Fluorescent iron lines as a probe of astrophysical black hole systems

Christopher S. Reynolds$^a$, Michael A. Nowak$^b$

$^a$Department of Astronomy, University of Maryland, College Park, MD 20742, USA
$^b$MIT, Center for Space Research, NE80-6077, 77 Massachusetts Ave., Cambridge, MA 02139, USA

Accepted 1 November 2002
editor: M.P. Kamionkowski

Abstract

With most physicists and astrophysicists in agreement that black holes do indeed exist, the focus of astrophysical black hole research has shifted to the detailed properties of these systems. Nature has provided us with an extremely useful probe of the region very close to an accreting black hole—X-ray irradiation of relatively cold material in the vicinity of the black hole can imprint characteristic features into the X-ray spectra of black hole systems, most notably the K$_\alpha$ fluorescent line of iron. Detailed X-ray spectroscopy of these features can be used to study Doppler and gravitational redshifts, thereby providing key information on the location and kinematics of the cold material. This is a powerful tool that allows us to probe within a few gravitational radii, or less, of the event horizon.

Here, we present a comprehensive review of relativistic iron line studies for both accreting stellar mass black holes (i.e., galactic black hole candidate systems, GBHCs), and accreting supermassive black holes (i.e., active galactic nuclei, AGN). We begin with a pedagogical introduction to astrophysical black holes, GBHCs, AGN, and accretion disks (including a brief discussion of recent work on the magnetohydrodynamical properties of accretion disks). We then discuss studies of relativistic iron lines in the AGN context, and show how differences between classes of AGN can be diagnosed using X-ray spectroscopy. Furthermore, through a detailed discussion of one particular object (MCG–6-30-15), we illustrate how the exotic physics of black hole spin, such as the Penrose and Blandford-Znajek processes, are now open to observational study. We proceed to discuss GBHCs, which turn out to possess rather more complicated X-ray spectra, making robust conclusions more difficult to draw. However, even in these cases, modern X-ray observatories are now providing convincing evidence for relativistic effects. We conclude by discussing the science that can be addressed by future X-ray observatories.

© 2003 Elsevier Science B.V. All rights reserved.

PACS: 97.60.Lf; 04.70.—s

* Corresponding author.
E-mail address: chris@astro.umd.edu (C.S. Reynolds).

0370-1573/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0370-1573(02)00584-7
Contents

1. Introduction: the astrophysics of relativistic compact objects .......................................................... 391
2. Black holes: theory and astrophysical setting .................................................................................. 393
   2.1. Review of some basic theoretical issues .................................................................................. 393
   2.2. Stellar mass black hole systems and X-ray binaries ............................................................. 395
       2.2.1. The identification of galactic black hole candidates (GBHCs) ...................................... 395
       2.2.2. The X-ray states of GBHCs ...................................................................................... 396
       2.2.3. X-ray transient sources ......................................................................................... 397
   2.3. Supermassive black hole systems ............................................................................................. 398
       2.3.1. The galactic center ..................................................................................................... 400
       2.3.2. Other kinematic studies of SMBHs, and the SMBH census ........................................ 400
       2.3.3. Active galactic nuclei and quasars ............................................................................. 401
3. Accretion disks and disk atmospheres ............................................................................................... 404
   3.1. Introduction to accretion and accretion disks .............................................................................. 404
   3.2. The structure of accretion disks ................................................................................................. 404
       3.2.1. Non-relativistic radiative accretion disks and the $x$-prescription .................................. 405
       3.2.2. Relativistic disks ........................................................................................................... 407
       3.2.3. The magneto-rotational instability and MHD turbulence ............................................. 409
       3.2.4. The role of magnetic fields in the innermost disk ........................................................... 409
   3.3. Accretion disk coronae .............................................................................................................. 411
   3.4. Radiatively inefficient accretion disks ......................................................................................... 415
   3.5. Disk atmospheres and X-ray reprocessing features ................................................................. 416
       3.5.1. Simple X-ray reflection models ..................................................................................... 417
       3.5.2. Physical X-ray reflection models .................................................................................. 418
4. Emission lines from accretion disks ................................................................................................... 421
   4.1. Formalisms for computing line profiles ...................................................................................... 421
   4.2. The emerging line profile .......................................................................................................... 422
5. The X-ray astronomer’s arsenal ......................................................................................................... 425
   5.1. Past and current ......................................................................................................................... 425
   5.2. Past and current X-ray observatories ......................................................................................... 425
       5.2.1. The Chandra X-ray observatory .................................................................................... 426
       5.2.2. XMM-Newton .............................................................................................................. 427
   5.3. Future missions .......................................................................................................................... 427
       5.3.1. Astro-E II ....................................................................................................................... 427
       5.3.2. Constellation-X ............................................................................................................ 427
       5.3.3. XEUS ............................................................................................................................ 428
       5.3.4. MAXIM ........................................................................................................................ 428
6. Iron lines from active galactic nuclei .................................................................................................. 428
   6.1. A case study of MCG–6-30-15 ................................................................................................. 429
       6.1.1. The first detection of a relativistic accretion disk ............................................................ 429
       6.1.2. Is the black hole rapidly spinning? .................................................................................. 431
       6.1.3. Is the spinning black hole energizing the disk? ............................................................... 434
       6.1.4. Relativistically broadened soft X-ray emission lines ....................................................... 435
   6.2. Other “normal” Seyfert nuclei ...................................................................................................... 436
   6.3. Very-high and very-low luminosity AGN .................................................................................... 440
   6.4. Iron lines from radio-loud AGN: a clue to the dichotomy? ......................................................... 441
7. Iron lines from galactic black hole candidates .................................................................................... 442
   7.1. A case study of Cygnus X-1 ...................................................................................................... 442
   7.2. Other GBHC and spectral correlations ....................................................................................... 448
8. Future prospects for iron line studies ................................................................................................ 451
1. Introduction: the astrophysics of relativistic compact objects

Gravitational collapse of normal matter can produce some of the most exotic objects in the universe—neutron stars and black holes. Proving that these objects exist in Nature occupied theoretical and observational astrophysicists for much of the 20th century. Most of the detailed debate centered around understanding the possible final states of massive stars. On his now famous sea voyage from India to England in 1930, Subrahmanyan Chandrasekhar considered the structure of white dwarf stars—compact stellar remnants in which gravitational forces are balanced by electron degeneracy pressure. He realized that, if the white dwarf was sufficiently massive, the degenerate electrons will become relativistic thereby rendering the star susceptible to further gravitational collapse [1,2]. Although hotly debated by Arthur Eddington, Chandrasekhar correctly deduced that a white dwarf would undergo gravitational collapse if its mass exceeded $M_{\text{ch}} \approx 1.4 M_\odot$ (where $1 M_\odot = 2.0 \times 10^{33} \text{ g}$ is the mass of the Sun), a limit now known as the Chandrasekhar limit.\footnote{The Russian physicist Lev Landau independently calculated the upper mass limit of a white dwarf star, and his results were published in 1932 [3].}

