Theory of Relativity
Michaelmas Term 2009: M. Haehnelt

13 Observational Evidence for Black Holes

13.1 Collapse of a Massive Star to form a Black Hole

Do black holes exist in Nature? A star like the Sun is held up by radiation pressure. It is burning hydrogen to make helium releasing about 26 Mev for each atom of He that is formed. What happens when all of the nuclear fuel is used up? The star must collapse to a high density. In fact, we expect that the Sun will collapse to a form a white dwarf – a star with a radius of about 5000 km and a spectacularly high mean density of about 10^6 g cm^{-3}.

Astronomers knew about white dwarfs a long time ago. For example, in 1915 Adams discovered that the companion of the bright star Sirius had a 'white' spectrum from which he deduced an effective temperature of 8000 K. The mass of the companion (known as Sirius B) was estimated to be about $0.8M_\odot$ from observations of the binary orbit. Given its observed flux and temperature, the blackbody radiation law could be used to infer a radius of $\sim 20,000$ km and so, given its mass, there could be no escape from the conclusion that Sirius B was a very compact star of extraordinarily high density. In 1925, Adams showed further that the spectral lines of Sirius B were gravitationally redshifted – as expected from General Relativity. Here is what Eddington said about this discovery in his classic book *The Internal Constitution of the Stars* written in 1926:

> "Prof. Adams has killed two birds with one stone; he has carried out a new test of Einstein's general theory of relativity and he has confirmed our suspicion that matter 2000 times denser than platinum is not only possible, but is present in the Universe."

This, incidentally, is a good example of why astronomy is so interesting. Astronomy turns up strange objects in which conditions are much more extreme than anything we encounter here on Earth. This allows us to study physics in unusual regimes, where we can hope to learn something fundamentally new . Examples include neutron stars, quasars, relativistic jets, accretion discs, gamma ray bursts and cosmological fluctuations in the early Universe.

Nobody knew how to explain white dwarfs. What holds them up? What provides the pressure? The answer had to await the development of quantum mechanics and the formulation of Fermi-Dirac statistics. Fowler realised in 1926 that white dwarfs were held up by electron degeneracy pressure. The electrons in a white dwarf behave like the free electrons in a metal. Because of the Pauli exclusion principle, the electrons completely fill phase space up to a characteristic Fermi-energy. It is the Pauli exclusion principle that holds up a white dwarf. Chandrasekhar, in 1930, was on his way to Cambridge to start work as a research student under Eddington's supervision. In transit, Chandrasekhar realized that the more massive a white dwarf, the denser it must be and so the stronger the gravitational field. For white dwarfs about a critical mass of $1.4M_\odot$ (now called the Chandrasekhar limit), gravity would overwhelm degeneracy pressure and no stable solution was possible. The white dwarf would collapse to a point. Eddington was not impressed (though Chandrasekhar later won the Nobel prize for this work). Here is another quote from Eddington,

> "The star apparently has to go on radiating and radiating and contracting and contracting until, I suppose, it gets down to a few kilometres radius when gravity becomes strong enough to hold the radiation and the star at last can find peace.... I think that there should be a law of Nature to prevent the star from behaving in this absurd way."

After the discovery of the neutron, people realized that at extremely high densities the electrons would react with the protons to form neutrons. A new stable configuration for a star was possible – neutron stars. A neutron star of one solar mass would have a radius of only 10 km. But as with white dwarfs, neutron stars have a maximum mass above which no stable configuration is possible. This maximum mass is believed to be about $3M_\odot$ or less – the exact value is uncertain because of uncertainties in the equation of state of matter at such high densities. Nevertheless, we now believe that Eddington was wrong – there is no law of Nature to prevent the gravitational collapse of a massive star. We believe that massive stars collapse to form *black holes*.

Figure 1 shows a space time diagram of a collapsing star. Eventually, the radius of the star collapses to less than the Schwarzschild radius, at which point the formation of a black hole becomes inevitable.

