Geometry of Spacetime and Mass in General Relativity

by Shing-Tung Yau*

Einstein's theory of general relativity is based on the desire to merge the newly developed theory of special relativity and Newton's theory of gravity. He accomplished this daunting task in 1915. Most physicists consider this to be the most creative work in science in the history of mankind. Let me now explain some part of this theory to you.

A very important ingredient is the concept of equivalence principle, the development of which had a long history:

Galileo used experiment to show that the acceleration of a test mass due to gravitation is independent of the amount of mass being accelerated.

Then in 1907, Einstein said:

> "We assume the complete physical equivalent of a gravitational field and a corresponding acceleration of the reference system. The gravitational motion of a small test body depends only on its initial position in spacetime and velocity, and not on its constitution. The outcome of any local experiment (gravitational or not) in a freely falling laboratory is independent of the velocity of the laboratory and its location in spacetime".

Hence Einstein realized that in the new theory of gravity that he would like to develop, the laws of gravity should be independent of the observers. But he needed a framework to build such a theory of gravity that can connect philosophy with observations.

Einstein's great work benefited from the help of many geometers. He, together with Grossmann, was student of the great geometer and physicist Minkowski. He also interacted with Levi-Civita, and eventually Hilbert and Noether.

But most importantly, Einstein owed his epoch-making contribution to the concept of space by the great 19th century mathematician Riemann.

Before Riemann, there were only three types of space: the Euclidean space, the sphere space, and the hyperbolic space which were all described by a single coordinate system.

This was very similar to the time of Newton where the universe was supposed to be static. Riemann, however, radically changed the notion of space in his famous essay "On the hypotheses which lie at the foundations of geometry" in 1854.

His space was totally different from the three spaces above, and it could exist without referring to a fixed coordinate system. He also knew that up to first order effect, we do not feel presence of curvature and therefore infinitesimally, the space should look like the flat Euclidean space. On the other hand, the second order effect of gravity should come from acceleration of the particles. Therefore our space should show curvature if it is used to describe dynamics of gravity.

We do not really know what the space should look like globally. On the other hand, our space should be general enough to allow many different observers without changing the essence of the physics of gravity. Observers can propagate information from one to another one.

Hence Riemann demanded that we can use a variety of different coordinate systems to observe the basic properties of the space. However, the only meaningful properties of space should be independent of choice of the coordinate systems. This point of view of space is very important because it is the crucial principle of equivalence in general relativity.

Riemann defined the concept of curvature in his introduction of abstract space. In fact, later development of gravitational field in general relativity is measured by the curvature, while the material distribu-

* Institute of High Energy Physics, CAS

tion is represented by a part of the curvature. The distribution of matter changes over time and so does the curvature.

The dynamics of curvature shows the effect of vibration of spacetime. And because of that, Einstein came to the conclusion that the gravitational wave, though small, should exist. In Einstein's equation, the gravitational field and the geometry of spacetime are inseparable, as a unified entity.

It is remarkable that already in 1854 in his speech, Riemann developed the new concept of space because of the need to understand physical phenomena. He even suggested that the smallest and greatest parts of space should be described in different ways. From a modern physical point of view, Riemann is looking for the possible structure of quantum space! Riemann once considered using discrete space to explain this problem.

Riemann started his scientific publication at the age of 25 and died of lung disease at the age of 39. Three years before his death, he went to Italy every year to escape the cold, thus affecting a number of Italian and Swiss geometers, including Christoffel, Ricci and Levi-Civita.

They generalized Riemann's ideas, defined tensors and connections rigorously, both of which were indispensable for general relativity and gauge field theory. Ricci introduced the Ricci curvature tensor, and proved that this tensor can produce a tensor that satisfies the conservation law. All of these works, accomplished by geometers in the mid-to-late 19th century, provided the most crucial tools for general relativity.

Einstein wrote a paper in 1934 entitled "Notes on the origin of the general theory of relativity" (see Mein Weltbild, Amsterdam: Querido Verlag), in which he reviews the development of general relativity.

The first stage, of course, is the special theory of relativity. In addition to Einstein himself, the main founders of this theory include Lorentz and Poincaré.

One of the most important results is that the distance is affected by time. But Einstein learned that the action at a distance between the special theory of relativity and Newton's theory of gravity is incompatible and must be rectified!

At first, physicists did not realize that the concept of space had undergone fundamental changes after the breakthrough of Riemann. They attempted to correct Newton's gravitational theory in the framework of three-dimensional space in line with the special relativity just discovered. This idea led Einstein to go astray three years!