Once gravity overwhelms electron degeneracy pressure, it is thought that neutron degeneracy pressure is the last, best hope for averting total gravitational collapse. Objects in which gravitational forces are balanced by neutron degeneracy pressure are called neutron stars. Although there was initial hope that nuclear forces would always be sufficient to resist gravity, the upper limit to the mass of a neutron star is now believed to be in the range $M \approx 1.8–2.2 M_\odot$ [4,5]. Uncertainties arising from the equation of state at super-nuclear densities continue to plague our determination of this critical mass, but an absolute upper limit of $M \approx 4 M_\odot$ arises from very general considerations, i.e., the validity of General Relativity and the principle of causality [6]. Above this mass, it is thought that complete gravitational collapse cannot be avoided. In particular, Hawking’s singularity theorems [7] show that the formation of a spacetime singularity is unavoidable (irrespective of the mass/energy distribution) once the object is contained within the light trapping surface. The result is a black hole, i.e., a region of spacetime bounded by an event horizon and, at its heart, possessing a spacetime singularity.

While the above considerations now have a firm theoretical base, observational astrophysics was, and continues to be, critically important in guiding our understanding of such extreme objects. In the case of both neutron stars and black holes, the very existence of these objects was only widely accepted when compelling observational evidence was forthcoming. For neutron stars, the pivotal observation was the discovery of pulsars by Jocelyn Bell and Anthony Hewish via radio observations taken from Cambridge. Black holes gained wide acceptance after it was demonstrated that the X-ray emitting compact object in the binary star system Cygnus X-1 did, in fact, possess a mass in excess of the maximum possible neutron star mass [8–12]. This made it the first of the so-called Galactic Black Hole Candidates (GBHCs), a class that has now grown to include some two dozen objects.
We now know of another class of black holes—the supermassive black holes, with masses in the range of $10^5-10^9 M_\odot$, that reside at the dynamical centers of most, if not all, galaxies.\(^2\) Today, by far the strongest case for a supermassive black hole can be made for our own Galaxy. Modern high-resolution, infra-red imaging reveals that the stars in the central-most regions of our Galaxy are orbiting an unseen mass of $2.6 \times 10^6 M_\odot$\(^{[14-16]}\). Furthermore, studies of the orbital dynamics (which now include measured accelerations as well as velocities, \(^{[17,18]}\)) constrain the central mass to be extremely compact. According to conventional physics, the only long-lived object with these properties is a supermassive black hole. Alternatives, such as a compact cluster of neutron stars, would suffer a dynamical collapse on short time scales \(^{[19]}\).

Having established beyond reasonable doubt that black holes exist, it is obviously interesting to perform detailed observational studies of them. The regions in the immediate vicinity of a black hole bear witness to complex interactions between matter moving at relativistic velocities, electromagnetic fields, and the black hole spacetime itself. Given that the apparent angular scales of even the biggest black hole event horizons are $\sim 10^{-6}$ arcsec, direct imaging studies of these regions will not be possible for many years.\(^3\) In the meantime, we must study these regions using more indirect methods, chief among which are spectroscopic methods.

As we will detail in this review, Nature has provided us with a well-understood and extremely useful spectral diagnostic of matter in the near vicinity of astrophysical black holes. In essence, relatively cold matter in the near vicinity of an astrophysical black hole will inevitably be irradiated by a spectrum of hard X-rays \(^{[21,22]}\). The result can be a spectrum of fluorescent emission lines, the most prominent being the K\(\alpha\) line of iron at an energy of 6.40–6.97 keV (depending upon the ionization state of the iron) \(^{[23–25]}\). Ever since the launch of the Advanced Satellite for Cosmology and Astrophysics (ASCA) in February 1993, X-ray astrophysicists have had the capability to identify this emission line and measure its spectral profile. Fig. 1 shows the iron line in the X-ray emissions originating near the supermassive black hole in the galaxy MCG–6-30-15 \(^{[20]}\). Bearing in mind that the line is intrinsically narrow with a rest-frame energy of 6.4 keV, it can be seen that the line has been dramatically broadened and skewed to low-energies. It is now widely accepted that the line originates from material that is just a few gravitational radii from the black hole, and possesses a profile that is shaped by (relativistic) Doppler shifts and gravitational redshift effects. Investigating these spectral features in X-ray luminous black hole systems has given us the clearest window to date on the physics that occurs in the immediate vicinity of astrophysical black holes.

The intent of this review is to describe our current understanding of black hole astrophysics, with an emphasis on what has been learnt by utilizing these X-ray spectral signatures. We begin by discussing the basic theoretical framework within which we understand the astrophysical environment around black holes. A central and important part of this discussion is an introduction to the modern theory of accretion disks. Hand-in-hand with the theoretical discussion, we will introduce
Fig. 1. Continuum subtracted iron line from the long July-1994 *ASCA* observation of the Seyfert-1 galaxy MCG–6-30-15 [20]. The dashed line shows a model consisting of iron line emission from a relativistic accretion disk around a non-rotating (Schwarzschild) black hole, with a disk inclination of $i = 30^\circ$, and an emissivity profile of $r^{-3}$ extending down to the radius of marginal stability ($6 GM/c^2$).

...the necessary phenomenology associated with stellar mass and supermassive black holes. We then describe the array of past, current, and future X-ray observatories which have bearing on relativistic studies of black holes before discussing how iron line spectroscopy has dramatically improved our current understanding of black hole astrophysics. We conclude by presenting the prospects for future research in this field.

2. Black holes: theory and astrophysical setting

2.1. Review of some basic theoretical issues

Throughout this review, we shall assume that standard General Relativity is valid for all regions of interest. Under this conservative assumption, it is a remarkable fact that the relatively simple Kerr metric [26] precisely describes the spacetime outside of almost any astrophysical black hole.\(^4\) The Kerr metric is described by only two parameters which can be chosen to be the mass $M$ and angular momentum $J$ of the black hole. The no-hair theorem of General Relativity tells us that black holes can also possess electrical charge. However, in any astrophysical setting, a black hole with appreciable charge will rapidly discharge via vacuum polarization. In Boyer–Lindquist coordinates

\(^4\) A dramatic exception to this statement occurs when two black holes of comparable mass merge. Large and complicated deviations from the Kerr metric can be readily obtained during such an event. However, such events are extremely rare and settle down to the Kerr metric on only a few light crossing times of the event horizon.
(the natural generalization of spherical polar coordinates), the line element of the Kerr metric is given as

\[ ds^2 = - \left( 1 - \frac{2Mr}{\Sigma} \right) dt^2 + \frac{4aMr \sin^2 \theta}{\Sigma} dt d\phi + \frac{\Sigma}{\Delta} dr^2 + \Sigma d\theta^2 + \left( r^2 + a^2 + \frac{2a^2Mr \sin^2 \theta}{\Sigma} \right) \sin^2 \theta d\phi^2 , \]

where \( a = \frac{J}{M} \), \( \Delta = r^2 - 2Mr + a^2 \), \( \Sigma = r^2 + a^2 \cos^2 \theta \) and we have chosen units such as to set \( G = c = 1 \). The event horizon of the black hole is given by the outer root of \( \Delta = 0 \), i.e., \( r_+ = (M + \sqrt{M^2 - a^2}) \). The special case of \( a = 0 \) corresponds to non-rotating black holes, reducing the metric to the Schwarzschild (1916) metric

\[ ds^2 = - \left( 1 - \frac{2M}{r} \right) dt^2 + \left( 1 - \frac{2M}{r} \right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 , \]

with the event horizon at \( r = 2M \).