Some theorists were very skeptical about the formation of black holes. The Schwarzschild solution is very special – it is exactly spherically symmetric by construction. In reality, a star will not be perfectly symmetric and so perhaps, as it collapses, the assymmetries amplify and avoid the formation of an event horizon. Perhaps black holes never form or are exceedingly rare in Nature. Penrose in the early 1960's applied global geometrical techniques to prove a famous series of 'singularity theorems'. These show that in realistic situations an event horizon (a closed trapped surface) will be formed and that there must exist a singularity within this surface, *i.e.* a point at which the curvature diverges and General Relativity ceases to be valid. I will touch on Penrose's approach in a later lecture. The singularity theorems were important in convincing people that black holes must form in Nature. In the rest of this section, I will look at some of the observational evidence for the existence of black holes. As we will see, there is compelling evidence that black holes do indeed exist. Furthermore, it should become possible within the next few years not only to measure the

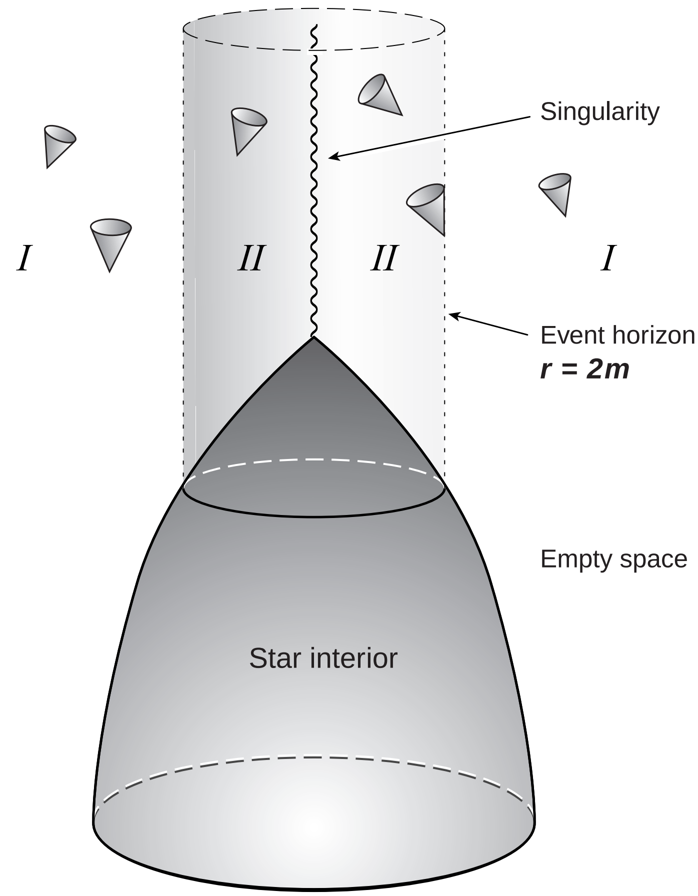

Figure 1: Collapse of a star to form a black hole.

masses of black holes, but also to measure their angular momenta using powerful X-ray telescopes! Direct experimental probes of the strong gravity regime are now possible.

13.2 Compact binary systems

Although radiation cannot escape from black holes, one of the best way of finding candidate black holes is to search for luminous compact X-ray sources. The reason for this is that if a black hole has a stellar companion, the intense tidal field can pull gas from the companion producing an accretion disc around the black hole. A schematic picture is shown below. I will show you that accretion discs can radiate very efficiently.

The following table summarizes the common classes of compact binaries. The compact object can be a white dwarf, neutron star, or a black hole.

Figure 2: Schematic picture of a compact binary system.

Compact accreting binary systems			
	Compact object		
Companion Star	White Dwarf	Neutron Star	Black Hole
Early type, massive	None known	Massive X-ray binaries	Cyg X-I, LMC X-3
Late-type, low mass	Cataclysmic variables (e.g., Dwarf novae)	Low mass X-ray binaries	A0620-00

If you find a compact binary system, then you can set limits on the mass of the compact object from the dynamics of the binary orbit. If you find evidence for a compact object that is more massive than the Chandrasekhar limit, then you have good evidence that the object might be a black hole. In fact it is not so straightforward. What observers actually measure is the *mass function*

$$f(M) = \frac{PK^3}{2\pi G},$$

where P is the orbital period, and K is the radial velocity amplitude. For example, for the low mass X-ray binary A0620-00, the period is $P = 7.7$ hours and $K = 457$ km s^{-1}. From Kepler's laws we can show that the mass function is related to the masses of the compact object (M_1), the companion star (M_2) and the inclination angle i of the orbit to the plane of the sky by