Einstein said in the essay "Notes of the origin of the general theory of relativity" (pp. 286–287):

> I was of course acquainted with Mach's view, according to which it appeared conceivable that inertial resistance counteracts is not acceleration as such but acceleration with respect to the masses of the other bodies existing in the world. There was something fascinating about this idea to me, but it provided no workable basis for a new theory.
>
> The simplest thing was, of course, to retain the Laplacian scalar potential of gravity, and to complete the equation of Poisson in an obvious way by a term differentiated with respect to time in such a way that the special theory of relativity was satisfied. The law of motion of the mass point in a gravitational field had also to be adapted to the special theory of relativity. The path was not so unmistakably marked out here, since the inert mass of a body might depend on the gravitational potential. In fact, this was to be expected on account of the principle of the inertia of energy.
>
> These investigations, however, led to a result which raised my strong suspicions.
>
> The principle of equality of inertial and gravitational mass could now be formulated quite clearly as follows: In a homogeneous gravitational field all motions take place in the same way as in the absence of a gravitational field in relation to a uniformly accelerated coordinate system.
>
> If this principle held good for any events whatever (The "principle of equivalence"), this was an indication that the principle of relativity needed to be extended to coordinate systems in non-uniform motion with respect to each other, if we were to reach a natural theory of the gravitational fields. Such reflections kept me busy from 1908 to 1911…

When Einstein was a student in Zurich, he was taught by Minkowski, who was a great mathematician on a par with Hilbert and Poincaré. Minkowski once said "There was a lazy student in my class who had recently done an important work which I had come up with a geometric interpretation".

Minkowski learned physics from Helmholtz, J.J. Thomson and Heinrich Hertz. He held that because of a "preestablished harmony between mathematics and nature", geometry could be used a key to physical insight. He ascribes physical reality to the geometry of spacetime.

This lecture entitled "Space and Time" was delivered by Minkowski in the eightieth meeting of the Assembly of Natural Scientists and Physicians in Cologne in Sep. 21, 1908.

The ideas of space and time developed here were applied in a major work on the laws of electrodynamics by Minkowski "The fundamental equations for electromagnetic phenomena in moving bodies" published in 1908. (Minkowski died in 1909.)

Minkowski wrote:

> "The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality."

It may be interested to note that Minkowski acknowledged his concept of spacetime owes a great deal to Poincaré's work in 1906, where Poincaré noticed that by changing time to imaginary time, Lorentz transformations become rotations.

However, Poincaré did not think the four-dimensional representation has much physical significance. Even at 1908, Poincaré said that "The language of three dimensions seems the better fitted to our description of the world although this description can be rigorously made in another idiom."

This is very different from Minkowski's point of view where he said directly "The world in space and time in a certain sense is a four-dimensional, non-Euclidean manifold. In truth, we are dealing with more than merely a new conception of space and time. The claim is that it is rather a quite specific natural law, which, because of its importance – since it alone deals with the primitive concepts of all natural knowledge, namely space and time – can claim to be called the first of all laws of nature."

In his essay of 1908, Minkowski constructed a four-dimensional space, introducing a metric tensor

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2,$$

following Riemann, to give geometric meaning of special relativity. The Lorenzian group, the fundamental symmetric group of special relativity, became the group of isometries of this spacetime of Minkowski.

For the first time in history, we learned from Minkowski that we live in a four-dimensional spacetime. Hence in 1908, Einstein got the most important inspiration for general relativity from Minkowski: that he has to construct his new theory of gravity based of the fact that the space should be four-dimensional.

It is generally believed that the most important thing Einstein did in the year is his thought experiment. The thought experiment taught Einstein the importance of equivalence principle and the need of new geometry to exhibit gravity. He knew that he need a new concept of space to achieve this. The static space of Newtonian gravity is not adequate any more.

Why is Minkowski's article so important? Not only that there is a conceptual breakthrough from three-dimensional space to four-dimensional space, but also that only within a four-dimensional spacetime, gravity can have enough room to show its dynamical nature! Newton's theory of gravity is static, in that a function is sufficient to describe the phenomenon of gravity.

Minkowski spacetime gave the most important reason why we need a tensor to describe gravity. Tensor is a newly invented concept consisting of many functions which together can transform consistently so that the principle of equivalence is obeyed. Minkowski's tensor perfectly describes the special theory of relativity, but Einstein wanted to further combine Newtonian mechanics with Minkowski space, so his new theory of spacetime should be equal to Minkowski spacetime infinitesimally.

Hence when two points are very close to each other, up to first order, the gravity rule governing them should be the one of Minkowski spacetime. However, this is no more true when we count second order effect of gravity, curvature becomes important. At the time, physicists knew nothing about the notion of tensors (in fact, only a few geometers knew about tensor analysis.)

Einstein knew from the principle of equivalence that the new potential of gravity should depend on a point and the tangent vector of space at that point (velocity vector), but he has no idea what kind of mathematical tool can be used. So he asked his classmate Marcell Grossmann for help and finally figured out that the gravitational field should be described by a metric tensor. The tensor varies in spacetime, but at every point it can be approximated by a first-order Minkowski metric.

Grossmann is a geometer, who helped Einstein to do homeworks in geometry in Zurich. He went to the library and found the ideas of tensor. However, the idea of introducing metric tensor alone is not enough to describe the gravitational field. We need to know how to differentiate tensor in a curved space. We would like the result of differentiation is also independent of choice of coordinate system (the requirement of equivalence principle). This is the connection theory of Christoffel and Levi-Civita.