A feature of the black hole spacetime that is crucially important for accretion disk models is the existence of a radius of marginal stability, \( r_{ms} \). This is the radius within which circular test particle orbits are no longer stable. We shall refer to the region \( r < r_{ms} \) as the plunging region, since, in most astrophysical settings, it will be occupied by material that is in the process of plunging into the black hole. For a given black hole mass, this radius is a function of the black hole’s angular momentum and, for orbits in the equatorial (\( \theta = \pi/2 \)) plane, is given by

\[ r_{ms} = M(3 + Z_2 \mp [(3 - Z_1)(3 + Z_1 + 2Z_2)])^{1/2} , \]

where the \( \mp \) sign is for test particles in prograde and retrograde orbits, respectively, relative to the spin axis of the black hole and we have defined

\[ Z_1 = 1 + \left( 1 - \frac{a^2}{M^2} \right)^{1/3} \left[ \left( 1 + \frac{a}{M} \right)^{1/3} + \left( 1 - \frac{a}{M} \right)^{1/3} \right] , \]

\[ Z_2 = \left( 3 \frac{a^2}{M^2} + Z_1^2 \right)^{1/2} . \]

Effects of the black hole spin that arise as a result of the \( dt d\phi \) term in the line element correspond to the dragging of inertial frames. An extreme manifestation of frame-dragging is the existence of a region (the ergosphere; given by \( r < M + \sqrt{M^2 - a^2 \cos^2 \theta} \)) in which all time-like particles must rotate in the same sense as the black hole as seen by an observer at infinity. The frame-dragging becomes more extreme as one approaches the event horizon until, at the horizon, all time-like particles orbit with an angular velocity of \( \Omega_H = a / 2Mr_+ \). This is referred to as the angular velocity of the event horizon. Another peculiar feature of the Kerr metric is the existence of negative energy orbits within the ergosphere. If a particle is placed on such an orbit via some interaction within the ergosphere, it will fall into the black hole and diminish the mass and angular momentum of the black hole. This gives rise to the Penrose-process for extracting the rotational energy of a spinning black hole [28]. A consideration of black hole thermodynamics readily shows that, in principle, an energy
of up to 29% of the mass energy of the hole, i.e., $0.29M$, can be extracted from a maximally rotating black hole, thereby reducing its angular momentum to zero (e.g., see the treatment in [29]).

In principle, a Kerr black hole can possess a spin parameter arbitrarily close to $a/M = 1$ without violating the cosmic censorship hypothesis. However, as argued by Kip Thorne [30], a rapidly spinning black hole that is actively accreting matter will preferentially capture photons with negative angular momentum, thereby limiting the spin parameter to $a/M \approx 0.998$ (depending somewhat on the angular distribution of the photons emitted by the accreting matter). For this reason, most of the astrophysical literature uses the term “extremal” or “near-extremal” Kerr black hole to refer to a black hole with $a/M = 0.998$. We shall follow this convention.

2.2. Stellar mass black hole systems and X-ray binaries

Now that we have introduced the basic theory of black holes, we present some of the phenomenology of real-life astrophysical black holes. In fact, the scientific communities studying stellar mass and supermassive black holes have remained rather distinct from each other, leading to the development of rather different languages and phenomenological descriptions. We begin by discussing the stellar mass black holes that we believe arise from the collapse of a massive star.

Although most observed stellar mass black hole candidates reside in binary star systems, it is expected that the vast majority of such black holes occur as isolated systems. They remain undetected due to their lack of any significant radiation. Current estimates indicate that there are between 0.1–1 billion such systems within our Galaxy, a very small fraction of which, $< 0.001\%$, might be detectable in a deep X-ray survey of the Galactic plane [31]. Even a so-called isolated black hole will accrete at some level from the surrounding gaseous interstellar medium (ISM)—this mode of accretion is called Bondi–Hoyle accretion after Herman Bondi and Fred Hoyle who calculated the details of this process [32]. Low-level X-ray emission is expected due to Bondi–Hoyle accretion, although there are large uncertainties regarding both the mass flux and radiative efficiency of the accretion flow [31]. Currently, the best prospects of detecting such isolated black holes is via their gravitational effects, specifically via gravitational micro-lensing—the amplification of the observed light from a background star due the gravitational focusing (i.e., bending) of that light in the potential of a foreground compact object. The amplification is expected to be time-symmetric, achromatic (i.e., independent of wavelength), and of a duration from weeks to months for expected stellar distances, velocities, and compact object masses. A number of surveys for such micro-lensing events are currently being undertaken, and there are claims of candidate detections of stellar mass black holes [33,34].

2.2.1. The identification of galactic black hole candidates (GBHCs)

By contrast to these isolated black hole systems, the first GBHC discovered was in a binary star system. A rocket borne X-ray detector, launched in 1964, discovered the persistently bright X-ray source Cygnus X-1 [35]. Its exact nature was not clear, however, until its optical identification with the O-type star HDE 226868 [36–38]. Based upon measurements of the optical companion’s orbital parameters, mass estimates for the compact object have ranged from 8 to 16 $M_{\odot}$, with recent estimates placing the mass at approximately 10 $M_{\odot}$ [9–12]. The chief uncertainties in the compact object mass determination are due to some uncertainty in the mass of the normal companion star, and due to the unknown inclination of the system’s orbital plane, with estimates of the latter being
scattered from 26 to $67^\circ$ \cite{39–42}. All estimates, however, place the mass of the compact object above the theoretical maximum mass for a neutron star. The tremendous X-ray luminosities of GBHCs (the X-ray luminosity of Cyg X-1 is approximately $10^9$ times larger than the total luminosity of the Sun) are thought to be due to accretion from the companion star onto the compact object. As we will discuss in Section 3, accretion can be a very efficient process for converting a mass flux into a persistent luminosity. Prior to the recent micro-lensing studies, all known GBHCs were discovered as X-ray sources in binary systems.

