Quantum Mechanics of Riemannian Geometry

Abhay Ashtekar
Center for Gravitational Physics and Geometry
University of Pennsylvania, University Park, USA

Abstract

Over the past four years, a detailed framework has been constructed to unravel the quantum nature of the Riemannian geometry of physical space. A review of these developments is presented at a level which should be accessible to advanced undergraduate students in physics. As an illustrative application, I indicate how this micro-structure of geometry can have a direct impact on physical processes such as the evaporation of black holes through the Hawking process.

1 Introduction

During his Göttingen inaugural address in 1854, Riemann [1] suggested that geometry of space may be more than just a fiducial, mathematical entity serving as a passive stage for physical phenomena, and may in fact have direct physical meaning in its own right. General relativity provided a brilliant confirmation of this vision: curvature of space now encodes the physical gravitational field. This shift is profound. To bring out the contrast, let me recall the situation in Newtonian physics. There, space forms an inert arena on which the dynamics of physical systems —such as the solar system— unfolds. It is like a stage, an unchanging backdrop for all of physics. In general relativity, by contrast, the situation is very different. Einstein's equations tell us that matter curves space. Geometry is no longer immune to change. It reacts to matter. It is dynamical. It has "physical degrees of freedom" in its own right. Thus, in general relativity, the stage disappears and joins the troupe of actors. Geometry is a physical entity, very much like matter.

Now, the physics of this century has shown us that matter has constituents and the 3-dimensional objects we perceive as solids are in fact made of atoms. The continuum description of matter is an approximation which succeeds brilliantly in the macroscopic regime but fails hopelessly at the atomic scale. It is therefore natural to ask: Is the same true of geometry? If so, what is the analog of the 'atomic scale?' We know that a quantum theory of geometry will feature three fundamental constants of Nature, $c, G, \hbar$, the speed of light, Newton's gravitational constant and Planck's constant. Now, as Planck pointed out in his celebrated paper that marks the beginning of quantum mechanics, there is a unique combination, $\ell_P = \sqrt{\hbar G/c^3}$, of these constants which has dimension of length. ($\ell_P \approx 10^{-33}$cm.) It is now called the Planck length. Experience has taught us that the presence of a distinguished scale in a physical theory marks a potential transition; physics below the scale can be very different from that above the scale. Now, all of our well-tested physics occurs at length scales much bigger than than ℓ_P. In this regime, the continuum picture works well. A key question then is: Will it break down at the Planck length? Does geometry have constituents at this scale? If so, what are its atoms? Its elementary excitations? Is the space-time continuum only a 'coarse-grained' approximation? Is geometry quantized? If so, what is the nature of its quanta?

To probe such issues, it is natural to look for hints in the procedures that have been successful in describing matter. Let us begin by asking what we mean by quantization of physical quantities. Take a simple example —the hydrogen atom. In this case, the answer is clear: while the basic observables —energy and angular momentum— take on a continuous range of values classically, in quantum mechanics their eigenvalues are discrete; they are quantized. So, we can ask if the same is true of geometry. Classical geometrical quantities such as lengths, areas and volumes can take on continuous values on the phase space of general relativity. Are the eigenvalues of corresponding quantum operators discrete? If so, we would say that geometry is quantized and the precise eigenvalues and eigenvectors of geometric operators would reveal its detailed microscopic properties.

Thus, it is rather easy to pose the basic questions in a precise fashion. Indeed, they could have been formulated soon after the advent of quantum mechanics. Answering them, on the other hand, has proved to be surprisingly difficult. The main reason, I believe, is the inadequacy of the standard techniques. More precisely, to examine the microscopic structure of geometry, we must treat Einsteinian gravity quantum mechanically, i.e., construct at least the basics of a quantum theory of the gravitational field. Now, in the traditional approaches to quantum field theory, one *begins* with a continuum, background geometry. To probe the nature of quantum geometry, on the other hand, we should *not* begin by assuming the validity of this picture. We must let quantum gravity decide whether this picture is adequate; the theory itself should lead us to the correct microscopic model of geometry.

With this general philosophy, in this article I will summarize the picture of quantum geometry that has emerged from a specific approach to quantum gravity. This approach is non-perturbative. In perturbative approaches, one generally begins by assuming that space-time geometry is flat and incorporates gravity —and hence curvature— step by step by adding up small corrections. In the non-perturbative approach, by contrast, there is no background metric at all. All we have is a bare manifold to start with. All fields —matter as well as gravity/geometry— are treated as dynamical from the beginning. Consequently, the description cannot refer to a background metric. Technically this means that the full diffeomorphism group of the manifold is respected; the theory is generally covariant.

As we will see, this fact leads one to Hilbert spaces of quantum states which are quite different from the familiar Fock spaces of particle physics. Now gravitons —the three-dimensional wavy undulations on a flat metric— do not represent fundamental excitations. Rather, the fundamental excitations are *one*-dimensional. Microscopically, geometry is rather like a polymer. Recall that, although polymers are intrinsically one-dimensional, when densely packed in suitable configurations they approximate a three-dimensional system. Similarly, the familiar continuum picture of geometry arises as an approximation. Indeed, one can regard the fundamental excitations as 'quantum threads' and construct from them 'weave states' which approximate continuum geometries. Gravitons are no longer the basic mediators of the gravitational interaction. They now arise only as approximate notions; they represent perturbations of weave states. Because states are polymer-like, geometrical observables turn out to have discrete spectra. They provide a rather detailed picture of quantum geometry from which physical predictions can be made.

