mills theory is strongly coupled, the space-time on which the closed strings move is weakly curved. This means that one can safely replace the complicated closed string theory by its simpler low energy effective field theory - General Relativity with additional massless fields. Remarkably, in this limit classical gravity is equivalent to the highly quantum strongly coupled gauge theory ! Furthermore not one, but six additional dimensions of space-time are manufactured from the internal degrees of freedom of the Yang-Mills theory.

There are in fact examples where *all* the spatial directions of the gravitational theory are manufactured, though five of them are compact. One interesting example concerns a collection of N D0 branes. The theory of these branes is the quantum mechanics of nine coordinates $X^i(t)$ and their supersymmetric partners. The gravitational theory now lives in 9+1 dimensions, with all the spatial directions non-compact.

In fact, prior to Maldacena's work Banks, Fischler, Shenker and Susskind had argued that in a suitable light front frame this theory in fact describes 11 dimensional M theory in a suitable light front frame. Around the same time, Ishibashi, Kawai, Kitazawa and Tshuchiya conjectured that the euclidean theory of D-instantons, which is a matrix model of 10 matrices and their supersymmetric partners give rise to Type IIB string theory in ten euclidean dimensions by a mechanism similar to the Eguchi-Kawai models.

Unlike its $c = 1$ Matrix Model predecessor, there is no explicit understanding how the internal degrees of freedom metamorphose into additional dimensions for most string theory examples of the AdS/CFT correspondence. However the AdS/CFT correspondence is more general than these string theory examples. There is one instance which is particularly interesting. This concerns models of fields which transform as fundamental representations of gauge groups like $SU(N)$ or $SO(N)$. These are vector models and serve as very useful models in the study of critical phenomena. In 2002, Klebanov and Polyakov argued that such models are dual not to string theories, but to a class of higher spin theories which include gravity developed earlier by Vasiliev. Soon after this paper, I and Jevicki [17] found the mechanism which leads to this. The invariant operators in vector models with fields $\phi^i(\mathbf{x})$ are bi-local fields

$$\sigma(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^{N} \phi^i(\mathbf{x}) \phi^i(\mathbf{y}) \tag{12}$$

where $\mathbf{x}, \mathbf{y}$ are points in d dimensional space-time. In terms of the center of mass and relative coordinates

$$\mathbf{u} = \frac{1}{2}(\mathbf{x} + \mathbf{y}) \qquad \mathbf{v} = (\mathbf{x} - \mathbf{y}) \tag{13}$$

one can decompose the bi-local as follows

$$\sigma(\mathbf{x}, \mathbf{y}) = \sum_{l, m_a} \sigma_{l, m_a}(\mathbf{u}, r) Y_{l, m_a}(\theta_a) \tag{14}$$

where $r^2 = \mathbf{v}^2$ and $Y_{l,m_a}(\theta_a)$ are spherical harmonics on S^{d-1}. The $\sigma_{l,m_a}(\mathbf{u}, r)$ then appears as a spin-l field in $(d + 1)$ dimensions. It turns out that when the vector model is conformal, a field re-definition maps these $\sigma_{l,m_a}(u, r)$ into spin-l fields in AdS_{d+1}. This is precisely the content of Vasiliev theory, with the spin-2 field being the graviton. In addition at leading order in $1/N$ there are an infinite tower of higher spin massless fields. It turns out, however, that these σ_{l,m_a} are not themselves the fields of Vasiliev theory. Rather they are related to the latter by a field redefinition which has been now derived.