$$f(M) = \frac{M_1^3 \sin^3 i}{(M_1 + M_2)^2}.$$

You can see from this equation that the mass function is a strict lower limit on the mass of the compact object. It is equal to the mass, $f = M_1$, *only* if $M_2 = 0$ and the orbit is viewed edge on ($\sin i = 1$). For example, for A0620-00, the lower limit on the mass of the compact object is $2.9M_\odot$, and this makes it a very good black hole candidate because this mass limit is very close to the theoretical *upper* limit for the mass of a neutron star. In fact, it is possible to make reasonable estimates[1] for M_2 and $\sin i$ in this system leading to a probable mass of $\approx 10M_\odot$ for the compact object – well into the black hole regime.

The following table from a review by Phil Charles summarizes the dynamical mass limits on some good black hole candidates (so called short X-ray transients). As you can see in several systems, V404 Cyg, G2000+25 and N Oph 77, the minimum mass inferred from the mass function is *well above* the theoretical maximum mass limit for a neutron star. As we understand things at present there can be no other explanation other than that the compact objects are black holes.

[1] An estimate of the mass M_2 can be made by measuring the spectral type and luminosity of the companion star. The inclination angle can be estimated from the shape of the stars light curve – searching for evidence of eclipsing by the compact object.

Table 2. Derived Parameters and Dynamical Mass Measurements of SXTs

Source	$f(M)$ ($M_\odot$)	ρ (g cm^{-3})	q ($= M_1/M_2$)	i	M_1 ($M_\odot$)	M_2 ($M_\odot$)	Ref.
V404 Cyg	6.08 ± 0.06	0.005	17 ± 1	55 ± 4	12 ± 2	0.6	[1-2]
G2000+25	5.01 ± 0.12	1.6	24 ± 10	56 ± 15	10 ± 4	0.5	[3-5]
N Oph 77	4.86 ± 0.13	0.7	> 19	60 ± 10	6 ± 2	0.3	[6-9]
N Mus 91	3.01 ± 0.15	1.0	8 ± 2	54^{+20}_{-15}	6^{+5}_{-2}	0.8	[13-15]
A0620-00	2.91 ± 0.08	1.8	15 ± 1	37 ± 5	10 ± 5	0.6	[16-18]
J0422+32	1.21 ± 0.06	4.2	> 12	$20 - 40$	10 ± 5	0.3	[19-20]
J1655-40	3.24 ± 0.14	0.03	3.6 ± 0.9	67 ± 3	6.9 ± 1	2.1	[10-12]
4U1543-47	0.22 ± 0.02	0.2	–	$20 - 40$	5.0 ± 2.5	2.5	[21]
Cen X-4	0.21 ± 0.08	0.5	5 ± 1	43 ± 11	1.3 ± 0.6	0.4	[22-23]

References

[1] Casares & Charles 1994; [2] Shahbaz et. al. 1994b; [3] Filippenko et. al. 1995a; [4] Beekman et. al. 1996; [5] Harlaftis et. al. 1996; [6] Filippenko et. al. 1997; [7] Remillard et. al. 1996; [8] Martin et. al. 1995; [9] Harlaftis et. al. 1997; [10] Orosz & Bailyn 1997; [11, 12] van der Hooft 1997, 1998; [13] Orosz et. al. 1996; [14] Casares et. al. 1997; [15] Shahbaz et. al. 1997; [16] Orosz et. al. 1994; [17] Marsh et al 1994; [18] Shahbaz et. al. 1994a; [19] Filippenko et. al. 1995b; [20] Beekman et. al. 1997; [21] Orosz et. al. 1998; [22] McClintock & Remillard 1990; [23] Shahbaz et. al. 1993.

13.3 Stable and Unstable orbits around a Schwarzschild black hole

In Newtonian dynamics the equation of motion of a particle in a central potential is

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 + V(r) = E,$$

where $V(r)$ is an "effective potential". For an orbit around a point mass, the effective potential is

$$V(r) = \frac{h^2}{2r^2} - \frac{GM}{r},$$

where h is the specific angular momentum of the particle. The concept of an effective potential is, I am sure, familiar to you. If you sketch the effective potential, you can see easily that bound orbits have two turning points and that a circular orbit corresponds to the special case where the particle sits at the minimum of the effective potential.