The identification of Cyg X-1 as a GBHC was followed by a series of discoveries of other persistent or quasi-persistent X-ray sources that were later identified as GBHCs. These discoveries included the systems LMC X-1 \cite{43–46}, LMC X-3 \cite{43,47}, GX 339-4 \cite{48,49}, and 4U 1957 + 11 \cite{43,50}, the “U” standing for Uhuru, an early 1970s X-ray satellite that scanned much of the sky. The first two objects reside approximately 50 kpc away in the nearby satellite galaxy, the Large Magellanic Cloud. Similar to Cyg X-1, the black hole classifications of LMC X-1 and LMC X-3 rest upon dynamical mass estimates of 6 and 9 M$_\odot$, respectively \cite{45,47}. The GBHC classifications of GX 339-4 and 4U 1957 + 11, on the other hand, come from the fact that their X-ray spectral and variability properties are very similar to other GBHCs that do have dynamical mass estimates. Specifically, GX 339-4 is very similar to Cyg X-1 as well as other systems \cite{51,52}, while 4U 1957 + 11 (whose black hole candidacy is still vigorously being debated) is very similar to LMC X-3 \cite{53,54}. These analogies among the properties of X-ray sources has led to the concept of ‘X-ray states’ with distinct spectral and temporal properties.

2.2.2. The X-ray states of GBHCs

Broadly speaking, the states of stellar mass black hole candidates are described in terms of their ‘spectral hardness’ and their short time scale ($\approx 10^{-3}$–$10^3$ s) variability properties. Typically, the spectra at energies greater than 10 keV are roughly described by a power law (which may or may not have a detectable high energy ‘break’). The slope of this power-law is often described via the photon index, $\Gamma$, where the photon number flux per unit energy (photons cm$^{-2}$ s$^{-1}$ keV$^{-1}$) is

$$F_N(E) \propto E^{-\Gamma},$$

where $E$ is the photon energy. The luminosities of ‘hard states’ are dominated by emission from energies above 10 keV, and their spectra have $\Gamma < 2$ \cite{55}. The luminosities of ‘soft states’ are dominated by emission from energies less than 3 keV. Their spectra are usually described by a quasi-thermal component with characteristic temperature between $\approx 0.5$ and 3 keV and, occasionally in addition to the low energy X-ray emission, high-energy power-law spectra with $\Gamma > 2$ \cite{56}. At low luminosities, stellar mass black hole systems tend to exhibit strongly variable X-ray lightcurves, with root mean square (rms) fluctuations of $\approx 40\%$, and ‘hard’ X-ray spectra. At high luminosities, these systems tend to show more weakly variable (rms $\leq 20\%$) X-ray lightcurves and soft X-ray spectra. There are a number of reviews describing in more detail the properties of these states, with the hard and soft states often being further divided into various sub-states \cite{20,57,58, and references therein}.

Cyg X-1, GX 339-4, and LMC X-3 each show transitions between the less variable soft states, and highly variable hard states \cite{52,60–66}. Whereas Cyg X-1 exhibits relatively little luminosity evolution during its state transitions \cite{61,62}, GX 339-4 exhibits large luminosity fluctuations \cite{52,67}, with the source at times being nearly undetectable in the X-ray sky \cite{63,68}. Fig. 2 illustrates these points with
radio and X-ray data taken during a typical state transition of GX 339-4. From this perspective, GX 339-4 is sometimes classified as a transient X-ray source. In contrast to sources such as Cyg X-1, in fact the vast majority of identified stellar mass black hole systems are transient X-ray sources, with many of these transients falling into the class of `X-ray novae’.

2.2.3. X-ray transient sources

As an example of an X-ray nova, we shall describe the system GS 1124 − 683, a.k.a. Nova Muscae. Its behavior was characteristic of many X-ray transients observed prior to the launch of the Rossi X-ray Timing Explorer (RXTE). In the span of less than 10 days, the 1–10 keV source flux rose by more than a factor of 100 [69]. At its maximum, the X-ray spectrum exhibited a strong, soft quasi-thermal component and a power law tail with $\Gamma \approx 2.5$ [56,70]. The flux then exponentially decayed with a time constant of $\approx 30$ days for the soft X-ray flux and $\approx 13$ days for the hard flux. The X-ray variability was also seen to decrease as the power-law component flux decreased [56,70]. Approximately 5 months after maximum flux, the source switched into a ‘hard state’ with $\Gamma \approx 1.7$ and then continued to decay.

After the X-ray source decayed away, optical observations revealed modulations on an orbital period of 10.4 h. Combining this with velocity measurements of the companion star, the lower limit
to the compact object mass was determined to be $3.1 \, M_\odot$ and its most likely value was determined to be $6 \, M_\odot$ [71]. Prior to those observations, several other quiescent X-ray novae had revealed dynamical evidence for massive compact objects. GS 2023+33, a.k.a. V404 Cyg was found to have a compact object mass of $10 \, M_\odot$, and A0620 − 00 was found to have a compact object mass of $7.3 \, M_\odot$ (lower limit $3.2 \, M_\odot$) [72,73]. There have been approximately two dozen such transients, roughly half of which have measured masses. To date, our greatest source of information about the mass distribution of stellar mass black holes has come from optical observations of transient systems in quiescence [74]. The X-ray nova outburst acts like a beacon calling attention to the system, while the subsequent decay into quiescence allows the orbital parameters of the binary system to be measured via observations of the typically very faint companion star.

X-ray transients that deviate strongly from the ‘fast-rise/exponential decay’ behavior of Novae Muscae also have been discovered. A wide variety of lightcurve morphologies have been observed by RXTE [75]. Among the most interesting examples are GRO J1655−40 and GRS 1915+105, both of which were found to exhibit relativistic jets of plasma that were observed in the radio band with inferred velocities in excess of $0.9c$ [76,77]. Both of these sources have shown dramatic variability (see Fig. 3), and the latter source remains in an active X-ray state to the present day. It has since been realized that the formation of radio jets or outflows is often associated with accretion in stellar mass black hole systems [78]. Typically, the radio flux is intermittent and ‘optically thin’ (i.e., $F_v \propto v^{-x}$, with $x = \Gamma - 1 > 0$, where $F_v$ is the radio energy flux per unit frequency, ergs cm$^{-2}$ s$^{-1}$ Hz$^{-1}$) at the highest fluxes within the soft state, becomes quenched at lower luminosities in the soft state, and then re-occurs in the hard state. The radio spectrum in the X-ray hard state is usually persistent and ‘flat’ or ‘inverted’ (i.e., $F_v \propto v^{-x}$ with $x \leq 0$), and the radio flux is positively correlated with the X-ray flux [59,78–82].

There is currently a great deal of controversy regarding the underlying mechanisms behind these state transitions, the formation of the radio outflow, as well as the basic geometry of the accretion flow itself. Observations of relativistically distorted X-ray emission lines, the prime focus of this review, are being used as probes of the accretion flow properties near the event horizons of stellar mass black holes in binary systems precisely for the purpose of gaining more information about these issues.