10 The world as a hologram

Even before the discovery of Hawking radiation, Bekenstein had shown that a black hole needs to be assigned an entropy propotional to the horizon area. Hawking's work established the proportionality constant to be $1/4G_N$. In 1981 Bekesntein wrote an intruiguing paper arguing for a universal upper bound of the entropy to energy ratio of *any* system - and this argument can be used to show that in any theory of gravity, the maximum possible entropy of any system is bounded by the area of a surface which encloses it. This is the Bekenstein bound. In 1993 ‘t Hooft and Susskind turned this argument around in an interesting way. The idea is the following. We know that if there is no gravity, the entropy of a system is extensive, i.e. proportional to its volume. ‘t Hooft and Susskind argued that this Bekenstein bound would naturally follow if one postulates that a theory containing gravity is equivalent to a theory *without gravity* which lives on the boundary. Thus the boundary theory acts as a hologram of everything which happens in the interior, which is why this proposal is called the Holographic Principle.

The AdS/CFT correspondence provided a concrete realization of this principle. This was realized by Gubser, Klebanov and Polyakov and by Witten (GKPW) [16]. The metric of $(d+1)$ dimensional AdS space-time can be written as

$$ds^2 = \frac{R^2}{z^2}[-dt^2 + dz^2 + d\mathbf{x}^2] \tag{15}$$

where $0 \leq z \leq \infty$ and $\mathbf{x}$ denote $(d-1)$ spacelike directions in R^{d-1}. We will call this space-time “bulk”. This space-time has a physical boundary at $z = 0$. Consider some set of bulk fields $\phi_i(t, z, \mathbf{x})$. In the semiclassical regime this fields satisfy the classical equations of motion. For example if ϕ is a scalar field with mas m, it satisfies the Klein Gordon equation. An analysis of this equation then shows that as we approach the boundary $z = 0$ the solution behaves as follows

$$\phi(z, t, \mathbf{x}) \to z^{\Delta}\phi_1(t, \mathbf{x})[1 + O(z^2)] + z^{d-\Delta}\phi_0(t, \mathbf{x})[1 + O(z^2)] \tag{16}$$

’ where Δ is

$$\Delta = \frac{1}{2}[d + \sqrt{d^2 + 4m^2}] \tag{17}$$

Then, GKPW argued that the AdS/CFT correspondence implies that such a solution corresponds to a state of a field theory defined on the boundary which has an action

$$S = S_{CFT} + \int d^d x \ \phi_0(t, \mathbf{x}) \ \mathcal{O}_{\Delta}(t, \mathbf{x}) \tag{18}$$

where S_{CFT} is the action of a conformal field theory and $\mathcal{O}_{\Delta}(t, \mathbf{x})$ is an operator with conformal dimension Δ. In this state

$$\langle \mathcal{O}_{\Delta}(t, \mathbf{x}) \rangle = \phi_1(t, \mathbf{x}) \tag{19}$$

Therefore, in this limit where the gravity theory is classical, the AdS/CFT correspondence provides a precise relationship between a field theory on the boundary of AdS with the gravitational theory in the bulk. The field theory is on a space-time $(t, \mathbf{x})$ whereas the bulk has one additional dimension z. In the CFT there is a conformal symmetry $SO(d-1, 2)$ - this is indeed the group of isometries of AdS_{d+1}. z is the emergent space dimension which arose out of the internal degrees of freedom in the CFT. This therefore provides a concrete realization of the holographic principle.

11 Back to QCD strings

It would be a pity if this kind of holographic correspondence is valid only for field theories which are strictly conformal. Fortunately there are examples when the field theories have mass gaps - the higher dimensional space-times are then not purely AdS. These examples of the holographic correspondence can be in fact used to understand the nature of the QCD string.

In the AdS/CFT correspondence it turns out that the emergent dimension is in fact the renormalization group scale of the field theory. As we proceed from the boundary $z = 0$ to large values of z, the field theory flows from the UV to the IR along a RG flow. The information about a point at a location z is then encoded in a region of the boundary which has roughly a size also given by z - further the point from the boundary, larger the size of the region in the hologram.

QCD is asymptotically free, i.e. the theory flows out of a conformal fixed point. Thus the geometry of the dual theory must be asymptotically AdS near the boundary. However since QCD has a dynamically generated mass gap, the geometry must differ significantly from AdS far from the boundary. In fact, there are many examples of the holographic correspondence where the dual geometry is like this - these represent quantum field theories with a dynamically generated mass gap.