Einstein said in his memoirs of general relativity mentioned above that this was his first question, and was found to have been solved by Levi-Civita and Ricci. Einstein's second question was how to generalize Newton's law of gravitation in this new framework. Newton's equation is simple, that is, the second derivative of the gravitational potential is equal to matter density.

At that time, neither Einstein nor Grossmann knew how to differentiate metric tensors so that the result is still a tensor which is independent of the choice of coordinate. Grossmann, at Einstein's repeated requests, managed to find Ricci's work in the library.

It turned out that Ricci had already contracted Riemann's curvature tensor to a symmetric second-order tensor. It is denoted by R_{ij}. It has the same degree of freedom as the metric tensor and in a suitable coordinate system, can be the second derivative of metric tensor g_{ij}. Einstein immediately realized that it must be the left-hand side of the field equation, while the right-hand side is the tensor T_{ij} of the general matter distribution (in flat space, this tensor has been well studied.) Einstein and Grossmann proposed the equation $R_{ij} = T_{ij}$ in two articles published in 1912 and 1913. This equation is similar to Newton's equation $\Delta u = \rho$ where u is the gravitational potential and ρ is the matter density.

However when Einstein tried to solve this equation by an asymptotic approach, he did not recover the astronomical phenomena (e.g., light deflection, Mercury's anomalous perihelion shift) that he was trying to explain. This made him very frustrated.

In the following years, in order to explain astronomical phenomena, he tried to choose special coordinates, essentially giving up the preciously simple principle of equivalence. The many communications between him and Levi-Civita could not help either.

The following was written by Einstein in: Notes of the origin of the general theory of relativity (1934, pp. 288–289).

> I soon saw that the inclusion of non-linear transformations, as the principle of equivalence demanded, was inevitably fatal to the simple physical interpretation of the coordinates – i.e. that it could no longer be required that coordinate differences should signify direct results of measurement with ideal scales or clocks.
>
> I was much bothered by this piece of knowledge, for it took me a long time to see what coordinates at all meant in physics. I did not find the way out of this dilemma until 1912, and then it came to me as a result of the following consideration:
>
> A new formulation of the law of inertia had to be found which in case of the absence of a "real gravitational field" passed over into Galileo's formulation for the principle of inertia if an inertial system was used as coordinate system. Galileo's formulation amounts to this: A material point, which is acted on by no force, will be represented in four-dimensional space by a straight line, that is to say, by a shortest line, or more correctly, an extremal line.
>
> This concept presupposes that of the length of a line element, that is to say, a metric. In the special theory of relativity, as Minkowski had shown, this metric was a quasi-Euclidean one, i.e., the square of the "length" ds of a line element was a certain quadratic function of the differentials of the coordinates.
>
> If other coordinates are introduced by means of a non-linear transformation, ds^2 remains a homogeneous function of the differentials of the coordinates, but the coefficients of this function ($g_{\mu\nu}$) cease to be constant and become certain functions of the coordinates. In mathematical terms this means that physical (four-dimensional) space has a Riemannian metric.
>
> The timelike extremal lines of this metric furnish that law of motion of a material point which is acted on by no force apart from the forces of gravity. The coefficients ($g_{\mu\nu}$) of this metric at the same time describe the gravitational field with reference to the coordinate system selected. A natural formulation of the principle of equivalence had thus been found, the extension of which to any gravitational field whatever formed a perfectly natural hypothesis.
>
> The solution of the above-mentioned dilemma was therefore as follows: A physical significance attaches not to the differentials of the coordinates but only to the Riemannian metric corresponding to them. A workable basis had now been found for the general theory of relativity. Two further problems remained to be solved, however.
>
> 1. If a field-law is expressed in terms of the special theory of relativity, how can it be transferred to the case of a Riemannian metric?
>
> 2. What are the differential laws which determine the Riemannian metric (i.e., $g_{\mu\nu}$) itself?
>
> As for problem 2, its solution obviously required the construction (from the $g_{\mu\nu}$) of the differential invariants of the second order. We soon saw that these had already been established by Riemann (the tensor of curvature). We had already considered the right field-equation for gravitation two years before the publication of the general theory of relativity, but we were unable to see how they could be used in physics.

Einstein struggled from 1913 to 1915. It is amusing that the equation that Einstein and Grossmann wrote down was actually correct if there is no matter. Indeed, Schwarzschild was able to solve Einstein equation for a spherical star in 1916, right after Einstein and Hilbert wrote down the right field equation.

Schwarzschild solution assumed that there is no matter. And it was enough to calculate light bending due to the gravity of the sun. Therefore Einstein and Grossmann could have made the observation in 1913, if they found the exact spherical symmetric solution. Apparently Einstein got discouraged when his approximate solution did not give him the right answer compatible with the physical observation. He was very depressed and was attempting to use special coordinate and hence gave up the principle of equivalence. The following writing of him shows his frustration:

> On the contrary, I felt sure that they could not do justice to experience. Moreover I believed that I could show on general considerations that a law of gravitation invariant with respect to arbitrary transformations of coordinates was inconsistent with the principle of causality. These were errors of thought which cost me two years of excessively hard work, until I finally recognized them as such at the end of 1915, and after having ruefully returned to the Riemannian curvature, succeeded in linking the theory with the facts of astronomical experience.
>
> In the light of knowledge attained, the happy achievement seems almost a matter of course, and any intelligent student can grasp it without too much trouble. But the years of anxious searching in the dark, with their intense longing, their alternations of confidence and exhaustion and the final emergence into light – only those who have experienced it can understand that.