2.3. Supermassive black hole systems

We now turn to the basic phenomenology of supermassive black holes (SMBHs). Although it is artificial to place strict bounds on any definition, the term SMBH usually refers to black holes with a mass greater than $10^5 \, M_\odot$. These objects are always found to reside at the dynamical center of their host galaxy. This is readily understood through the action of dynamical friction [83]. Suppose the SMBH was not located at the center of the host galaxy and, instead, was orbiting within the galaxy’s gravitational potential. As the SMBH moves through and gravitationally interacts with the background of stars (which have much lower mass) and dark matter, it gravitationally induces a wake in its trail in which the stellar and dark matter densities are enhanced. The resulting gravitational force between the wake and the SMBH itself acts as a drag force on the SMBH. The effective drag force scales as $M^2$ and hence, for a sufficiently massive black hole, is able to extract energy from the orbit of the SMBH on a fairly short time scale and make it come to rest at the bottom of the gravitational potential well, i.e., at the dynamical center of the galaxy.
Unlike the case of typical stellar mass black holes, it is far from clear how supermassive black holes formed. While it is clear that accretion has been responsible for significant growth of supermassive black holes [84], the origin of the initial massive “seed” black holes is highly controversial. The interested reader is pointed to the seminal review by Martin Rees [85].
2.3.1. The galactic center

While the proximity of our Galactic Center (only 8 kpc away) is obviously a big advantage for any study of the SMBH residing there, it is a difficult region to study for other reasons. The substantial quantities of dust in the plane of the Galactic disk (and hence between us and the Galactic Center) extinguish the optical light from the Galactic Center by a factor of $10^{10}$, rendering it basically invisible. Progress must be made by utilizing wavelengths that can more readily penetrate the Galactic dust—in particular, radio, infra-red, hard X-rays and $\gamma$-rays.

In recent years, dramatic progress has been made by high spatial resolution studies of the near infra-red emission from bright stars in the central-most regions of the Galaxy. By achieving diffraction-limited near infra-red (K-band) images of the Galactic Center (with resolutions of 60 milli-arcsecs) over the time span of several years, one can directly observe the motions of rapidly-moving stars within the central stellar cluster that inhabits the Galactic Center. One finds that the velocities of stars within the central regions of this cluster drops as the square-root of the distance from the Galactic Center, implying a gravitationally dominant mass of $2.6 \times 10^6 M_\odot$ confined to the central $10^{-6}$ pc$^3$ [14,15]. A supermassive black hole is the only long-lived object of this mass and compactness—a compact cluster of stellar-mass objects would undergo dynamical collapse on a time scale of $10^{7}$ yr or less [19]. More recently, these proper motion studies have also detected acceleration of several stars in the Galactic Center region [17,18]. The supermassive black hole hypothesis survives the powerful consistency check allowed by a vector analysis of these accelerations. Any remaining doubt seems to have been removed by recently reported observations that show a star passing within 17 light hours of the putative SMBH [16]. The star remains on a Keplerian orbit, even at the peri-center of the orbit where it achieves a velocity exceeding 5000 km s$^{-1}$. This raises the implied mass-density to more than $10^{17}$ M$_\odot$ pc$^{-3}$—if this were a cluster of compact stellar remnants, it would evaporate or collapse on a timescale of less than $10^5$ yr.

2.3.2. Other kinematic studies of SMBHs, and the SMBH census

Kinematic studies of stars around the SMBH in the center of our own Galaxy have only been possible in recent years with the development of diffraction limited near infra-red imaging on large ground-based telescopes. Before that, studies of the central regions of other galaxies (where the obscuration problem is much less severe) were crucial for establishing the existence of SMBHs. In particular, there were two studies published in the mid-1990s that were pivotal in providing compelling evidence for SMBHs—the Hubble Space Telescope (HST) observations of M87 and the radio observations of the H$_2$O MASERs in NGC 4258.

The center of M87 was an important target for HST since a central SMBH had long been suspected based upon previous ground-based optical observations of the galaxy’s central regions [86,87], as well as due to the presence of a prominent synchrotron emitting jet of plasma that flows away from the galaxy’s core at relativistic speeds. HST imaged a disk of ionized gas, with a radius of $\sim 50$ pc centered on the galactic core [88]. The high resolution of HST allowed the spectrum of this ionized gas to be measured as a function of position across the gas disk, thereby allowing the kinematics of the disk to be determined [89]. It was found that the velocity profile of the central 20 pc of the gas disk possessed a Keplerian profile (i.e., $v \propto r^{-1/2}$) as expected if the gas was orbiting in the gravitational potential of a point-like mass [89,90]. From the measured velocities of this disk ($\sim 1000$ km s$^{-1}$), the central mass was determined to be $3 \times 10^9 M_\odot$. The only known and
long-lived object to possess such a large mass in a small region of space, and be as under-luminous as observed, is a SMBH.

Similar conclusions were drawn for the center of the galaxy NGC 4258, albeit using very different observational techniques [91,92]. The center of this galaxy contains molecular gas which is subjected to heating by a central X-ray source. Collisional pumping of water within the molecular gas leads to a population inversion thereby driving a naturally occurring maser. Both the spatial position and line-of-sight velocity of the masing blobs can be measured very accurately using radio observations (in particular, the Very Long Baseline Array, VLBA). It is found that the masers lie in a very thin disk (oriented almost edge-on to us) that is orbiting a central object of mass $3.6 \times 10^7 M_\odot$ with an almost perfect Keplerian velocity profile (with the velocity of the inner masing region being $\sim 1000 \text{ km s}^{-1}$). The accuracy with which the velocity profile follows a Keplerian law tightly constrains the spatial extent of the central mass, again rendering any explanation other than a SMBH very problematic. The direct detection of centripetal accelerations within the masing regions provides a powerful consistency check [93].

While the Galactic Center, M87 and NGC 4258 are important cases in the argument for the existence of SMBHs, it is difficult to generalize these observational techniques to all galaxies. A more generally applicable technique is to examine the ensemble velocity distribution of stars in the central region of a galaxy via observations of stellar absorption lines in galactic spectra. By carefully comparing detailed galactic models (that include the distribution of stars across the possible phase space of orbits within a given gravitational potential) with high quality imaging and spectral data, one can constrain the mass of any central black hole. Using these techniques, several authors have performed relatively large surveys of nearby galaxies in order to examine the demographics of SMBHs. There are two exciting results from these studies. Firstly, it appears that every galaxy that possesses a well defined bulge contains a SMBH. Secondly, there is a good correlation between the mass of the central black hole $M$ and the mass of the galactic bulge ($M_{\text{bulge}}$), with $M/M_{\text{bulge}} \sim 0.003$ [94,95]. In fact, it has been shown that the more fundamental (and better) correlation is between the mass of the central SMBH and the velocity dispersion $\sigma$ (or “temperature”, considering an analogy between stellar kinematics in a galaxy and particle kinematics in a gas) of the stellar population; $M \propto \sigma^b$, with $b = 4-5$ [96,97]. The underlying cause for this correlation is still the subject of intense work and much debate, since these stars are far enough from the center of the galaxy to have negligible direct influence from the gravitational field of the SMBH. It is widely regarded that this correlation argues for a connection between the formation of the SMBH and the galaxy itself.

