The Penrose Inequality : Differential Geometry and Black Holes

One of the richest and most useful areas of mathematics is the interplay between the calculus of variations, geometry, and mathematical physics. This interplay is keenly illustrated in Einstein's theory of general relativity, or "curved space".

Today, it is well established that strong gravitational fields don't just move things around, but warp the very shape of space itself. This manifests itself in many phenomena now being observed in the heavens: the bending of light, or "gravitational lensing"; the slowing of time as one climbs out of a gravitational well; discrepancies in the global positioning system; precession of orbits; the gradual spindown of orbiting pulsars.

One of the most striking phenomena, for which the evidence is rapidly mounting, is a *black hole*: a region, perhaps around a collapsed star or galaxy core, where the force of gravity is so strong that light itself cannot escape.

To make physical predictions for comparison against evidence from the skies, we must understand more fully the geometric consequences of Einstein's theory. The Riemannian Penrose Inequality, recently proven by Huisken and Ilmanen (of the MPI), is a mathematical theorem that limits the possible shapes of black holes. Roughly speaking, it says that the surface area of a black hole is limited by its total mass, thus clearly establishing a relationship long conjectured by physicists. In order to explain it, let us first take a look at the historical background.

Background

Underlying the physics lies *differential geometry*: the mathematical theory of curved space. What shapes of space are possible? Let us take a look at the evolution of this subject from mathematical speculation to concrete description of our world.

In 1854, Riemann proposed a bold program: to replace the simple, linear axioms of Euclidean geometry by a scheme in which the local nature of space – the local concept of distance – is allowed to vary freely from point to point.

A simple example is given by a curved surface in space - for example, the surface of a vase. (See Figure 1.) A tiny bug that crawls on the vase will experience a 2-dimensional geometry that seems on a small scale very much like a flat, Euclidean plane.

More detailed measurements, however, reveal tiny discrepanies! For example, the sum of the angles of a small triangle will differ slightly from 180°. At this point, the bug will know that Euclid's flat geometry is not quite right. The deviation of the surface near a given point x from a Euclidean plane is measured by a quantity K, called the *(Gaussian) curvature* of the surface. If K is positive, the angle sum is greater than 180°, a so-called "angle excess". Similarly, if K is negative, the angle sum is less than 180°

Globally, Riemann's concept gives rise to the concept of a *manifold*: an abstract n-dimensional "curved space" that is complete in itself, requiring no larger space to contain it.

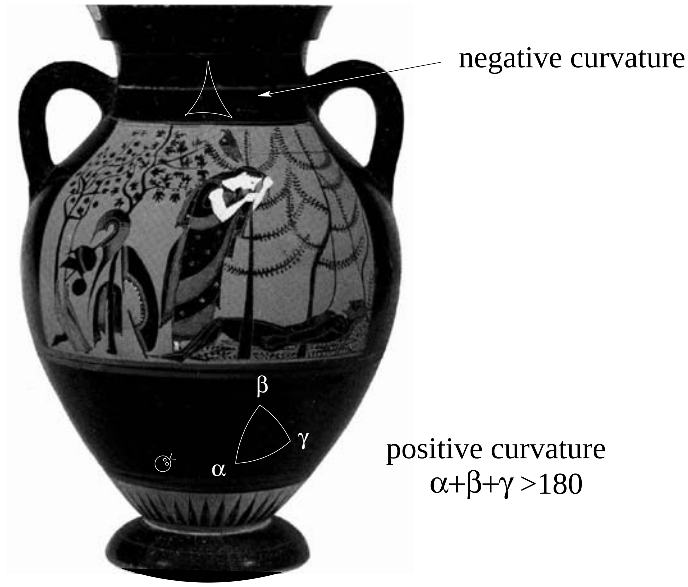

Figure 1

A manifold need not represent our physical, 3-dimensional world. Indeed, manifolds are widely employed today to represent all the possible states of any system, such as a spinning top, robot arm, chemical process, or market model, whose variables (angle, momentum, concentration, price) are constrained by some nonlinear relationship.

The entities of differential geometry took on a direct, physical reality in 1915 with the appearance of Einstein's theory of general relativity. He showed that the gravitational effects of mass-energy express themselves by affecting the geometric structure of space itself. Locally, matter causes curvature, which in turn determines the shape of space. Gravitational attraction is represented by positive curvature, which tends to make free-falling particles seem to accelerate toward one another. In other words, we ourselves live in a curved space as envisioned by the mathematicians.

Like the bug, we must infer the properties of the overall space from limited, local measurements. Indeed, one of the central, guiding questions in the study of manifolds is the following.

> The Problem of Integration: How does local geometry (that is, curvature) affect global geometry (for example, general shape, overall size, or topology)?