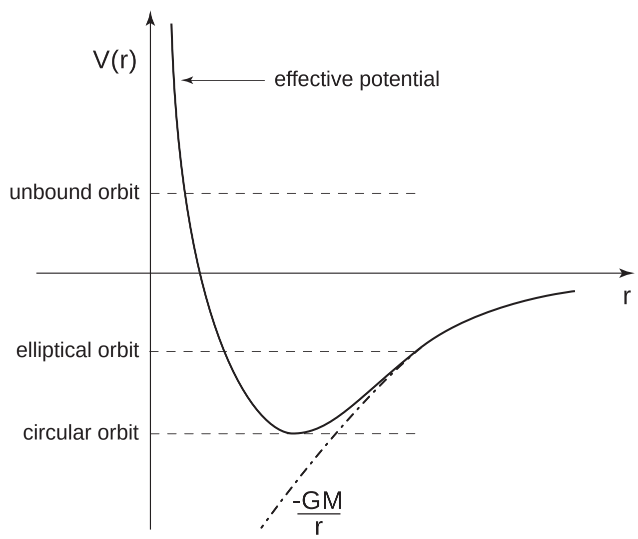

Figure 3: The Newtonian effective potential showing how an angular momentum barrier prevents particles reaching $r = 0$.

Furthermore, you can see that in Newtonian dynamics, a finite angular momentum provides an *angular momentum barrier* preventing a particle reaching $r = 0$. This is not true in General Relativity.

In GR, the equation of motion of a particle around a central mass point is

$$\beta(r)\left(\frac{dr}{dp}\right)^2 + \frac{J^2}{r^2} - \frac{c^4}{\alpha(r)} = -\mathcal{E}, \tag{1}$$

where

$$\alpha(r) = \frac{c^2}{\beta(r)} = \left(1 - \frac{2\mu}{r}\right) c^2.$$

Since $c^2 d\tau^2 = \mathcal{E} dp^2$ we can rewrite equation (1) as

$$\left(\frac{dr}{d\tau}\right)^2 + V(r) = \frac{c^4}{\mathcal{E}} \tag{2}$$

where $h = Jc/\sqrt{\mathcal{E}}$ and $V(r)$ is an effective potential

$$V(r) = \left(1 - \frac{2\mu}{r}\right)\left(1 + \frac{h^2}{r^2c^2}\right) c^2. \tag{3}$$

If we transform to the dimensionless variables

$$\begin{aligned} x &= \frac{2r}{r_s} = \frac{rc^2}{GM}, \\ h'^2 &= \frac{4h^2}{r_s^2 c^2}, \end{aligned}$$

we can write the effective potential as

$$V(x) = c^2\left(1 - \frac{2}{x}\right)\left(1 + \frac{h'^2}{x^2}\right) = \left(1 - \frac{2}{x} - \frac{2h'^2}{x^3} + \frac{h'^2}{x^2}\right) c^2. \tag{4}$$

Differentiating this expression,

$$\frac{dV}{dx} = \left(\frac{2}{x^2} + \frac{6h'^2}{x^4} - \frac{2h'^2}{x^3}\right) c^2,$$

and so the extrema of the effective potential are located at the solutions of the quadratic equation

$$x^2 - h'^2 x + 3h'^2 = 0,$$

i.e. at

$$x = \frac{h'}{2}\left\{h' \pm \sqrt{h'^2 - 12}\right\}.$$

IF $h' = \sqrt{12} = 2\sqrt{3}$ *there is only one extremum*, and there are no turning points in the orbit for lower values of h'. Figure 4 shows the effective potential for several values of h'. The dots show the locations of stable circular orbits. The maxima in the potential are the locations of *unstable* circular orbits.