### 2.3.3. Active galactic nuclei and quasars

Active galactic nuclei (AGN) were known and studied before the concept of SMBHs became firmly established. Observationally, an AGN is defined as a galactic nucleus which displays energetic phenomena such as large electromagnetic luminosities and/or powerful jets. The first known AGN was the “radio quasar” 3C 273. This object was first identified by radio observations, but radio imaging alone was insufficient to localize its position on the sky well enough to allow follow-up investigation with optical telescopes. A major breakthrough was the use of lunar occultation techniques (i.e., precision measurement of the time at which the radio emissions from 3C 273 were blocked as the Moon passed in front of it) by Cyril Hazard and collaborators [98] which localized 3C 273 to within 1 arcsec. This allowed an identification and subsequent spectroscopy of the corresponding optical object. The optical spectrum was initially confusing, displaying emission lines at wavelengths
that did not correspond to any known atomic transition, until it was realized by M. Schmidt that the spectrum was redshifted by a seemingly enormous factor \[99\]

\[
z \equiv \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}} = 0.158.
\] (7)

Interpreting such a redshift as cosmological (i.e., due to the expansion of the Universe), suggests the object to be extremely distant and, therefore, extremely luminous. These enormous luminosities were the clue that led some researchers, most notably Salpeter, Zeldovich and Lynden-Bell, to suggest the existence of accreting SMBHs \[100–102\]. Accretion of matter into a relativistically deep gravitational potential well is one of the few processes in nature efficient enough to produce such luminosities without having to process unreasonably large amounts of mass. The supermassive nature of the black holes follows from the requirement for the AGN to remain gravitationally bound (and hence long-lived) despite the enormous outward pressure exerted by the electromagnetic radiation.

After four decades of study, there is a large and complex phenomenology associated with AGN. It has proven useful to classify AGN according to their luminosity, electromagnetic spectrum, and spatial (radio) morphology.\(^5\) A basic dichotomy seems to exist between radio-loud AGN and radio-quiet AGN. Radio-loud AGN (of which 3C 273 and M87 are examples) possess back-to-back twin jets of plasma that are produced in the vicinity of the SMBH and propagate outwards at relativistic velocities (with Lorentz factors of 5–10 or more) for, in some cases, distances of \(10^5–10^6\) pc. These jets tend to be copious sources of radio and X-ray emission due to, respectively, synchrotron and inverse Compton emission by high-energy electrons within the magnetized plasma. The primary physical mechanisms initiating, accelerating and collimating these jets are still far from clear, although it seems likely that they are launched from the inner accretion disk or ergosphere of the SMBH and are accelerated/collimated by hydromagnetic processes. In the rare cases where one of the jets is directed straight at us, special relativistic beaming strongly enhances the observed jet emission, often diluting and thus making it difficult to observe any other emissions from the AGN. Such objects are known as blazars (of which the well-studied BL-Lac objects are a sub-class) and can be sources of extremely energetic emissions (with photon energies up to \(\sim 10\) TeV having been detected from the BL-Lac objects Mrk 501 and Mrk 421 \[104,105\]).

Radio-quiet objects, on the other hand, seem not to possess these powerful jets. Radio-quiet AGN with moderate electromagnetic luminosities (\(L \sim 10^{43}–10^{45}\) erg s\(^{-1}\)), commonly referred to as Seyfert galaxies, are a particularly important and well studied sub-class. This is due to the fact that they are reasonably common (constituting 1–10\% of all major galaxies), leading to some relatively nearby and easily studied examples. Seyfert galaxies themselves have been classified into two broad categories. Those objects in which we can directly view the energetic regions immediately around the SMBH are called type-1 Seyfert galaxies (often referred to as Seyfert-1 galaxies). On the other hand, if the SMBH region is obscured by large amounts of dust and gas (situated near to the AGN and/or within the host galaxy), it is called a type-2 Seyfert galaxy. Spectropolarimetry of some nearby Seyfert galaxies (in particular NGC 1068 \[106\]) provides evidence that the same AGN may appear as a type-1 or type-2 depending upon the orientation at which we view the system. Two possible geometries that are often discussed are the dusty torus geometry or the warped disk geometry.

\(^5\) We will only describe the most coarse aspects of the classification scheme here—the interested reader is directed to Chapter 1 of the book by Krolik \[103\] for further details.
In the absence of strongly beamed jet emissions, the overall electromagnetic spectrum of type-1 radio-quiet and radio-loud AGN are qualitatively similar. There is often a distinct bump in the spectrum at optical/UV wavelengths, generically referred to as the “big blue bump”, that is thought to correspond to thermal radiation from the optically thick portions of the accretion flow. Furthermore, a power-law spectral component extends from the big blue bump up to hard X-ray energies. As we will discuss in Section 3.3, we believe that this component arises from inverse Compton scattering.
of optical/UV photons in a very hot corona associated with the accretion flow. A high-energy exponential cutoff at $kT \sim 100–200$ keV, corresponding to the temperature of the coronal plasma, has been observed directly in some Seyfert-1 galaxies.

3. Accretion disks and disk atmospheres

3.1. Introduction to accretion and accretion disks

As we have discussed in Sections 2.2 and 2.3, many of the observable effects from black hole systems arise due to the accretion of gas. At its most basic level, gravitational potential energy can be released from the accreting matter and produce electromagnetic radiation or power bi-polar outflows or jets. More exotic is the possibility that accreting matter may produce and support magnetic fields which interact with the rotating black hole thereby facilitating the extraction of its spin energy via, for example, the process elucidated by Roger Blandford and Roman Znajek [107].

In any realistic scenario, matter will be accreting from a substantial distance and will possess significant angular momentum. In this case, we expect the accreting gas to form a rotationally flattened structure orbiting the black hole, an accretion disk. A particular element of gas in the accretion disk will then lose angular momentum through a process which is loosely referred to as viscosity, thereby allowing it to flow slowly inwards towards the black hole and eventually accrete. While the nature of this angular momentum transport was mysterious for some time, we now believe it to be due to magneto-hydrodynamical (MHD) turbulence driven by a well-studied instability, the magnetorotational instability (MRI). In this section, we shall present the “standard” black hole accretion disk model and discuss these MHD phenomena. We then proceed to discuss disk atmospheres and the X-ray spectral signatures that result from these surface layers. Readers who are already familiar with accretion disk physics and X-ray reflection may wish to skip to Section 4.

An excellent review of disk hydrodynamics and MHD, with emphasis on turbulent transport within astrophysical disks, is given by Steven Balbus and John Hawley [108].