The article is divided into two parts. In the first, I will indicate how one can reformulate general relativity so that it resembles gauge theories. This formulation provides the starting point for the quantum theory. In particular, the one-dimensional excitations of geometry arise as the analogs of 'Wilson loops' which are themselves analogs of the line integrals $\exp i \oint A.d\ell$ of electro-magnetism. In the second part, I will indicate how this description leads us to a quantum theory of geometry. I will focus on area operators and show how the detailed information about the eigenvalues of these operators has interesting physical consequences, e.g., to the process of Hawking evaporation of black holes.

I should emphasize that this is *not* a technical review. Rather, the article is written at the level of colloquia in physics departments in the United States. Thus, I will purposely avoid technicalities and try to make the discussion intuitive. I will also make some historic detours of general interest. At the end, however, I will list some references where the details of the central results can be found.

2 From metrics to connections

2.1 Gravity versus other fundamental forces

General relativity is normally regarded as a dynamical theory of metrics —tensor fields that define distances and hence geometry. It is this fact that enabled Einstein to code the gravitational field in the Riemannian curvature of the metric. Let me amplify with an analogy. Just as position serves as the configuration variable in particle dynamics, the three-dimensional metric of space can be taken to be the configuration variable of general relativity. Given the initial position and velocity of a particle, Newton's laws provide us with a trajectory of particle in the position space. Similarly, given a three-dimensional metric and its time derivative at an initial instant, Einstein's equations provide us with a four-dimensional space-time which can be regarded as a trajectory in the space of 3-metrics[1].

However, this emphasis on the metric sets general relativity apart from all other fundamental forces of Nature. Indeed, in the theory of electro-weak and strong interactions, the basic dynamical variable is a (matrix-valued) vector potential, or a connection. Like general relativity, these theories are also geometrical. The connection enables one to parallel-transport objects along curves. In electrodynamics, the object is a charged particle such as an electron; in chromodynamics, it is a particle with internal color, such as a quark. Generally, if we move the object around a closed loop, we find that its state does not return to the initial value; it is rotated by an unitary matrix. In this case, the connection is said to have curvature and the unitary matrix is a measure of the curvature

[1]Actually, only six of the ten Einstein's equations provide the evolution equations. The other four do not involve time-derivatives at all and are thus constraints on the initial values of the metric and its time derivative. However, if the constraint equations are satisfied initially, they continue to be satisfied at all times.

in a region enclosed by the loop. In the case of electrodynamics, the connection is determined by the vector potential and the curvature by the electro-magnetic field strength.

Since the metric also gives rise to curvature, it is natural to ask if there is a relation between metrics and connections. The answer is in the affirmative. Every metric defines a connection —called the Levi-Civita connection of the metric. The object that the connection enables one to parallel transport is a vector. (It is this connection that determines the geodesics, i.e. the trajectories of particles in absence of non-gravitational forces.) It is therefore natural to ask if one can not use this connection as the basic variable in general relativity. If so, general relativity would be cast in a language that is rather similar to gauge theories and the description of the (general relativistic) gravitational interaction would be very similar to that of the other fundamental interactions of Nature. It turns out that the answer is in the affirmative. Furthermore, both Einstein and Schrödinger gave such a reformulation of general relativity. Why is this fact then not generally known? Indeed, I know of no textbook on general relativity which even mentions it. One reason is that in this formulation the basic equations are somewhat complicated —but not much more complicated, I think, than the standard ones in terms of the metric. A more important reason is that we tend to think of distances, light cones and causality as fundamental. These are directly determined by the metric and in a connection formulation, the metric is a 'derived' rather than a fundamental concept. But in the last few years, I have come to the conclusion that the real reason why the connection formulation of Einstein and Schrödinger has remained so obscure lies in an interesting historical episode. I will return to this point at the end of this section.

2.2 Metrics versus connections

Modern day researchers re-discovered connection theories of gravity after the invention and successes of gauge theories for other interactions. Generally, however, these formulations lead one to theories which are quite distinct from general relativity and the stringent experimental tests of general relativity often suffice to rule them out. There is, however, a reformulation of standard general relativity whose basic equations, furthermore, are simpler than the standard ones: while Einstein's equations are non-polynomial in terms of the metric and its conjugate momentum, they turn out to be low order polynomials in terms of the new connection and its conjugate momentum. Furthermore, just as the simplest particle trajectories in space-time are given by geodesics, the 'trajectory' determined by the time evolution of this connection according to Einstein's equation turns out to be a geodesic in configuration space of connections.

In this formulation, the phase space of general relativity is identical to that of the Yang-Mills theory which governs weak interactions. Recall first that in electrodynamics, the (magnetic) vector potential constitutes the configuration variable and the electric field serves as the conjugate momentum. In weak interactions and general relativity, the configuration variable is a matrix-valued vector potential; it can be written as $\vec{A}_i\tau_i$ where $\vec{A}_i$ is a triplet of vector fields and τ_i are the Pauli matrices. The conjugate momenta are represented by $\vec{E}_i\tau_i$ where $\vec{E}_i$ is a triplet of vector fields[2]. Given a pair $(\vec{A}_i, \vec{E}_i)$ (satisfying appropriate conditions as noted in footnote 1), the field equations of the two theories determine the complete time-evolution, i.e., a dynamical trajectory.