Let us now go back to what happened in the final stage of Einstein's work on general relativity. In the spring of 1915, he visited the great mathematician David Hilbert in Göttingen. Hilbert certainly knows geometry well, but above all, he is the founder of modern geometric invariant theory. He also gathered a large group of outstanding mathematicians in Göttingen. Some of them can be described in the following:

Felix Klein was a pioneer in classifying geometries by using symmetry groups, Hilbert's student Hermann Weyl was the founder of gauge field theory, along with Emmy Noether, the greatest female mathematician in history.

During the period of 1915 to 1918, Noether was developing her theory of current where one can use group of continuous symmetries to deduce equations of motions. (In general relativity, the continuous group of symmetry is the group of coordinate transformations.)

Einstein's visit was just at the right time! Hilbert discovered the Hilbert action in November of the same year, and deriving the correct gravitational equation quickly from this action. Upon hearing the news and receiving Hilbert's postcard on the equation, Einstein quickly got his equation, and based on this equation, deduced astronomical phenomena he had been trying to solve. The equation is

$$R_{ij} - \frac{R}{2} g_{ij} = T_{ij}.$$

The left hand side can be derived from the Hilbert action $\int R$ in a simple manner. This was actually known to Bianchi and Ricci in 1901.

In the beginning, Einstein was unhappy with Hilbert's priority. But Hilbert quickly declared that the work should belong entirely to Einstein, and that turned Einstein happy. This is an epoch-making work. Later generations of physicists and mathematicians should all pay their highest tribute to Einstein. But I shall hope history will remember the group of Geometers who helped Einstein achieved his great theory of gravity. Much of what I discussed here is written by Einstein himself. It is unfortunate that in that article, he did not mention the contribution of Hilbert.

Looking backward, the correct equation of motion derived by Hilbert and Einstein could have been found by Grossmann and Einstein in 1913. The left-hand side of the equation in 1913 consists of Ricci tensor while the right-hand side is the matter tensor. The right-hand side is familiar and it satisfies conservation law. But the left-hand side of the 1913 equation is only the Ricci tensor which does not satisfy the conservation law. Hence they cannot be equal.

The left-hand side should therefore be replaced by some form of curvature tensor that satisfies conservation law. This was actually found by Ricci using Bianchi identity. One simply subtracts the Ricci tensor by some multiple of the metric tensor by the trace of the Ricci tensor. If Einstein and Grossmann trust the beauty of geometry and tried to complete the equations based on its internal consistency, Einstein would not have to wait until 1915 to write down the right equations.

After completing the general theory of relativity, Einstein believed that the most basic part of physics should be guided by thought experiment and the elegance of mathematics. At the end of the article, he said that after finding the equation of general relativity, everything came so natural and so simple that it was a breeze for a capable scholar. However, before finding the truth, he tried his best, after years of hard work, suffered pains day and night, which was hard to tell. Einstein's work can be said to be the greatest scientific work ever undertaken by mankind.

The success of general relativity left us another daunting task to explain natural phenomena of gravity. The task is difficult because the system of equation is truly nonlinear and the background spacetime is changing dynamically. Physics of gravity does not give a precise description of the initial data or the boundary conditions of the complicated nonlinear system.

There is no global symmetry of the dynamical changing spacetime. Nonexistence of global time, or nonexistence of timelike translation symmetry, gave great difficulty to define many important physical quantities that we learned in Newtonian mechanics. Noether's theory of current allows us to define mass and linear momentum four-vector if we have timelike translation that preserves the system. But for a generic system in general relativity, continuous group of symmetry does not exist!

Nonexistence of continuous symmetries caused difficulties to define classical concepts such as mass, linear momentum and angular momentum that are fundamental in understanding physics of gravity. When we watch two neutron stars interacting with each other, we need to know the mass of each star and the binding energy of the whole system counting contributions from matter and gravity together. This problem arises in general relativity because the concept of energy density is not possible in this theory of gravity.

The reason is that if the density exists, it will depend only on the first order information of the potential of the gravity which is the metric tensor. Yet we can always find a coordinate system so that the first order differentiation of the metric tensor is zero at one point. This will mean that the energy density is zero.

Einstein already realized such questions one hundred years ago. He proposed a definition of energy based on a concept of pseudo-tensor drawing analog with the definition of Newtonian mechanics. This definition was clarified more precisely by the work of Arnowitt, Deser, and Misner in 1962. Nowadays it is called ADM mass.

This definition works well for isolated physical system of gravity where the total mass of the whole system is defined. From the point of view of Noether, this is natural because for an isolated physical system, we expect existence of asymptotic symmetry at infinity and the time translation at infinity captures the total energy of the system. This is a good definition of total energy. However, it captures the total energy only and there are detailed information of partial energy we need to explore.