Before launching into a detailed discussion of accretion disks, we must introduce the concept of the Eddington limit. Consider spherically symmetric accretion onto an object of mass $M$, producing a radiative luminosity of $L$. If the accreting material is totally ionized, its opacity will be dominated by electron-scattering. If $L$ exceeds a critical luminosity (the Eddington luminosity) given by

$$L_{\text{Edd}} = \frac{4\pi c GM \mu_e}{\sigma_T},$$

where $\mu_e$ is the mass per electron and $\sigma_T$ is the electron scattering cross section, then the outward pressure of radiation will exceed the inward force of gravity making prolonged accretion an impossibility. While the Eddington luminosity strictly applies to spherically symmetric systems, it also probably sets a practical limit for the luminosity of disk-like accretion flows (although there are recent suggestions that MHD processes may allow the limit to be somewhat violated [109]).

3.2. The structure of accretion disks

We begin by discussing the structure and dynamics of the so-called standard model of radiatively efficient accretion disks. We shall focus on such disks since they turn out to be the most relevant for
iron Kα studies of black hole systems. Our discussion will be brief and aims at providing a review of the salient points required for the observationally oriented discussion later. The interested reader is referred to the articles cited near the beginning of each segment of this discussion.

3.2.1. Non-relativistic radiative accretion disks and the $\alpha$-prescription

Initially, we shall discuss non-relativistic models of accretion disks. Here, we follow the presentations and arguments presented in [110–112]. Throughout this discussion, we assume cylindrical polar coordinates $(r, \phi, z)$.

We consider a rotating mass of gas around a point mass $M$, which represents the black hole. We assume that the gas can radiate efficiently and, hence, that it maintains a temperature well below the virial temperature of the system. Furthermore, we assume that an angular momentum transport mechanism operates and can be characterized by a kinematic viscosity $\nu$, but we suppose that the time scale for redistributing angular momentum is long compared with either the orbital or radiative time scales. Under these assumptions, the gas should form a geometrically thin disk (with vertical scale height $h$) in which fluid elements are orbiting in almost circular paths with angular velocity $\Omega(r)$. A small radial velocity $v_r(r,t)$ corresponds to the actual accretion flow.

If we denote the surface mass density of the disk as $\Sigma(r,t)$, the conservation of mass and angular momentum can be written as

$$ r \frac{\partial \Sigma}{\partial t} + \frac{\partial}{\partial r} (r \Sigma v_r) = 0 , \quad (9) $$

and

$$ r \frac{\partial}{\partial t} (\Sigma r^2 \Omega) + \frac{\partial}{\partial r} (r^3 \Sigma v_r \Omega) = \frac{1}{2\pi} \frac{\partial \mathcal{G}}{\partial r} , \quad (10) $$

where

$$ \mathcal{G}(r,t) = 2\pi r^3 v \frac{\partial \Omega}{\partial r} \quad (11) $$

is the torque exerted on the disk interior to radius $r$ by the flow outside of this radius. These equations can be combined, using the fact that in Newtonian gravity $\Omega \propto r^{-3/2}$, to obtain

$$ \frac{\partial \Sigma}{\partial t} = \frac{3}{r} \frac{\partial}{\partial r} \left[ r^{1/2} \frac{\partial}{\partial r} (v \Sigma r^{1/2}) \right] . \quad (12) $$

Noting that $v$ may be a function of the local variables in the disk, this has the form of a non-linear diffusion equation governing the behavior of $\Sigma(r,t)$ given some initial state. Given a solution for $\Sigma(r,t)$, the radial velocity is

$$ v_r = -\frac{3}{\Sigma r^{1/2}} \frac{\partial}{\partial r} (v \Sigma r^{1/2}) . \quad (13) $$

We need to know the behavior of $\nu$ in order to close these equations and permit a full determination of the radial structure of the accretion disk. All of the complexities currently ignored (such as the $z$-structure of the accretion disk, and the detailed microscopic physics) enter into the problem via $\nu$. However, in the case of time-independent accretion disks (i.e., $\partial/\partial t = 0$), it is possible to determine
some of the most interesting observables independent of our ignorance of $v$. In particular, it is readily shown (e.g., [112, Chapter 5]) that the viscous dissipation per unit disk face area is

$$D(r) = \frac{3G\dot{M}M}{8\pi r^3} \left[ 1 - \left( \frac{r_{\text{in}}}{r} \right)^{1/2} \right],$$  \hspace{1cm} (14)

where $\dot{M} = 2\pi r \Sigma(-v_r)$ is the rate at which mass is flowing inwards through the disk, and $r_{\text{in}}$ is some inner boundary to the disk where the torque $\mathcal{G}$ goes to zero. One peculiarity of viscous accretion flows becomes apparent upon an examination of this last expression—for $r \gg r_{\text{in}}$, the rate that energy is lost from the disk is three times the local rate of change of binding energy. That is to say that for a ring of the disk located at $r \gg r_{\text{in}}$, two thirds of the dissipation is attributable to energy that is ‘viscously’ transported from the interior disk, while the remaining one-third is attributable to the local rate of change of binding energy. Of course, to compensate for this effect, the innermost regions of the disk radiate less than the local rate of change of binding energy. Integrating over the whole disk, the total luminosity of the disk is $L_{\text{disk}} = GM\dot{M}/2r_{\text{in}}$. An equal amount of energy remains in the form of kinetic energy of the accreting material as it flows through $r = r_{\text{in}}$. If there is a (slowly rotating) star at the center of the disk (as opposed to a black hole), this remaining energy can be radiated in the disk-star boundary layer.

If we suppose that the viscously dissipated energy is radiated as a black body spectrum, energy conservation ($\sigma_{\text{SB}} T^4 = D(r)$, where $\sigma_{\text{SB}}$ is the Stephan–Boltzman constant) gives the temperature of the disk surface as a function of radius,

$$T(r) = \left( \frac{3G\dot{M}M}{8\pi\sigma_{\text{SB}} r^3} \left[ 1 - \left( \frac{r_{\text{in}}}{r} \right)^{1/2} \right] \right)^{1/4}. \hspace{1cm} (15)$$

For a mass accretion rate that scales with mass (i.e., for a fixed ratio between the source luminosity and the Eddington limit) and a scaled radius $r/M$, the temperature of the disk scales weakly with the black hole mass $T \propto M^{-1/4}$. The accretion disks around supermassive black holes generally possess a characteristic disk temperature of $T \sim 10^5–10^6$ K, making them optical and ultraviolet sources. On the other hand, accretion disks around stellar mass black holes are appreciably hotter with $T \sim 10^7$ K, making them sources of thermal soft X-rays.

For a thin disk, there are no appreciable motions or accelerations in the $z$-direction. Hence the $z$-structure (i.e., the vertical structure) of the disk is determined by a hydrostatic equilibrium condition,

$$\frac{1}{\rho} \frac{\partial p}{\partial z} = \frac{\partial}{\partial z} \left[ \frac{GM}{(r^2 + z^2)^{1/2}} \right], \hspace{1cm} (16)$$

where $p$ is the pressure. Noting that the sound speed is $c_s \sim \sqrt{p/\rho}$, the hydrostatic equilibrium expression can be used to see that the vertical scale height $h$ is,

$$h \sim \frac{c_s r}{v_K}, \hspace{1cm} (17)$$

where $v_K \equiv (GM/r)^{1/2}$ is the local Keplerian velocity. Thus, an accretion disk is geometrically thin ($h/r \ll 1$) if the Keplerian velocity is highly supersonic.