For example, a small sphere has very large, positive curvature, whereas a large sphere has small curvature. In general, a surface with positive curvature will tend to curve around, reconverge, and ultimately close up, forming a finite, bounded surface. (See Figure 2.) This tendency can be given a quantitative form, in which the diameter, d, of the surface is bounded by

$$d \leq \pi/\sqrt{c}$$

where c is the minimum value of the curvature. Thus: a surface with high curvature will be very small. Such a relationship, known as a *geometric inequality*, gives us powerful information about the general shape of the surface. Discovering such geometric inequalities is a major task of differential geometry.

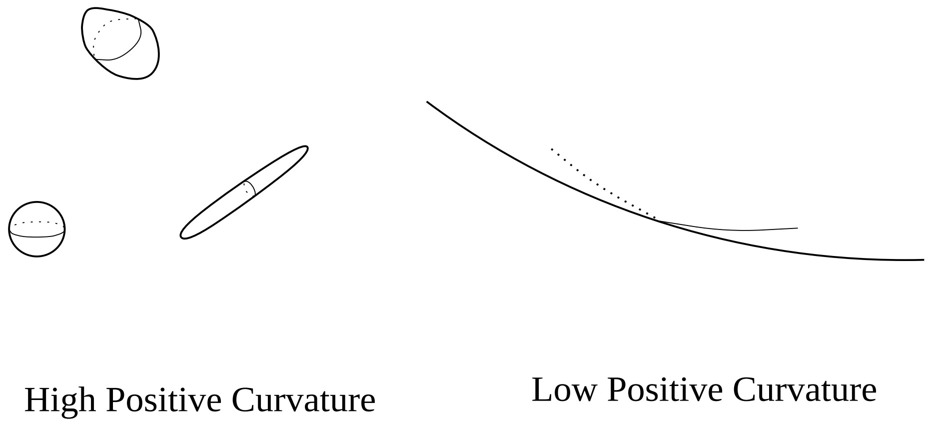

Figure 2

Here is another example of a geometric inequality. A soap bubble surrounding a blob of air will wobble around for awhile, but eventually settle down to the familiar spherical shape. Now, any surface enclosing a volume V must have have a certain minimum area A, according to the inequality

$$A \geq \sqrt[3]{36\pi V^2},$$

known as the *Isoperimetric Inequality.* A round sphere achieves exact equality. Thus, the sphere is a "perfect", highly symmetrical shape that minimizes surface area among all possible surfaces that enclose the given volume V.

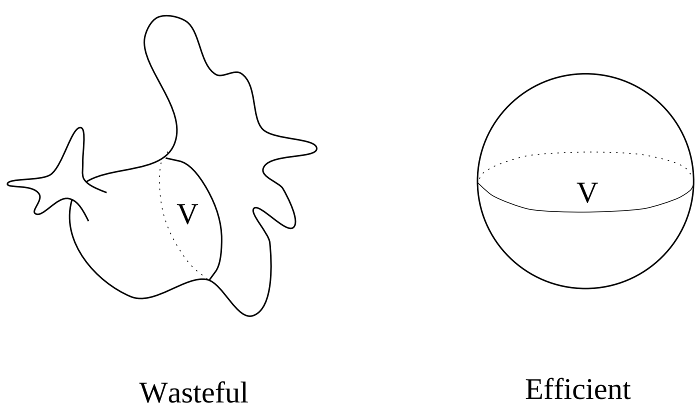

Figure 3

Something analogous happens during the gravitational collapse of a massive star or star cluster. After wobbling about for awhile, the space around the crushing matter will, under suitable conditions, settle down to a "perfect" solution of Einstein's equations, found in 1916 by the astronomer Schwarzschild: a static, highly symmetrical space-time modelling the gravitational field around a nonrotating, isolated body.

Unlike the original star, however, the massive center of the Schwarzschild solution is hidden away where it cannot be seen, because the gravity has become so strong that even light cannot escape. The boundary of this "region of no escape", or black hole, is a surface called the *event horizon*. (See Figure 4.) Somewhat later it was noticed that beyond the event horizon is a location where space-time itself becomes sharply curved and ultimately singular. Here, the usual laws of geometry and physics break down, and infalling matter more or less ceases to exist.

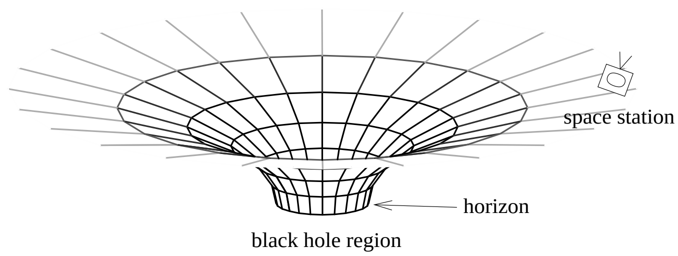

Figure 4

In the 1960's the study of general relativity took on a new life with the improvement of astrophysical observations and the systematic application of methods of differential geometry. Using the positive curvature of space-time, Hawking and Penrose demonstrated that in Einstein's theory, singularities are not rare, but typical. The presence of a so-called *apparent horizon* – a kind of area-minimizing surface present at the starting time – serves as a harbinger to predict the later development of a singularity.