The field equations —and the Hamiltonians governing them— of the two theories are of course very different. In the case of weak interactions, we have a background space-time and we can use its metric to construct the Hamiltonian. In general relativity, we do not have a background metric. On the one hand this makes life very difficult since we do not have a fixed notion of distances or causal structures; these notions are to arise from the solution of the equations we are trying to write down! On the other hand, there is also tremendous simplification: Because there is no background metric, there are very few mathematically meaningful, gauge invariant expressions of the Hamiltonian that one can write down. (As we will see, this theme repeats itself in the quantum theory.) It is a pleasant surprise that the simplest non-trivial expression one can construct from the connection and its conjugate momentum is in fact the correct one, i.e., is the Hamiltonian of general relativity! The expression is at most quadratic in $\vec{A}_i$ and at most quadratic in $\vec{E}_i$. The similarity with gauge theories opens up new avenues for quantizing general relativity and the simplicity of the field equations makes the task considerably easier.

What is the physical meaning of these new basic variables of general relativity? As mentioned before, connections tell us how to parallel transport various physical entities around curves. The Levi-Civita connection tells us how to parallel transport vectors. The new connection, $\vec{A}_i$, on the other hand, determines

[2]A summation over the repeated index i is assumed. Also, technically each $\vec{A}_i$ is a 1-form rather than a vector field. Similarly, each $\vec{E}_i$ is a vector density of weight one, i.e., natural dual of a 2-form

the parallel transport of *left handed spin-$\frac{1}{2}$ particles* (such as neutrinos) —the so called *chiral fermions*. These fermions are mathematically represented by spinors which, as we know from elementary quantum mechanics, can be roughly thought of as 'square roots of vectors'. Not surprisingly, therefore, this connection is not completely determined by the metric alone. It requires additional information which roughly is a square-root of the metric, or a tetrad. The conjugate momenta $\hat{E}_i$ represent restrictions of these tetrads to space. They can be interpreted as spatial triads, i.e., as 'square-roots' of the metric of the 3-dimensional space. Thus, information about the Riemannian geometry of space is coded directly in these momenta. The (space and) time-derivatives of the triads are coded in the connection.

To summarize, there *is* a formulation of general relativity which brings it closer to theories of other fundamental interactions. Furthermore, in this formulation, the field equations simplify greatly. Thus, it provides a natural point of departure for constructing a quantum theory of gravity and for probing the nature of quantum geometry non-perturbatively.

2.3 Historical detour

To conclude this section, let me return to the piece of history involving Einstein and Schrödinger that I mentioned earlier. In the forties, both men were working on unified field theories. They were intellectually very close. Indeed, Einstein wrote to Schrödinger saying that he was perhaps the only one who was not 'wearing blinkers' in regard to fundamental questions in science and Schrödinger credited Einstein for inspiration behind his own work that led to the Schrödinger equation. During the years 1946-47, they had periods of intense correspondence on unified field theory and, in particular, on the issue of whether connections should be regarded as fundamental or metrics. Einstein was in Princeton and Schrödinger in Dublin. But starting January 1946, they exchanged their ideas and latest results very frequently. In fact the dates on their letters often show that the correspondence was going back and forth with astonishing speed. It reveals how quickly they understood the technical material the other had sent, how they hesitated, how they teased each other. Here are a few quotes:

The whole thing is going through my head like a millwheel: To take Γ [the connection] *alone as the primitive variable or the* g*'s* [metrics] *and* Γ*'s ? ...*

—Schrödinger, May 1st, 1946.

How well I understand your hesitating attitude! I must confess to you that inwardly I am not so certain ... We have squandered a lot of time on this thing, and the results look like a gift from devil's grandmother.
—Einstein, May 20th, 1946

Einstein was expressing doubts about using the Levi-Civita connection alone as the starting point which he had advocated at one time. Schrödinger wrote back that he laughed very hard at the phrase 'devil's grandmother'. In another letter, Einstein called Schrödinger 'a clever rascal'. Schrödinger was delighted and took it to be a high honor. This continued all through 1946. Then, in the beginning of 1947, Schrödinger thought he had made a breakthrough. He wrote to Einstein:

Today, I can report on a real advance. Maybe you will grumble frightfully for you have explained recently why you don't approve of my method. But very soon, you will agree with me...
—Schrödinger, January 26th, 1947

Schrödinger sincerely believed that his breakthrough was revolutionary [3]. Privately, he spoke of a second Nobel prize. The very next day after he wrote to Einstein, he gave a seminar in the Dublin Institute of Advanced Studies. Both the Taoiseach (the Irish prime minister) and newspaper reporters were invited. The day after, the following headlines appeared:

Twenty persons heard and saw history being made in the world of physics. ... The Taoiseach was in the group of professors and students. ..[To a question from the reporter] *Professor Schrödinger replied "This is the generalization. Now the Einstein theory becomes simply a special case ..."*
—*Irish Press*, January 28th, 1947

Not surprisingly, the headlines were picked up by *New York Times* which obtained photocopies of Schrödinger's paper and sent them to prominent physicists —including of course Einstein— for comments. As Walter Moore, Schrödinger's biographer puts it, Einstein could hardly believe that such grandiose claims had been made based on a what was at best a small advance in an area of work that they both had been pursuing for some time along parallel lines. He prepared a carefully worded response to the request from *New York Times:*

[3]The 'breakthrough' was to drop the requirement that the (Levi-Civita) connection be symmetric, i.e., to allow for torsion.