What is the physical significance of this result? The smallest stable circular orbit has

$$x_{\min} = 6, \quad i.e. \;\; r_{\min} = 6\frac{GM}{c^2}.$$

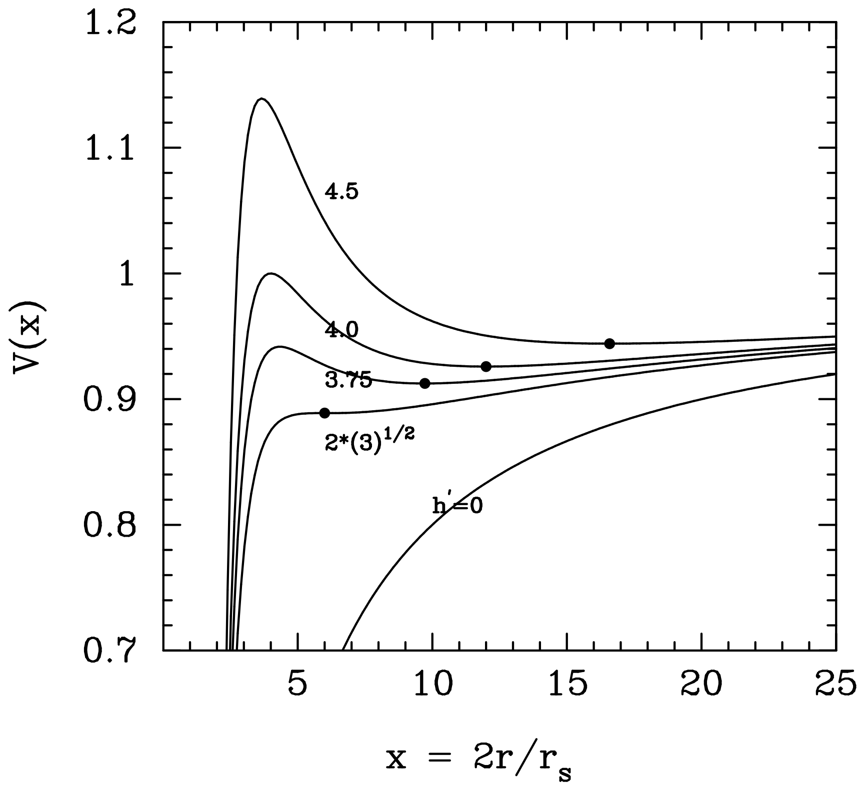

Figure 4: The effective potential (4) plotted for several values of the angular momentum parameter h'.

Gas in an accretion disc settles into circular orbits around the compact object as shown schematically in Figure 2. However, the gas slowly loses angular momentum because of turbulent viscosity (the turbulence is thought to be generated by magnetohydrodynamic instabilities). As the gas loses angular momentum it moves slowly towards the black hole, gaining gravitational potential energy and heating up. Eventually it loses enough angular momentum that it can no longer follow a stable circular orbit and so it falls into the black hole. We can, therefore, make a rough energy of the efficiency of energy radiation in an accretion disc. The maximum efficiency is of order the gravitational binding energy at the smallest stable circular orbit divided by the rest mass enery of the gas

$$\epsilon_{\rm acc} \approx \frac{1}{2}\frac{GMm}{r_{\rm min}}\frac{1}{mc^2} \simeq \frac{1}{12} \sim 8\%.$$

An accretion disc can convert perhaps a few percent of the rest mass energy of the gas into radiation. Compare this with the efficiency of nuclear burning of hydrogen to helium (26 MeV per He nucleus),

$$\epsilon_{\rm nuclear} \sim 0.7\%.$$

Accretion discs are capable of converting rest mass energy into radiation with an efficiency that is about 10 times greater than the efficiency of nuclear burning of hydrogen. The 'accretion power' of black holes cause some of the most energetic phenomena known in the Universe.

13.4 Supermassive black holes

The first quasar[2] (3C273) was discovered in 1963 by Maarten Schmidt. He measured a cosmological redshift of $z = 0.15$ for this object – unprecedently high at the time. (Quasars have since been discovered with redshifts as high as $z = 5.8$) Quasars are very luminous, typically 100– 1000 times brighter than a large galaxy. However, they are *compact*, so compact in fact that quasars look like stars in photographs. In fact, from variability and other studies one can infer that the size of the continuum emitting region of a quasar is of order a few parsecs or less. How can we explain such a phenomenon? Imagine an object radiating many times the luminosity of an entire galaxy from a region smaller than the Solar System. Donald Lynden-Bell was one of the first to suggest that the quasar phenomenon is caused by accretion of gas on to a *supermassive* black hole residing at the centre of a galaxy. The black hole masses required to explain the high luminosities of quasars are truly spectacular – we require black holes with a few million to a few billion times the mass of the Sun.