These examples illuminate one of the main reasons why the earlier attempts to construct string theory of QCD failed. Consider a heavy quark-antiquark pair in the Yang-Mills theory on the boundary. The string or flux tube which joins them should have the minimal length. However, the minimal length string is no longer restricted to the boundary but will go into the interior of the AdS or the deformed AdS. If the Yang Mills theory is conformal the bulk is pure AdS and it turns out that this string goes all way to the IR. The relationship between the z coordinate and the scale in the field theory means that from the point of view of the field theory, this string is an infinitely fat string. The lines of force are spread out, leading to a Coulomb law. This is exactly what one would expect in a conformal theory. In a theory like QCD, the bulk is capped off at some value of $z = z_0 \sim 1/m_g$ where m_g is the mass gap. Therefore the string has to turn back at this value of z. In the boundary field theory this would mean that the string has a width which is roughly $1/m_g$. Indeed this is what one expects in QCD. And of course the fact that the string must be thick was realized long time ago. Coming up with a mathematical description of a thick string is of course much more difficult, though there has been a lot of recent progress.

In the new light of holography, we have learnt that one should try to construct an elementary string in one higher dimension. In the correct geometry this would automatically lead to a thick string. This is the main new insight. We now know that the QCD string is in fact the same string as the string which provides a theory of gravity - only the background is different.

While the qualitative aspects of the above discussion are clear and convincing, it is not yet possible to use this framework to perform a reliable quantitative calculation in large-N QCD. Over the years, there are many examples of AdS/CFT where the field theory has many of the essential features of QCD - though these theories are not exactly QCD. It turns out that for the non-supersymmetric cases, in the continuum limit (i.e. where all the physical masses are much smaller than the UV cutoff) the curvature of the dual geometry becomes large. This implies that a supergravity approximation is no longer valid. AdS/CFT has been spectacularly useful in the situations where the bulk theory may be approximated by supergravity, and unfortunately this is not the case for QCD like theories. Recall that at large N, we only need a classical string theory which is dual to QCD. Therefore one needs to construct a consistent classical string theory of elementary strings in these backgrounds. So far this has not been possible, though there has been a lot of progress towards this goal.

It is quite likely that such a string theory can be constructed in the forseeable future. Then we will be closer to the holy grail: one would have a back of the envelope calculation of key quantities in QCD, albeit at large N.

12 A note

I had the great fortune of being a student of Nambu. He taught me to think of physics as a unified field, free of boundaries. And he shaped my taste in physics. I have been lucky to be able to be in touch with him and meet him regularly till his last days, and every time I came back with excitement, still wondering how can someone think like that.

13 Acknowledgements

I would like to thank Bindu Bambah for organizing this conference and bringing out a centennial volume. I am honored to make this contribution to this volume. I would also like to thank my many collaborators over the years in areas related to the subject of this article: Antal Jevicki, John Kogut, Gautam Mandal, Samir Mathur, Sandip Trivedi and Spenta Wadia. This work was partially supported by a National Science Foundation grant NSF/PHY-2111673 and the Jack and Linda Gill Foundation.