It seems undesirable to me to present such preliminary attempts to the public. ... Such communiqués given in sensational terms give the lay public misleading ideas about the character of research. The reader gets the impression that every five minutes there is a revolution in Science, somewhat like a coup d'état in some of the smaller unstable republics. ...

Einstein's comments were also carried by the international press. On seeing them, Schrödinger wrote a letter of apology to Einstein citing his desire to improve the financial conditions of physicists in the Dublin Institute as a reason for the exaggerated account. It seems likely that it only worsened the situation. Einstein never replied. He also stopped scientific communication with Schrödinger.

The episode must have been shocking to those few who were exploring general relativity and unified field theories at the time. Could it be that this episode effectively buried the desire to follow up on connection formulations of general relativity until an entirely new generation of physicists who were blissfully unaware of this episode came on the scene?

3 Quantum geometry

3.1 General setting

Now that we have a connection formulation of general relativity, let us consider the problem of quantization. Recall first that in the quantum description of a particle, states are represented by suitable wave functions $\Psi(\vec{x})$ on the configuration space of the particle. Similarly, quantum states of the gravitational field are represented by appropriate wave functions $\Psi(\vec{A}_i)$ of connections. Just as the momentum operator in particle mechanics is represented by $\hat{P} \cdot \Psi_I = -i\hbar\,(\partial\Psi/\partial x_I)$ (with $I = 1, 2, 3$), the triad operators are represented by $\hat{\vec{E}}_i \cdot \Psi = \hbar G(\delta\Psi/\delta\vec{A}_i)$. The task is take geometric quantities such as lengths of curves, areas of surfaces and volumes of regions, express them in terms of triads using ordinary differential geometry and then promote these expressions to well-defined operators on the Hilbert space of quantum states. In principle, the task is rather similar to that in quantum mechanics where we first express observables such as angular momentum or Hamiltonian, express them in terms of configuration and momentum variables, $\vec{x}, \vec{p}$ and then promote them to quantum theory as well-defined operators on the quantum Hilbert space.

In quantum mechanics, the task is relatively straightforward; the only potential problem is the choice of factor ordering. In the present case, by contrast, we are dealing with a *field theory*, i.e., a system with an infinite number of degrees of freedom. Consequently, in addition to factor ordering, we face the much more difficult problem of regularization. Let me explain qualitatively how this arises. A field operator, such as the triad mentioned above, excites infinitely many degrees of freedom. Technically, its expectation values are distributions rather than smooth fields. They don't take precise values at a given point in space. To obtain numbers, we have to integrate the distribution against a test function, which extracts from it a 'bit' of information. As we change our test or smearing field, we get more and more information. (Take the familiar Dirac δ-distribution $\delta(x)$; it does not have a well-defined value at $x = 0$. Yet, we can extract the full information contained in $\delta(x)$ through the formula: $\int \delta(x) f(x) dx = f(0)$ for all test functions $f(x)$.) Thus, in a precise sense, field operators are distribution-valued. Now, as is well known, product of distributions is not well-defined. If we attempt naively to give meaning to it, we obtain infinities, i.e., a senseless result. Unfortunately, all geometric operators involve rather complicated (in fact non-polynomial) functions of the triads. So, the naive expressions of the corresponding quantum operators are typically meaningless. The key problem is to regularize these expressions, i.e., to extract well-defined operators from the formal expressions in a coherent fashion.

3.2 Geometric operators

This problem is not new; it arises in all physically interesting quantum field theories. However, as I mentioned in the Introduction, in other theories one has a background space-time metric and it is invariably used in a critical way in the process of regularization. For example, consider the electro-magnetic field. We know that the energy of the Hamiltonian of the theory is given by $H = \int \vec{E} \cdot \vec{E} + \vec{B} \cdot \vec{B}\, d^3x$. Now, in the quantum theory, $\hat{\vec{E}}$ and $\hat{\vec{B}}$ are both operator-valued distributions and so their square is ill-defined. But then, using the background flat metric, one Fourier decomposes these distributions, identifies creation and annihilation operators and extracts a well-defined Hamiltonian operator by normal ordering, i.e., by physically moving all annihilators to the right of creators. This procedure removes the unwanted and unphysical infinite zero point energy from the formal expression and the subtraction makes the operator well-defined. In the present case, on the other hand, we are trying to construct a quantum theory of geometry/gravity and do not have a flat metric —or indeed, any metric— in the background. Therefore, many of the standard regularization techniques are no longer available.