The very important question went back to Einstein again. He proposed the concept of gravitation radiation: the vibration of spacetime will radiate wave which gives energy. The energy comes from the binding energy of the gravity of the system. This concept was clarified by Bondi-van der Burg-Metzner and Trautman where they defined a mass along some null hypersurface. Such a mass is called Bondi mass and it has a pleasant property that it decreases when the null hypersurface moves to the future.

The decreases of the Bondi mass is interpreted as the energy carried away by the gravitational radiation. The definition of Bondi mass is important as it describes the dynamics of spacetime. However, the definition presumes some structure of spacetime that depends on the dynamics of the Einstein equation.

Both ADM and Bondi mass are total mass in nature. It cannot capture mass of bodies that are interacting with some other bodies. An important case is how to define binding energy of two neutron stars interacting with each other. Hence we need a concept of quasilocal mass: Given a closed two-dimensional (spacelike) surface S in spacetime, what is the total energy it encloses?

If S is the boundary of a three-dimensional spacelike three-manifold M in spacetime, we like to measure the total mass enclosed by S within M. Since we like to make sure the energy to be conserved, the quantity that we want should depend only on the information of S in spacetime and independent of the choice of M. This is the conservation law for quasilocal mass.

The existence of such quantity has been a serious problem for a long time. The very first thing that Einstein and later workers in general relativity was interested in whether the total ADM mass for an isolated physical system is positive?

In fact, in 1957, Bondi and other well known physicists had a meeting and discussed the possibility of negative mass in general relativity. Einstein's theory could not tell whether this is possible or not. But if the total mass is negative, the system may collapse and it will mean Einstein's theory of gravity may create a rather undesirable effect of unstable system.

The positivity of ADM mass was proved by Schoen and myself in 1979, the full proof published in 1981. Our proof is more geometric in nature.

Subsequently, Witten gave a proof depending on Dirac operator which is more transparent to physicists. Shortly afterwards, Bondi mass was also proved to be positive and the state of affair for total mass of an isolated physical system in gravity is pretty satisfactory.

Schoen and I also used our method to prove in an effective way that when matter density is large enough, black hole will form. It is the first rigorous statement that black hole forms when matter density is large.

The concept of Black Hole was proposed by P.S. Laplace in 18th century who proposed that there can be object whose gravitational fields are so strong that even light can not escape. But nothing can be done about this proposal. In 1916, right after Einstein-Hilbert wrote down the Einstein equation, Karl Schwarzschild wrote down a solution

$$ds^2 = -\left(1-\frac{2m}{r}\right)dt^2 + \left(1-\frac{2m}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

This solution has a singular point when $r = 0$, where the curvature goes to infinity. Physical laws can not be interpreted at such a point.

Since this solution has a large group of symmetries, many physicists thought that such singularities would not occur in general. This view was held by Lifshitz, Khalatnikov and coworkers.

This view was changed in 1965 by the work of Penrose. He invented the concept of closed trapped surface Σ.

It is a spacelike two dimensional surface Σ such that the two families of null geodesics orthogonal to Σ are convergent at Σ. (The outgoing light rays are dragged back and converge.)

Penrose proved that existence of closed trapped surface implies that spacetime is incomplete. The theory is further developed by Hawking-Penrose.

Penrose founded closed trapped surface exists in the above Schwarzschild solution and such closed trapped surface exists in any spacetime which is close to Schwarzschild, the singularity will therefore be there for spacetime close to Schwarzschild.

However, Penrose and Hawking could not explain how closed trapped surface arise in general spacetime and by what mechanism.

It was in the paper of Schoen-Yau, existence of black hole due to condensation of matter (1983) that gave the derivation of existence of closed trapped surface from first physical principle. When matter density is large in a fixed region, closed trapped surface will form. The argument is mathematical rigorous based on the theory of general relativity.

The demonstration of existence of a good definition of quasilocal mass took a long time, after works of many people including Penrose, Hawking, Brown-York, Geroch, Bartnik, Horowitz and Shi-Tam.

Since it is supposed to be trivial for any closed surface in the flat Minkowski spacetime and yet nonnegative for general spacetime, it is a miracle that such a definition can exist which is compatible with the previous works of ADM and Bondi. Two important definitions were proposed: one is due to Robert Bartnik and the other due to Mu-Tao Wang and myself.

The quasilocal mass allows us to define binding energy related to binary black holes and is related to the energy of the gravitational radiation. The approach of Wang-Yau allows them to define quasilocal angular momentum with Po-Ning Chen. It helps to clarify the former definitions of total angular momentum.

By the works of Richard Schoen and his coauthors, we know the Bartnik mass is different from Wang-Yau mass. It would be interesting to know which one is more useful to describe physical dynamics of gravity.

The concept of quasilocal mass and angular momentum has opened a window on studying the physics and geometry of spacetime. A great deal more efforts need to be spent in their study.

The definitions are most successful for objects within an isolated physical system. It would still be useful to understand a more general situation including higher dimensional analogue. Rather intricate geometry are involved in the study of such concepts.