Do such supermassive black holes exist? The evidence in recent years has become extremely strong. Using the Hubble Space Telescope it is possible to probe the velocity dispersions of stars in the central regions of galaxies. According to Newtonian dynamics, we would expect the characteristic velocities to vary as

$$v^2 \sim \frac{GM}{r}. \tag{5}$$

If the central mass is dominated by a supermassive black hole, then we expect the typical velocities of stars to *increase* as we go closer to the centre. This is indeed what is found in a number of galaxies. From the rate of increase of the velocities with radius, we can estimate the mass of the central object which seems to be correlated with the mass of the bulge component of the galaxy:

$$M_{bh} \approx 0.003 M_{bulge}.$$

It seems as though at the time of galaxy formation, about half a percent of the mass of the bulge material collapses right to the very centre of a galaxy to form a supermassive black hole. During this phase the infalling gas radiates efficiently producing a quasar. When the gas supply is used up, the quasar quickly fades away leaving a dormant massive black hole that is starved of fuel. Nobody has yet developed a convincing theory of how this happens, or of what determined the masses of the central black holes.

[2] *Quasi-stellar radio sources.* We now know that the majority of quasars are radio quiet, and so they are often called QSOs for *quasi-stellar object*.

Table 2: Supermassive Black Holes

		$M_{bh}/M_\odot$	Evidence
$***$	M87	2×10^9	stars & optical disc
$**$	NGC 3115	1×10^9	stars
$**$	NGC 4594 (Sombrero)	5×10^8	stars
$**$	NGC 3377	1×10^9	stars
$*****$	NGC 4258	4×10^7	masing $H_2 0$ disc
$**$	M31 (Andromeda)	1×10^8	stars
$**$	M32	3×10^6	stars
$****$	Galactic Centre	3.6×10^6	stars & 3D motions

A skeptic might argue that these observations merely prove that a dense compact object exists at the centres of galaxies, not necessarily a black hole. But there are two beautiful observational results that probe compact objects on parsec scales – making it almost certain that the central objects are black holes. In our own Milky Way Galaxy it is possible to measure the *proper motions* of stars in the Galactic centre (using infrared wavelengths to penetrate through the dense dust that obscures optical light). This has allowed astronomers to actually see the stars moving and so infer their three dimensional motions (See Figure 5). These observations imply that there exists a black hole of mass $3.6 \times 10^6 M_\odot$ at the centre of our Galaxy.

In a remarkable set of observations, a disc of H_2O masars has been detected in the galaxy NGC 4258 using VLBI. The VLBI observations measure the velocities of the masing clouds on scales of $\sim 0.3 - 2$ parsec and are well fitted by a thin (actually slightly warped) disc in circular motion (See Figure 6). The mass of the central black hole is estimated to be $4 \times 10^7 M_\odot$.

Table 2 lists the masses of some supermassive black holes. I have given the observations a five star rating. The masing disc of NGC 4528 gets a full five stars – this is the strongest observational evidence for a supermassive black hole. The stellar motions in the Galactic Centre get four stars, though some astronomers might argue that this evidence is so strong that it should rate five stars. Most of the other observations are based on measurements of stellar velocity dispersions. This is fairly strong evidence, but not completely convincing[3] and so rate only two stars.

Finally, to end this section, I will show you some results from X-ray spectroscopy. The following picture shows spectral line of iron at X-ray wavelengths measured in the Seyfert galaxy MCG-6-30-15. The iron lines come from the inner parts of the accretion disc around the black hole and the line profile allows one to probe the strong gravitational regime. The

[3]The interpretation of velocity dispersion measurements requires some assumptions about the degree of velocity anisotropy.