Fortunately, however, between 1992 and 1995, a new functional calculus was developed on the space of connections $\vec{A}_i$ —i.e., on the configuration space of the theory. This calculus is mathematically rigorous and makes no reference at all to a background space-time geometry; it is generally covariant. It provides a variety of new techniques which make the task of regularization feasible. First of all, there is a well-defined integration theory on this space. To actually evaluate integrals and define the Hilbert space of quantum states, one needs a measure: given a measure on the space of connections, we can consider the space of square-integrable functions which can serve as the Hilbert space of quantum states. There is, however, a preferred measure, singled out by the physical requirement that the (gauge invariant versions of the) configuration and momentum operators be self-adjoint. This measure is diffeomorphism invariant and thus respects the underlying symmetries coming from general covariance. Thus, there is a natural Hilbert space of states to work with [4]. Let us denote it by $\mathcal{H}$. Differential calculus enables one to introduce physically interesting operators on this Hilbert space and regulate them in a generally covariant fashion. As in the classical theory, the absence of a background metric is both a curse and a blessing. On the one hand, because we have very little structure to work with, many of the standard techniques simply fail to carry over. On the other hand, at least for geometric operators, the choice of viable expressions is now severely limited, which greatly simplifies the task of regularization.

The general strategy is the following. The Hilbert space $\mathcal{H}$ is the space of square-integrable functions $\Psi(\vec{A}_i)$ of connections $\vec{A}_i$. A key simplification arises because it can be obtained as the (projective) limit of Hilbert spaces associated with systems with only a finite number of degrees of freedom. More precisely, given any graph γ (which one can intuitively think of as a 'floating lattice') in the physical space, using techniques which are very similar to those employed in lattice gauge theory, one can construct a Hilbert space $\mathcal{H}_\gamma$ for a quantum mechanical system with $3N$ degrees of freedom, where N is the number of edges of the graph. Roughly, these Hilbert spaces know only about how the connection parallel transports chiral fermions along the edges of the graph and not elsewhere. That

[4] This is called the kinematical Hilbert space; it enables one to formulate the quantum Einstein's (or supergravity) equations. The final, physical Hilbert space will consist of states which are solutions to these equations.

is, the graph is a mathematical device to extract $3N$ 'bits of information' from the full, infinite dimensional information contained in the connection, and $\mathcal{H}_\gamma$ is the sub-space of $\mathcal{H}$ consisting of those functions of connections which depend only on these $3N$ bits. (Roughly, it is like focussing on only $3N$ components of a vector with an infinite number of components and considering functions which depend only on these $3N$ components, i.e., are constants along the orthogonal directions.) To get the full information, we need all possible graphs. Thus, a function of connections in $\mathcal{H}$ can be specified by specifying a function in $\mathcal{H}_\gamma$ for *every* graph γ in the physical space. Of course, since two distinct graphs can share edges, the collection of functions on $\mathcal{H}_\gamma$ must satisfy certain consistency conditions. These lie at the technical heart of various constructions and proofs.

The fact that $\mathcal{H}$ is the (projective) limit of $\mathcal{H}_\gamma$ breaks up any given problem in quantum geometry into a set of problems in quantum mechanics. Thus, for example, to define operators on $\mathcal{H}$, it suffices to define a *consistent family of* operators on $\mathcal{H}_\gamma$ for each γ. This makes the task of defining geometric operators feasible. I want to emphasize, however, that the introduction of graphs is only for technical convenience. Unlike in lattice gauge theory, we are not *defining* the theory via a continuum limit (in which the lattice spacing goes to zero.) Rather, the full Hilbert space $\mathcal{H}$ of the continuum theory is already well-defined. Graphs are introduced only for practical calculations. Nonetheless, they bring out the one-dimensional character of quantum states/excitations of geometry. It is because 'most' states in $\mathcal{H}$ can be realized as elements of $\mathcal{H}_\gamma$ for some γ that quantum geometry can be regarded as polymer-like.

Let me now outline the result of applying this procedure for geometric operators. Suppose we are given a surface S, defined in local coordinates by $x_3 = \text{const.}$ The classical formula for the area of the surface is: $A_S = \int d^2x \sqrt{E_i^3 E_i^3}$, where E_i^3 are the third components of the vectors $\vec{E}_i$. As is obvious, this expression is non-polynomial in the basic variables $\vec{E}_i$. Hence, off-hand, it would seem very difficult to write down the corresponding quantum operator. However, thanks to the background independent functional calculus, the operator can in fact be constructed rigorously.

To specify its action, let us consider a state which belongs to $\mathcal{H}_\gamma$ for *some* γ. Then, the action of the final, regularized operator $\hat{A}_S$ is as follows. If the graph has no intersection with the surface, the operator simply annihilates the state. If there are intersections, it acts at each intersection via group theory. *This simple form is a direct consequence of the fact that we do not have a background* *geometry*: given a graph and a surface, the diffeomorphism invariant information one can extract lies in their intersections. To specify the action of the operator in detail, let me suppose that the graph γ has N edges. Then the state Ψ has the form: $\Psi(\vec{A}_i) = \psi(g_1, ...g_N)$ for some function ψ of the N variables $g_1, ..., g_N$, where g_k ($\in SU(2)$) denotes the spin-rotation that a chiral fermion undergoes if parallel transported along the k-th edge using the connection $\vec{A}_i$. Since g_k represent the possible rotations of spins, angular momentum operators have a natural action on them. In terms of these, we can introduce 'vertex operators' associated with each intersection point v between S and γ:

$$\hat{O}_v \cdot \Psi(A) = \sum_{I,J} k(I,L)\vec{J}_I \cdot \vec{J}_L \cdot \psi(g_1, ..., g_N) \tag{1}$$