• I shall now discuss the quasilocal angular momentum and center of mass of Chen-Wang-Yau, which are defined based on the theory of optimal isometric embedding, and their applications to spatial and null infinity. This is based on joint work with Po-Ning Chen, Jordan Keller, Mu-Tao Wang, and Ye-Kai Wang.

• (Wang-Yau, 2009) To evaluate the quasilocal mass of a 2-surface Σ with the physical data (σ, H), one solves the optimal isometric embedding equation, which gives an embedding of Σ into the Minkowski spacetime with the image surface Σ_0 that has the same induced metric as Σ, i.e. σ. One then compares the extrinsic geometries of Σ and Σ_0 and evaluate the quasilocal mass from σ, H and H_0.

• The physical surface Σ with physical data $(\sigma, \mathbf{H})$ gives $(\sigma, |\mathbf{H}|, \alpha_{\mathbf{H}})$.

• Given an isometric embedding $X : \Sigma \to \mathbb{R}^{3,1}$ of σ. Let Σ_0 be the image $X(\Sigma)$ and $(\sigma, |\mathbf{H}_0|, \alpha_{\mathbf{H}_0})$ be the data of Σ_0.

• Let T be a future timelike unit Killing field of $\mathbb{R}^{3,1}$ and define $\tau = -\langle X, T\rangle$.

• Define a function ρ and a 1-form j_a on Σ:

$$\rho = \frac{\sqrt{|\mathbf{H}_0|^2 + \frac{(\Delta\tau)^2}{1+|\nabla\tau|^2}} - \sqrt{|\mathbf{H}|^2 + \frac{(\Delta\tau)^2}{1+|\nabla\tau|^2}}}{\sqrt{1+|\nabla\tau|^2}}$$

$$j_a = \rho\nabla_a\tau - \nabla_a\left(\sinh^{-1}\left(\frac{\rho\Delta\tau}{|\mathbf{H}_0||\mathbf{H}|}\right)\right) - (\alpha_{\mathbf{H}_0})_a + (\alpha_{\mathbf{H}})_a.$$

• The optimal isometric embedding equation is

$$\begin{cases} \langle dX, dX\rangle = \sigma \\ \nabla^a j_a = 0 \end{cases}$$

• For a solution (X, T), the quasi-local mass is then

$$E(\Sigma, X, T) = \frac{1}{8\pi}\int_\Sigma \rho.$$

• Σ_0 is the “unique” surface in the Minkowski spacetime that best matches the physical surface Σ. If the original surface Σ happens to be a surface in the Minkowski spacetime, the above procedure identifies $\Sigma_0 = \Sigma$ up to a global isometry.

• $E(\Sigma, X, T)$ is positive in general, and zero for surfaces in the Minkowski spacetime.

• In addition, in joint work with Po-Ning Chen, we also defined quasilocal conserved quantities.

• For an optimal isometric embedding (X, T), by restricting a rotation (or boost) Killing field K of $\mathbb{R}^{3,1}$ to $\Sigma_0 = X(\Sigma) \subset \mathbb{R}^{3,1}$, the quasi-local conserved quantity is defined to be:

$$-\frac{1}{8\pi}\int_\Sigma \langle K, T\rangle\rho + (K^\top)^a j_a,$$

where $K^\top$ is the component of K that is tangential to Σ_0.

• In general, the optimal isometric embedding equation is difficult to solve. However, in a perturbative configuration, when a family of surfaces limit to a surface in the Minkowski spacetime, the optimal isometric embedding equation is solvable, subject to the positivity of the limiting ρ.

• An important application of the theory is to anchor the definition of total mass and total angular momentum of an asymptotically flat initial data set.

• If (M, g, k) is an asymptotically flat initial data set with $g - \delta = O(r^{-p})$ and $k = O(r^{-q})$ as $r \to \infty$, where $p > \frac{1}{2}$ and $q > \frac{3}{2}$, the ADM mass

$$\frac{1}{16\pi}\int_{S^2_\infty}\sum_{i,j}(g_{ij,j} - g_{jj,i})\nu^i.$$

• The ADM angular momentum is defined as

$$\frac{1}{8\pi}\int_{S^2_\infty}\sum_{i,j}(k_{ij} - tr_g k g_{ij})K^i\nu^j$$

where K^i is an asymptotic rotation Killing field on (M, g, k).

• Note that, however, the calculation of angular momentum is more subtle, as the expression diverges apparently.

• There are proposals (Regge-Teitelboim) of parity condition on (g, k) to ensure finiteness, and important gluing constructions and density theorems for prescribing angular momentum by Corvino-Schoen, Chruściel-Delay, Chruściel-Corvino-Isenberg, Huang, Huang-Schoen-Wang etc. under such a condition.

• (Chen-Huang-Wang-Yau) There exist asymptotically flat spacelike hypersurfaces in $\mathbb{R}^{3,1}$ or the

Schwarzschild spacetime with finite, nonzero ADM angular momentum such that $g-\delta = O(r^{-\frac{4}{3}})$ and $k = O(r^{-\frac{5}{3}})$.

• (Chruściel) If $p+q>3$, then the ADM angular momentum is finite.

• To what extent is the ADM definition a valid one?