According to the somewhat controversial *Cosmic Censorship Hypothesis*, a black hole then forms "around" the singularity, shielding it from view, and preventing it from affecting an external observer – not just for the Schwarzschild solution, but generally. Many physicists consider such a principle essential in order to explain the stability of space-time near a singularity.

The Penrose Inequality

From the beginning, relativity has been beset by the difficulty of choosing coordinates. Measurements depend on the yardstick – the local reference frame of the observer – and only by considering systematically all possible yardsticks can the full picture be obtained. This is the true meaning of the term "relativity".

A principal example of this difficulty is the problem of mass. The local mass density of space-time is represented by the positive curvature at each point. The total mass, m, of an isolated gravitating system, on the other hand, is defined by the gravitational effect of the system at large distances – for example, as measured from a space station far from the highly curved region of space. In this context, the Problem of Integration becomes the following.

> THE PROBLEM OF MASS: How does the local mass density determine the total mass m?

In classical Newtonian mechanics, this problem has an easy solution: you just add up, or integrate, the local mass density ρ to obtain the total mass according to the formula

$$m = \int_{\text{space}} \rho.$$

In Einstein's theory, the problem is far more difficult, because there is no preferred coordinate system in which to add up the little pieces of mass.

A deceptively simple test case was resolved in 1979 by the *Positive Mass Theorem* of Schoen and Yau: if the local curvature is positive, then m is positive. (So the isolated system exerts attraction rather than repulsion.) Though the statement is straightforward, the mathematical techniques developed are very deep, and have led to many developments in conformal geometry and in the topology of positive scalar curvature. Since then, several other proofs have appeared, based on rather different physical and mathematical ideas. The diversity of proofs mysteriously echoes the multiplicity of coordinate systems.

In 1973 Penrose had put forward a more precise version of the Positive Mass Theorem, namely

$$m \geq \frac{c^2}{G}\sqrt{\frac{A}{16\pi}},$$

where A is the surface area of a minimal surface, or apparent horizon, present in space, G is Newton's gravitational constant, and c is the speed of light. Roughly speaking, this says: the area of a black hole is limited by its total mass. (For example, the area of a black hole with the mass of an apple would be at most 3×10^{-51} square centimeters.)

This inequality gives a universal geometric relationship analogous to the Isoperimetric Inequality. The optimal shape is given by the highly symmetric Schwarzschild manifold, which realizes the smallest mass consistent with the given area A.

Penrose did not give a mathematical proof of this inequality, however. Instead, he gave an argument based on physical reasoning starting from the Cosmic Censorship Hypothesis. He envisaged the inequality as a mathematical test of Cosmic Censorship: if an example were found that violates the inequality, it would most likely indicate failure of Cosmic Censorship, and the existence of a so-called "naked singularity" – which would be quite a shock.

Just in the past year, singnificant progress has been made on this inequality by Gibbons, Bray, and Herzlich. Recently, Huisken and Ilmanen have been able to prove the Penrose Inequality for a 3-dimensional Riemannian manifold of positive scalar curvature, which arises in the special case of a time-reversible space-time. This rules out one avenue to naked singularities.

The proof draws together techniques from differential geometry, the calculus of variations, and mathematical relativity. The principal tool is an evolution equation related to minimal surfaces, discovered by the physicist Geroch in 1973. According to this scheme, a surface N_t starts at the apparent horizon, and evolves progressively outward with speed equal to the *inverse* of its average curvature. During this evolution, a masslike quantity associated to the surface is monotone increasing. The initial value of this quantity is the surface area A, and the final value, as the surface becomes a large, round sphere enclosing the system, may be compared to the mass m. Certain analytic difficulties arise, because the surface N_t itself may become singular and jump around in the manifold. By resolving these difficulties, we establish the Penrose Inequality.

Determining the geometric consequences of Einstein's equations – in effect, solving the geometric Integration Problem – is an essential step in making physical predictions that may be compared against observational evidence in searching the heavens for black holes and other effects of curved space. Geometric inequalities and curvature flows have an important role to play in the mathematical underpinnings of this project. Such techniques, among many others, illustrate how methods of modern mathematics can serve to verify and illuminate heuristic principles coming from physics.

T. Ilmanen, Leipzig, April 1998