where I, L run over the edges of γ at the vertex v, $k(I,J) = 0, \pm 1$ depending on the orientation of edges I, L at v, and $\vec{J}_I$ are the three angular momentum operators associated with the I-th edge. (Thus, $\vec{J}_I$ act only on the argument g_I of ψ and the action is via the three left invariant vector fields on $SU(2)$.) *Thus, the vertex operators resemble the Hamiltonian of a spin system, $k(I,L)$ playing the role of the coupling constant.* The area operator is just a sum of the square-roots of the vertex operators:

$$\hat{A}_S = \frac{G\hbar}{2c^3} \sum_v |O_v|^{\frac{1}{2}} \tag{2}$$

Thus, the area operator is constructed from angular momentum-like operators. Note that the coefficient in front of the sum is just $\frac{1}{2}\ell_P^2$, the square of the Planck length. This fact will be important later.

Because of the simplicity of these operators, their complete spectrum —i.e., full set of eigenvalues— is known explicitly: Possible eigenvalues a_S are given by

$$a_S = \frac{\ell_P^2}{2} \sum_v \left[2j_v^{(d)}(j_v^{(d)}+1) + 2j_v^{(u)}(j_v^{(u)}+1) - j_v^{(d+u)}(j_v^{(d+u)}+1)\right]^{\frac{1}{2}} \tag{3}$$

where v labels a finite set of points in S and $j^{(d)}, j^{(u)}$ and $j^{(d+u)}$ are non-negative half-integers assigned to each v, subject to the usual inequality

$$j^{(d)} + j^{(u)} \geq j^{(d+u)} \geq |j^{(d)} - j^{(u)}| \,. \tag{4}$$

Thus the entire spectrum is discrete; *areas are indeed quantized!* This discreteness holds also for the length and the volume operators. Thus the expectation that the continuum picture may break down at the Planck scale is borne out fully. Quantum geometry is *very* different from the continuum picture. This may be the fundamental reason for the failure of perturbative approaches to quantum gravity.

Let us now examine a few properties of the spectrum. The lowest eigenvalue is of course zero. The next lowest eigenvalue may be called the *area gap*. Interestingly, area-gap is sensitive to the topology of the surface S. If S is open, it is $\frac{\sqrt{3}}{4}\ell_P^2$. If S is a closed surface —such as a 2-torus in a 3-torus— which fails to divide the spatial 3-manifold into an 'inside' and an 'outside' region, the gap is larger, $\frac{2}{4}\ell_P^2$. If S is a closed surface —such as a 2-sphere in R^3— which divides space into an 'inside' and an 'outside' region, the area gap is even larger; it is $\frac{2\sqrt{2}}{4}\ell_P^2$. Another interesting feature is that in the large area limit, the eigenvalues crowd together. This follows directly from the form of eigenvalues given above. Indeed, one can show that for large eigenvalues a_S, the difference Δa_S between consecutive eigenvalues goes as $\Delta a_S \leq (exp - \sqrt{a_S/\ell_P^2})\ell_P^2$. Thus, Δa_S goes to zero very fast. (The crowding is noticeable already for low values of a_S. For example, in the case of trivial topology, there is only one non-zero eigenvalue with $a_S < 0.5\ell_P^2$, seven with $a_S < \ell_P^2$ and 98 with $a_S < 2\ell_P^2$.) Intuitively, this explains why the continuum limit works so well.

3.3 Physical consequences: details matter!

We will now see that if Δa_S had failed to vanish sufficiently fast, one would have been forced to conclude that the semi-classical approximation to quantum gravity must fail in an important way. To bring out this point, let me backtrack a bit. Let us consider not the most general eigenstates of the area operator $\hat{A}_S$ but —as was first done chronologically— the simplest ones. These correspond to graphs which have simple intersections with S. For example, n edges of the graph may just pierce S, each one separately, so that at each vertex there is just a straight line passing through. For these states, the eigenvalues are $a_S = (\sqrt{3}/2)n\ell_P^2$. Thus, here, the level spacing is uniform, like that of the Hamiltonian of a simple harmonic oscillator. Even if we restrict ourselves to the simplest eigenstates, even for large eigenvalues, the level spacing does not go to zero. Suppose for a moment that this is the *full* spectrum of the area operator. Then, as I will indicate below, Hawking's semi-classical derivation of black hole evaporation would have been incorrect. That is, the effects coming from area quantization would have implied that even for large macroscopic black holes of, say, a thousand solar masses, we can not trust semi-classical arguments.

Let me explain this point in some detail. The original derivation of Hawking's was carried out in the framework of quantum field theory in curved space-times which assumes that there is a specific underlying continuum space-time and explores the effects of the curvature of this space-time on quantum matter fields. In this approximation, Hawking found that the black hole geometries are such that there is a spontaneous emission which has a Planckian spectrum at infinity. Thus, black holes, seen from far away, resemble black bodies and the associated temperature turns out to be inversely related to the mass of the hole. Now, physically one expects that, as it evaporates, the black hole must lose mass. Since the radius of the horizon is proportional to the the mass, the area of the horizon must decrease. If one uses a classical picture for the underlying space-time, one would conclude that the process is continuous. However, if in a more fundamental theory of quantum gravity area is quantized, one would expect that the black hole evaporates in discrete steps by making a transition from one area eigenvalue to another, smaller one. The process would be very similar to the way an excited atom descends to its ground state through a series of discrete transitions.