• On the other hand, the limit of CWY quasilocal angular momentum can be viewed as a total angular momentum on an asymptotically flat initial data set.

• To justify our definition, we prove an invariance of CWY angular momentum theorem in Kerr: any strictly spacelike hypersurface in the Kerr spacetime has the same total CWY angular momentum.

• "Strictly spacelike" means, in Boyer-Lindquist coordinates, $t = O(cr)$ for $|c|<1$. In particular, the CWY angular momentum vanishes for such hypersurfaces in $\mathbb{R}^{3,1}$ or the Schwarzschild spacetime.

• The proof relies on a gravitational conservation law.

• We also computed the limit of quasilocal conserved quantities for spacelike hypersurface of harmonic asymptotics (Corvino-Schoen) and asymptotical hyperboloids.

• I shall now discuss the definition of total angular momentum at future null infinity $\mathfrak{I}^+$.

• There were various definitions and proposals: Ashtekar-Hansen, Dray-Streubel, Rizzi (Christodoulou-Klainerman), Chruściel-Jezierski-Kijowski, etc.

• How do we justify these definitions?

• The spacetime near $\mathfrak{I}^+$ is described in terms of the Bondi-Sachs coordinate system.

• $(u,r,x^a, a=2,3)$ is a Bondi-Sachs coordinate system if near $r=\infty$, the spacetime metric takes the form

$$-UVdu^2 - 2Ududr + \sum_{a,b=2,3} r^2 h_{a,b}(dx^a + W^a du)(dx^b + W^b du) = g_{\alpha\beta}dx^\alpha dx^\beta,$$

such that

$$\det h_{ab} = \det \sigma_{a,b},$$

where $\sigma_{a,b}$ is the standard round metric on (S^2, x^a). U, V, h_{ab}, W^a depend on $u, r, x^a, a=2,3$.

• Each $u = constant$ is a null hypersurface and r is indeed the inverse mean curvature flow (from $r=\infty$) parameter on $u = constant$.

• The asymptotically flat condition and the vacuum Einstein equation imply that as $r\to\infty$,

$$\begin{aligned} U &= 1 + O(r^{-2}) \\ V &= 1 - \frac{2m}{r} + O(r^{-2}) \\ W^a &= O(r^{-2}) \\ h_{ab} &= \sigma_{ab} + \frac{C_{ab}}{r} + O(r^{-2}) \end{aligned}$$

where $m = m(u,x^a)$ is the mass aspect and $C_{ab} = C_{ab}(u,x^a)$ is the shear tensor which is symmetric and traceless with respect to σ.

• As $r\to\infty$, $r^{-2}g_{\alpha\beta}dx^\alpha dx^\beta \to \sigma_{ab}dx^a dx^b$, the null metric on $\mathfrak{I}^+$.

$$\mathfrak{I}^+ \sim I \times (S^2, \sigma_{ab})$$

with $u\in I$, $x^a \in S^2$.

• The Bondi-Sachs-Trautman energy-momentum is

$$e(u) = \frac{1}{4\pi}\int_{S^2_\infty} m(u,x^a)dv_\sigma,$$
$$p_i(u) = \frac{1}{4\pi}\int_{S^2_\infty} m(u,x^a)Y_i dv_\sigma, \ i=1,2,3$$

where $\{Y_i = Y_i(x^a), i=1,2,3\}$ is an orthonormal basis of the (-2) eigenspace of $\Delta = \Delta_\sigma$.

• Each Bondi-Sachs coordinate system (u,r,x^a) induces a limiting coordinate system (u,x^a) on $\mathfrak{I}^+$ together with the mass aspect $m(u,x^a)$ and the shear $C_{ab}(u,x^a)$.

• Such a Bondi-Sachs coordinate system is not unique and the BMS group, which consists of diffeomorphisms that preserve the gauge and boundary conditions, acts on the space of Bondi-Sachs coordinate systems.

• A BMS group element induces a diffeomorphism on $\mathfrak{I}^+$ that is of the following form:

$$(u,x^a)\mapsto(\bar{u},\bar{x}^A)$$

such that

$$\begin{cases} \bar{x}^A = g^A(x^a) \\ \bar{u} = K(x^a)(u+f(x^a)) \end{cases}$$

where $g:(S^2,\sigma)\to(S^2,\bar{\sigma})$ is a conformal isometry, i.e. $g^*\bar{\sigma} = K^2\sigma$ and $K = \frac{1}{(\alpha^0 + \sum_i \alpha^i Y_i)}$ with (α^0,α^i) a future timelike unit vector.

• $f(x^a)$ is any smooth function on S^2 that is called a "supertranslation". $f(x^a) = \sum a_i Y_i$ corresponds to an actual translation in the Poincaré group.

• $K = \frac{1}{(\alpha^0 + \sum_i \alpha^i Y_i)}$ corresponds to a boost in $O(3,1)$.

• Choices of $Y_i, i=1,2,3$ correspond to $O(3)\subset O(3,1)$.