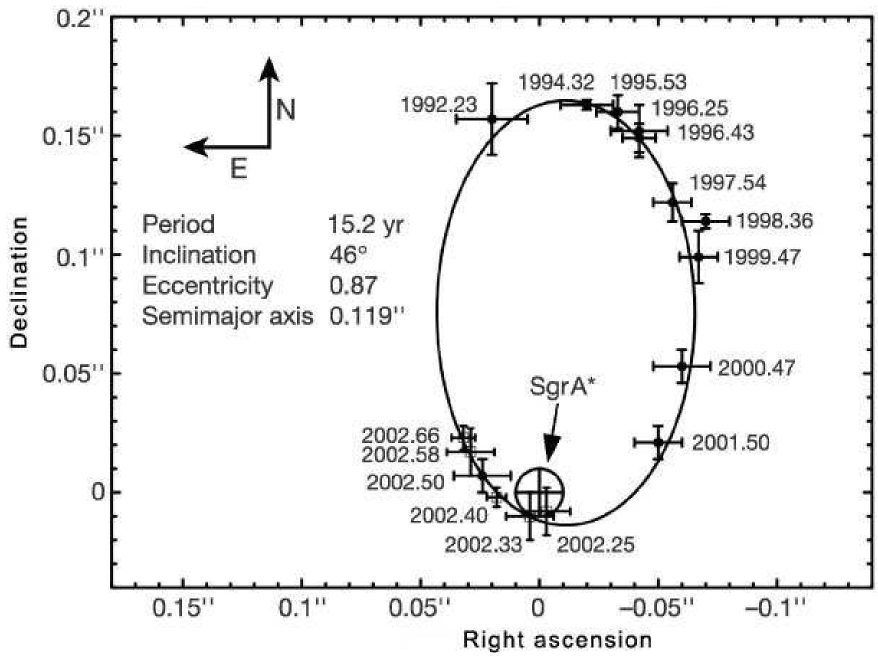

Figure 5: Star moving around black hole with mass of $3.6 \times 10^6 M_\odot$ at the centre of our own Milky Way Galaxy (Schödel et al., Nature, 2002).

lines are assymetric as a result of special relativistic beaming and general relativity. The detailed shape of the line profile depends on the metric and hence on whether the black hole is rotating (see the notes on the Kerr metric). The hope is that the line profiles can be measured and modelled in detail and so be used to infer the *angular momenta* of black holes as well as their masses.