Let us look at this process in some detail. For simplicity let us use units with $c = 1$. *Suppose, to begin with, that the level spacing of eigenvalues of the area operator is the naive one, i.e. with* $\Delta a_S = \sqrt{3}/2\ell_P^2$. Then, the fundamental theory would have predicted that the smallest frequency, ω_o of emitted particles would be given by $\hbar\omega_o = \Delta M \sim (1/G^2M)\Delta a_H \sim \hbar/GM$, since the area A_H of the horizon goes as G^2M^2. Thus, the 'true' spectrum would have emission lines only at frequencies $\omega = N\omega_o \sim N\omega_p$, for $N = 1, 2, \ldots$ corresponding to transitions of the black hole through N area levels. How does this compare with the Hawking prediction? As I mentioned above, according to Hawking's semi-classical analysis, the spectrum would be the same as that of a black body at temperature T given by $kT \sim \hbar/GM$, where k is the Boltzmann constant. Hence, the peak of this spectrum would appear at ω_p given by $\hbar\omega_p \sim kT \sim \hbar/GM$. But this is precisely the order of magnitude of the minimum frequency ω_o that would be allowed if the area spectrum were the naive one. Thus, in this case, a more fundamental theory would predict that the spectrum would not resemble a black-body spectrum. The most probable transition would be for $N = 1$ and so the spectrum would be peaked at ω_p as in the case of a black body. However, there would be no emission lines at frequencies low compared with ω_p; this part of the black body spectrum would be simply absent. The part of the spectrum for $\omega > \omega_p$ would also not be faithfully reproduced since the discrete lines with frequencies $N\omega_o$, with $N = 1, 2, ...$ would *not* be sufficiently near each other —i.e. crowded— to yield an approximation to the continuous black-body spectrum.

The situation is completely different for the correct, full spectrum of the area operator if the black hole is macroscopic, i.e., large. Then, as I noted earlier, the area eigenvalues crowd and the level spacing goes as $\Delta a_H \leq (\exp - \sqrt{a_H/\ell_P^2})\ell_P^2$. As a consequence, as the black hole makes transition from one area eigenvalue to another, it would emit particles at frequencies equal to or larger than $\sim \omega_p \exp - \sqrt{a_H/\ell_P^2}$. Since for a macroscopic black-hole the exponent is very large (for a solar mass black-hole it is $\sim 10^{71}$!) the spectrum would be well-approximated by a continuous spectrum and would extend well below the peak frequency. Thus, the precise form of the area spectrum ensures that, for large black-holes, the potential problem with Hawking's semi-classical picture disappears. Note however that as the black hole evaporates, its area decreases, it gets hotter and evaporates faster. Therefore, a stage comes when the area is of the order of ℓ_P^2. Then, there *would* be deviations from the black body spectrum. But this is to be expected since in this extreme regime one does not expect the semi-classical picture to continue to be meaningful.

This argument brings out an interesting fact. Since the Planck length ℓ_P is so small, one would have thought that even if the area spectrum were the naive one —with equal level spacing $\Delta a_S = (\sqrt{3}/2)\ell_P^2$— one would not run in to a problem with classical or semi-classical approximations while dealing with large, macroscopic objects. Indeed, there are several iconoclastic approaches to quantum geometry in which one simply begins by postulating that geometric quantities should be quantized. Then, having no recourse to first principles from where to derive the eigenvalues of these operators, one simply postulates them to be multiples of appropriate powers of the Planck length. For area then, one would say that the eigenvalues are integral multiples of ℓ_P^2. The above argument shows how this innocent looking assumption can contradict semi-classical results even for large black holes. In our case, we did not begin by postulating the nature of quantum geometry. Rather, we *derived* the spectrum of the area operator from first principles. As we see, the form of these eigenvalues is rather complicated and could not have been guessed apriori. More importantly, the detailed form does carry rich information and in particular removes the conflict with semi-classical results in macroscopic situations.

3.4 Future directions

Exploration of quantum Riemannian geometry continues. Last year, it was found that geometric operators exhibit certain unexpected non-commutativity. This reminds one of the features explored by Alain Connes in his non-commutative geometry. Indeed, there are several points of contact between these two approaches. For instance, the Dirac operator that features prominently in Conne's theory is closely related to the connection $\vec{A}_i$ used here. However, at a fundamental level, the two approaches are rather different. In Conne's approach, one constructs a non-commutative analog of entire differential geometry. Here, by contrast, one focuses only on Riemannian geometry; the underlying manifold structure remains classical. In three space-time dimensions, it is possible to get rid of this feature in the final picture and express the theory in purely combinatorial fashion. Whether the same will be possible in four dimensions remains unclear. However, combinatorial methods continue to dominate the theory and it is quite possible that one would again be able to present the final picture without any reference to an underlying smooth manifold.

Another promising direction for further work is to construct better and better candidates for 'weave states' which can be regarded as non-linear analogs of coherent states approximating smooth, macroscopic geometries. Once one has an 'optimum' candidate to represent Minkowski space, one would develop quantum field theory on these weave quantum geometries. Because the underlying basic excitations are one-dimensional, the 'effective dimension of space' for these field theories would be less than three. Now, in the standard continuum approach, we know that quantum field theories in low dimensions tend to be better behaved because their ultra-violet problems are softer. Hence, there is hope that these theories will be free of infinities. If they are renormalizable in the continuum, their predictions at large scales cannot depend on the details of the behavior at very small scales. Therefore, theories based on weaves would not only be finite but their predictions may well agree with those of renormalizable theories at the laboratory scale.