• Invariance and monotonicity of mass are best described in terms of a modified mass aspect:

$$\hat{m} = m - \frac{1}{4}\nabla^a\nabla^b C_{ab}.$$

• Under a BMS transformation (K,f), $(u,x^a,m,\hat{m}) \mapsto (\bar{u},\bar{x}^A,\bar{m},\hat{\bar{m}})$, the two modified mass aspects $\hat{m}$ and $\hat{\bar{m}}$ are related by

$$\hat{m} - \frac{1}{4}\Delta(\Delta+2)f = K^3\hat{\bar{m}}.$$

• This implies $(e(u),p_i(u))$ and $(e(\bar{u}),p_i(\bar{u}))$ differ by exactly the boost associated with K.

• In addition, the vacuum Einstein equation implies

$$\partial_u \hat{m} = -|\partial_u C|^2_\sigma,$$

and the modified mass aspect is pointwise non-increasing (mass loss formula). Note that m and $\hat{m}$ define the same energy-momentum. $\partial_u C_{ab} := N_{ab}$ is called the news tensor.

• There is also an angular momentum aspect from the expansion of W^a.

$$r^2 h_{ab} W^b = g_{ua} = \frac{1}{2}\nabla^b C_{ab} + \frac{g_{ua}^{(-1)}}{r} + O(r^{-2}).$$

• The angular momentum aspect is defined to be

$$N_a = -\frac{1}{2}\left(g_{ua}^{(-1)} + \frac{1}{16}\nabla_a |C|^2\right)$$

• The vacuum Einstein equation implies

$$3\frac{\partial N_a}{\partial u} = -\nabla_a m + \frac{1}{4}\epsilon_a^b \nabla_b(\epsilon^{ec}\nabla_c \nabla^d C_{de}) - \frac{3}{4} C_{ab}\nabla_c N^{bc} + \frac{1}{4} N^{cd}\nabla_d C_{ac}.$$

• Does there exist a modified angular momentum aspect that satisfies similar properties of the modified mass aspect?

• Unfortunately, the transformation of the angular momentum aspect is extremely complicated.

• Under a supertranslation f, N_a transforms as

$$N_a = \bar{N}_a + m\nabla_a f + \frac{1}{2}(\nabla^c f)(\nabla_c \nabla_b C^b_a - \nabla_a \nabla_b C^b_c)$$
$$+ 15 \text{ other terms that involve } f, C_{ab}, N_{ab} \text{ and their derivatives}$$

• CJK, from the Hamiltonian theory associated with Bondi-Sachs coordinates, defined the total Lorentz charge to be the integral

$$\int_{S^2} (24 N_a + 2\nabla_c (C_{ab} C^{bc}) + \frac{1}{2}\nabla_a |C|^2)(\cdot),$$

$(\cdot) = \epsilon^{ab}\nabla_b Y_i$ corresponds to angular momentum and $(\cdot) = \nabla^a Y_i$ corresponds to center of mass.

• Theorem (CJK) The Lorentz charge is equivariant under the BMS group if there exists a Bondi-Sachs coordinate system such that $m = constant$, $N_a = constant$, and $C_{ab} = 0$.

• The assumptions correspond to a stationary spacetime.

• The limit of the CWY quasilocal conserved quantities was recently computed by Keller-Wang-Yau. The expression depends on the Hodge decomposition of C_{ab}. Write

$$C_{ab} = \nabla_a \nabla_b \mathbf{c} - \frac{1}{2}\sigma_{ab}\Delta \mathbf{c} + \frac{1}{2}(\epsilon_{ad}\nabla^d \nabla_b \underline{\mathbf{c}} + \epsilon_{bd}\nabla^d \nabla_a \underline{\mathbf{c}})$$
$$= F_{ab} + \underline{F}_{ab}$$

• The limit of the CWY quasilocal center of mass is

$$\int_{S^2} (\nabla^a Y_i)\left(\frac{3}{2}N_a - u\nabla_a m - c\nabla_a m + 2\epsilon_a^b(\nabla_b \underline{\mathbf{c}})m\right) + \int_{S^2} Y_i \left(-\frac{1}{16}|\nabla(\Delta+2)\underline{\mathbf{c}}|^2 - \frac{1}{2}\nabla_d F^{ad}\nabla^b \underline{F}_{ab} - \frac{1}{4}F^{ab}\underline{F}_{ab}\right)$$

• The expression coincide with CJK for a stationary spacetime and thus is invariant under the BMS group in this case as well.

• In the more general case, the expression involves more refined structure of $\mathfrak{I}^+$ (Hodge decomposition of the shear C_{ab}).

• All previous definitions of angular momentum on $\mathfrak{I}^+$ depend on a specific gauge (a null frame or a spacetime coordinate system). In contrast, our definition is geometric and coordinate independent (depends only on σ, $|H|$, α_H).

• In addition, solving the optimal isometric equation is a canonical procedure that is free from any ad hoc referencing. As we are just calculating the same conserved quantities of a surface and specialize to different Bondi-Sachs coordinate, general invariant/equivariant properties are expected.

Einstein's theory of gravitation has initiated a deep understanding of geometry through physical insight and vice versa, in the last century. We expect this to continue in this century.