References

1. New York Times, July 18, 2015
2. For an account of the early days of String Theory with references to the original literature see e.g. J. H. Schwarz, "String theory: The Early years," [arXiv:hep-th/0007118 [hep-th]]
3. For an excellent early review see S. Coleman, "1/N" , Erice Subnuclear 1979:0011 Contribution to: 17th International School of Subnuclear Physics: Pointlike Structures Inside and Outside Hadrons, 0011.
4. For a comprehensive collection of original literature on large-N field theories, see E. Brezin and S. R. Wadia, "The Large N expansion in quantum field theory and statistical physics: From spin systems to two-dimensional gravity," (World Scientific, 1993)
5. Y. Nambu, "QCD and the String Model," Phys. Lett. B **80**, 372-376 (1979) doi:10.1016/0370-2693(79)91193-6
6. For a review and references, see A. A. Migdal, "Loop Equations and 1/N Expansion," Phys. Rept. **102**, 199-290 (1983) doi:10.1016/0370-1573(83)90076-5.;
   Papers reprinted in Section 5 of [4]
7. See the papers reprinted in Section 8 of [4].
8. S. R. Das, "Some Aspects of Large N Theories," Rev. Mod. Phys. **59**, 235 (1987) doi:10.1103/RevModPhys.59.235
9. For a review of this activity and references, see V. A. Kazakov and A. A. Migdal, "Recent Progress in the Theory of Noncritical Strings," Nucl. Phys. B **311**, 171 (1988) doi:10.1016/0550-3213(88)90146-0.
10. See e.g. T. Yoneya, "Gravity from strings: Personal reminiscence on early developments," [arXiv:0901.0079 [hep-th]].
11. J. H. Schwarz, "The second superstring revolution," [arXiv:hep-th/9607067 [hep-th]].
12. This point of view is reviewed in S. R. Das, "THE LIOUVILLE MODE AND STRINGS," TIFR/TH/89/38.
13. For a review of such matrix models, see P. Di Francesco, P. H. Ginsparg and J. Zinn-Justin, "2-D Gravity and random matrices," Phys. Rept. **254**, 1-133 (1995) doi:10.1016/0370-1573(94)00084-G [arXiv:hep-th/9306153 [hep-th]].
14. For reviews, see S. R. Das, "The one-dimensional matrix model and string theory," [arXiv:hep-th/9211085 [hep-th]];
    I. R. Klebanov, "String theory in two-dimensions," [arXiv:hep-th/9108019 [hep-th]];
    P. H. Ginsparg and G. W. Moore, "Lectures on 2-D gravity and 2-D string theory," [arXiv:hep-th/9304011 [hep-th]];
    A. Jevicki, "Development in 2-d string theory," doi:10.1142/9789814447072_0004 [arXiv:hep-th/9309115 [hep-th]].
15. S. R. Das and S. D. Mathur, "The quantum physics of black holes: Results from string theory," Ann. Rev. Nucl. Part. Sci. **50**, 153-206 (2000) doi:10.1146/annurev.nucl.50.1.153 [arXiv:gr-qc/0105063 [gr-qc]];

J. R. David, G. Mandal and S. R. Wadia, "Microscopic formulation of black holes in string theory," Phys. Rept. **369**, 549-686 (2002) doi:10.1016/S0370-1573(02)00271-5 [arXiv:hep-th/0203048 [hep-th]].

16. A relatively recent book with a comprehensive bibliography which is accessible to a wide spectrum of physicists is M. Ammon and J. Erdmenger, "Gauge/gravity duality: Foundations and applications,". ;
For a slightly older review, see J. Polchinski, "Introduction to Gauge/Gravity Duality," doi:10.1142/9789814350525_0001 [arXiv:1010.6134 [hep-th]].
17. S. R. Das and A. Jevicki, "Large N collective fields and holography," Phys. Rev. D **68**, 044011 (2003) doi:10.1103/PhysRevD.68.044011 [arXiv:hep-th/0304093 [hep-th]];
For a recent acccount see R. de Mello Koch, A. Jevicki, K. Suzuki and J. Yoon, "AdS Maps and Diagrams of Bi-local Holography," JHEP **03**, 133 (2019) doi:10.1007/JHEP03(2019)133 [arXiv:1810.02332 [hep-th]];
O. Aharony, S. M. Chester and E. Y. Urbach, "A Derivation of AdS/CFT for Vector Models," JHEP **03**, 208 (2021) doi:10.1007/JHEP03(2021)208 [arXiv:2011.06328 [hep-th]].