Another major direction of research is devoted to formulating and solving quantum Einstein's equations using the new functional calculus. Over the past year, there have been some exciting developments in this area. The methods developed there seem to be applicable also to supergravity theories. In the coming years, therefore, there should be further work in this area. Finally, since this quantum geometry does not depend on a background metric, it provides a natural arena for other problem, in particular, that of obtaining a background independent formulation of string theory.

Acknowledgments

The work summarized here is based on contributions from many researchers especially John Baez, Alejandro Corichi, Roberto DePitri, Chris Isham, Jerzy Lewandowski, Renate Loll, Don Marolf, Jose Mourao, Thomas Thiemann, Carlo Rovelli, Lee Smolin and José-Antonio Zapata. Special thanks are due to Jerzy Lewandowski for long-range collaboration. This work was supported in part by the NSF Grant PHY95-14240 and by the Eberly fund of the Pennsylvania State University.

References

[1] B. Riemann, *Über die Hypothesen, welche der Geometrie zugrunde liegen* (1854).

Monographs and Reviews on Non-perturbative Quantum Gravity:

[2] A. Ashtekar, *Lectures on Non-perturbative Canonical Gravity*, Notes prepared in collaboration with R. S. Tate. (World Scientific, Singapore, 1991).

[3] R. Gambini and J. Pullin, *Loops, Knots, Gauge Theories and Quantum Gravity.* (Cambridge University Press, Cambridge, 1996).

[4] A. Ashtekar, in *Gravitation and Quantizations*, ed B. Julia and J. Zinn-Justin (Elsevier, Amsterdam, 1995).

Background-independent Functional Calculus:

[5] A. Ashtekar and C. J. Isham, *Class. Quant. Grav.* **9**, 1433 (1992).

[6] A. Ashtekar and J. Lewandowski, "Representation theory of analytic holonomy $C^\star$ algebras", in *Knots and quantum gravity*, J. Baez (ed), (Oxford University Press, Oxford 1994).

[7] J. Baez, *Lett. Math. Phys.* **31**, 213 (1994); "Diffeomorphism invariant generalized measures on the space of connections modulo gauge transformations", hep-th/9305045, in the Proceedings of the conference on quantum topology, D. Yetter (ed) (World Scientific, Singapore, 1994).

[8] A. Ashtekar and J. Lewandowski, *J. Math. Phys.* **36**, 2170 (1995).

[9] D. Marolf and J. M. Mourão, *Commun. Math. Phys.* **170**, 583 (1995).

[10] A. Ashtekar and J. Lewandowski, J. Geo. & Phys. **17**, 191 (1995).

[11] J. Baez, "Spin network states in gauge theory", Adv. Math. (in press); "Spin networks in non-perturbative quantum gravity," pre-print gr-qc/9504036.

[12] C. Rovelli and L. Smolin, *Phys. Rev.* D 52 (1995) 5743-5759

[13] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão and T. Thiemann, *J. Math. Phys.* **36**, 6456 (1995).

[14] J. Baez and S. Sawin, *Functional integration on spaces of connections, preprint q-alg/9507023.*

[15] J. A. Zapata, *Combinatoric space from loop quantum gravity*, Penn State Pre-print (1997).

Geometric Operators

[16] A. Ashtekar, C. Rovelli and L. Smolin, *Phys. Rev. Lett.* **69**, 237 (1992).

[17] J. Iwasaki and C. Rovelli, *Int. J. Modern. Phys.* D **1**, 533 (1993); *Class. Quant. Grav.* **11**, 2899 (1994).

[18] C. Rovelli and L. Smolin, *Nucl. Phys.* B **442**, 593 (1995).

[19] A. Ashtekar, J. Lewandowski, D. Marolf, J. Mourão and T. Thiemann, *J. Funct. Analysis*, **135**, 519 (1996).

[20] R. Loll, *Phys. Rev. Lett.* **75**, 3084 (1995).

[21] A. Ashtekar, J. Lewandowski, *Class. Quant. Grav.* **14**, A55-A81 (1997).

[22] R. Loll, "Spectrum of the volume operator in quantum gravity", *Nucl. Phys. B* (to appear).

[23] A. Ashtekar, A. Corichi, J. Lewandowski and J. A. Zapata, *Quantum theory of geometry II: Non-commutativity of Riemannian structures*, Penn State Pre-print (1997).

[24] A. Ashtekar and J. Lewandowski, *Quantum theory of geometry III: Volume operators*, Penn State Pre-print (1997).

Black Hole Evaporation:

[25] S. W. Hawking, *Comun. Math. Phys.* **43**, 199 (1975).

[26] J. Bekenstein and V. F. Mukhanov, "Spectroscopy of quantum black holes", *Physics Letters* B **360**, 7 (1995).

[27] S. Fairhurst, *Properties of the Spectrum of the Area Operator* (unpublished Penn State Report)

[28] A. Ashtekar, to appear in: *Geometric Issues in the Foundation of Science*, edited by L. Mason et al (Oxford University Press).