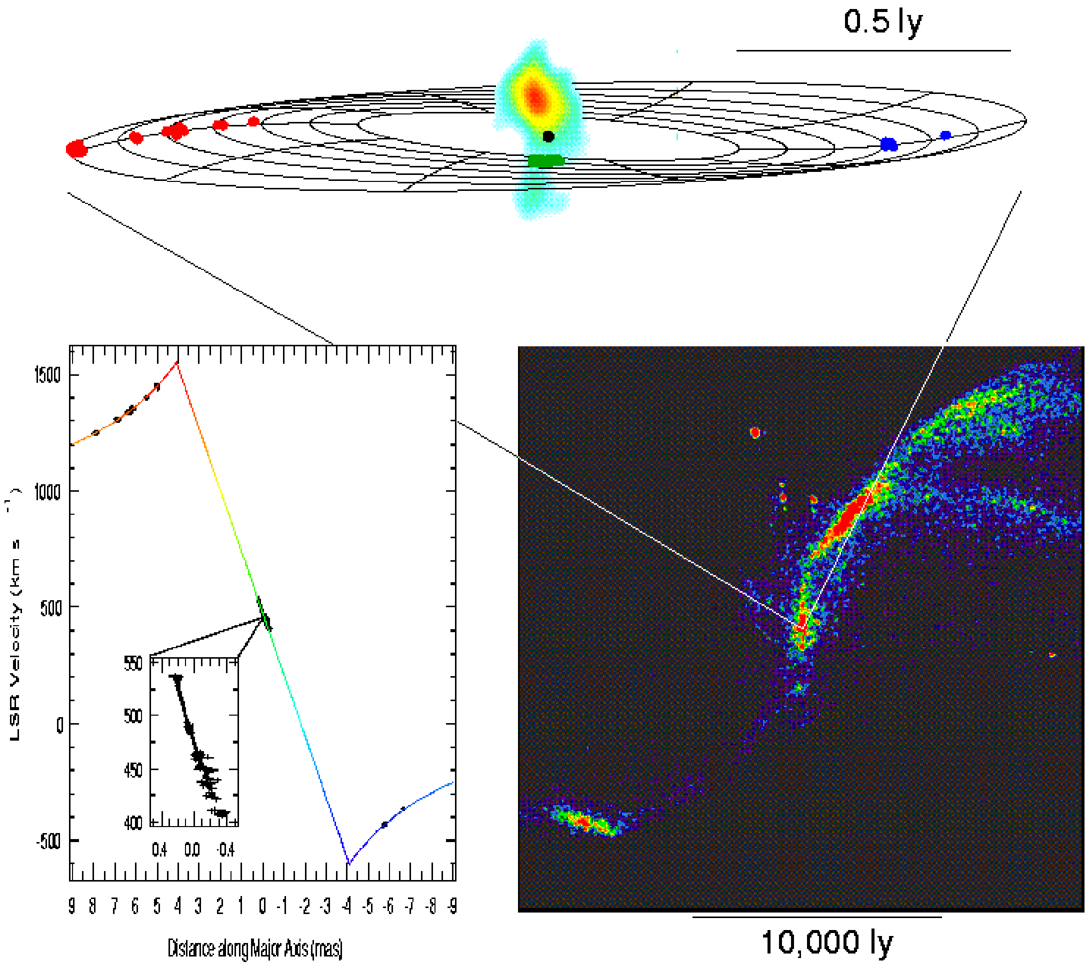

Figure 6: The masing H_20 disc in the centre of NGC4258.

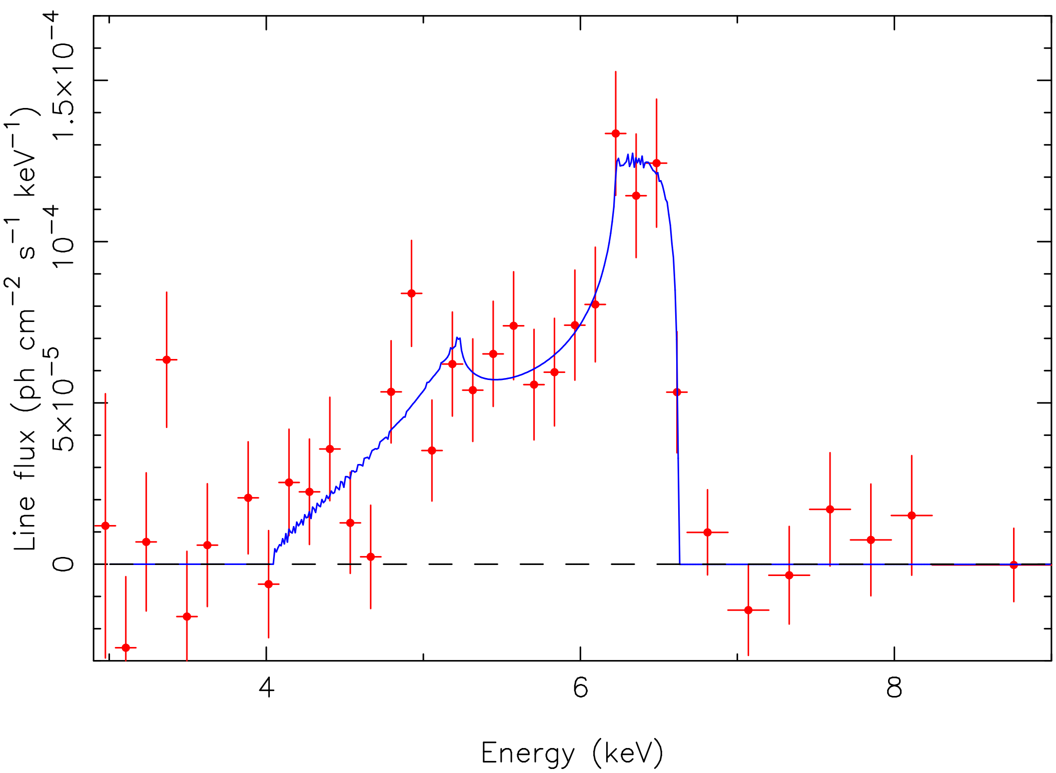

Figure 7: The line profile of iron K-alpha from MCG-6-30-15 observed by the ASCA satellite (Tanaka et al, 1995, Nature, 375, 659). The emission line is extremely broad, with a width indicating velocities of order one-third of the speed of light. There is a marked asymmetry towards energies lower than the rest-energy of the emission line (6.4 keV). This asymmetry is most likely caused by gravitational and relativistic-Doppler shifts near the black hole at the center of the galaxy. The solid line shows the model profile expected from a disk of matter orbiting the hole, extending between 3 and 10 Schwarzschild radii.