In order to study the detailed physical structure of accretion disks, or any aspect of their time-variability (including the stability of accretion disks), we require a knowledge of the viscosity. In an extremely successful application of dimensional analysis, Shakura and Sunyaev [110] parameterized
the viscosity as \( \nu = \alpha c_s h \), where \( c_s \) is the sound speed in the disk, \( h \) is the vertical scale height and \( \alpha \) is a dimensionless parameter which, on general hydrodynamical grounds, is argued to be less than unity. Furthermore, they made the supposition that \( \alpha \) is a constant for a given accretion disk. The resulting family of disk models are referred to as \( \alpha \)-models (or Shakura–Sunyaev models). Given the \( \alpha \)-prescription, we can solve for the detailed radial disk structure (the results of this exercise are given in [110] and nicely reviewed by [112]). In addition, the assumptions underpinning the basic equations of the radiatively efficient thin accretion disk can now be checked post-priori. From Eq. (13) it can be seen that the radial velocity is given by

\[
\frac{v_r}{R} \sim \frac{\alpha c_s h}{r} \approx c_s .
\]

(18)

Thus, the radial inflow is very subsonic. Using the radial force equation, this also implies that the orbital velocity is very close to Keplerian. Furthermore, we can relate the various fundamental time scales governing the disk. Let the dynamical time scale be defined as \( t_{\text{dyn}} = \Omega^{-1} \). It can be shown that the time scale on which the dissipated energy is radiated from the disk, the thermal time scale, is given by \( t_{\text{th}} \approx t_{\text{dyn}}/\alpha > t_{\text{dyn}} \). Finally the time scale on which angular momentum is redistributed, the viscous time scale, is given by \( t_{\text{visc}} \sim \frac{r^2}{\nu} \sim (r/h)^2 t_{\text{th}} > t_{\text{th}} > t_{\text{dyn}} \). The ordering of these time scales, together with the fact that the disk material orbits the black hole in almost Keplerian orbits, justifies the assumptions listed at the beginning of this section.

While the \( \alpha \)-prescription is extremely useful, allowing us to construct simple analytic models of accretion disks, it is important to realize the severe limitations of this approach. As will be discussed in Section 3.2.3, we now believe that MHD turbulence is responsible for the angular momentum transport that drives accretion. While the angular momentum transport resulting from MHD turbulence can be approximately parameterized in terms of an “effective \( \alpha \)”, there are circumstances where it is incorrect to treat the accretion flow as an unmagnetized gas with an anomalously high kinematic viscosity [108,113]. Even when used within its domain of validity, the actual value of \( \alpha \) applicable to real systems is very unclear with estimates covering full range from essentially zero to unity (see discussion in [112]).

3.2.2. Relativistic disks

Of course, the Shakura–Sunyaev disk models presented above are non-relativistic and, hence, can only be viewed as a crude approximation to real black hole accretion disks. The formal relativistic generalization of the steady-state Shakura–Sunyaev disk model to a relativistic accretion disk in a Kerr metric was made in 1974 by Kip Thorne, Don Page and Igor Novikov [114–116], whereas time-dependent relativistic disks were first treated by Douglas Eardley and Alan Lightman [117]. We shall refer to these relativistic, geometrically thin and radiatively efficient disks as the “standard model” for black hole accretion disks.

Several new features to the model arise once the fully relativistic situation is considered. Most importantly, there is a natural choice of inner boundary. In the standard model of a geometrically thin relativistic accretion disk, the torque is assumed to vanish at the radius of marginal stability, \( r = r_{\text{ms}} \); we shall refer to this standard assumption as the zero-torque boundary condition (ZTBC). It is envisaged that, within \( r = r_{\text{ms}} \), the material plunges into the black hole conserving its energy and angular momentum. When the ZTBC is employed, the viscous dissipation per unit proper disk
The face area per unit proper time is \[115\]

\[D_{\text{PT}}(r; a) = \frac{\dot{M}}{4\pi r F}, \tag{19}\]

where we have defined the function \(F\)

\[F = \frac{3}{2M} \frac{1}{x^2(x^3 - 3x + 2a^*)} \left[ x - x_0 - \frac{3}{2} a^* \ln \left( \frac{x}{x_0} \right) \right.\]

\[\left. - \frac{3(x_1 - a^*)^2}{x_1(x_1 - x_2)(x_1 - x_3)} \ln \left( \frac{x - x_1}{x_0 - x_1} \right) \right.\]

\[\left. - \frac{3(x_2 - a^*)^2}{x_2(x_2 - x_1)(x_2 - x_3)} \ln \left( \frac{x - x_2}{x_0 - x_2} \right) \right.\]

\[\left. - \frac{3(x_3 - a^*)^2}{x_3(x_3 - x_1)(x_3 - x_2)} \ln \left( \frac{x - x_3}{x_0 - x_3} \right) \right], \tag{20}\]

with,

\[x = \sqrt{r/M}, \tag{21}\]

\[x_0 = \sqrt{r_{\text{ms}}/M}, \tag{22}\]

\[x_1 = 2 \cos \left( \frac{1}{3} \cos^{-1} a^* - \frac{\pi}{3} \right), \tag{23}\]

\[x_2 = 2 \cos \left( \frac{1}{3} \cos^{-1} a^* + \frac{\pi}{3} \right), \tag{24}\]

\[x_3 = -2 \cos \left( \frac{1}{3} \cos^{-1} a^* \right). \tag{25}\]

Here, we have defined the dimensionless spin parameter to be \(a^* = a/M\). This is the relativistic generalization of Eq. (14). As we shall discuss in Section 3.2.4, there are likely situations in which the ZTBC does not apply and that the region in which angular momentum and energy transport occurs penetrates significantly within \(r = r_{\text{ms}}\). Nevertheless, the “standard model” is still a useful reference point against which more sophisticated disk models can be judged.

The properties of the inner edge of the disk are important in determining the efficiency of the accretion disk. We define the efficiency of the disk, \(\eta\), by \(L = \eta \dot{M} c^2\), where \(\dot{M}\) is the mass accretion rate onto the black hole and \(L\) is the total power output of the accretion disk (summing over all forms including radiation and the kinetic energy of any outflows). The assumptions of the standard model (and, in particular, the ZTBC), fixes the efficiency \(\eta\) to be the change in specific binding energy as a fluid element moves from infinity down to \(r = r_{\text{ms}}\). For a non-rotating black hole, the radius of marginal stability is relatively distant \((r_{\text{ms}} = 6M)\) and the radiative efficiency is only \(\eta \approx 0.06\). On the other hand, a standard disk around a Kerr black hole with \(a^* \equiv a/M = 1\) has an efficiency of \(\eta \approx 0.42\). For a realistic rapidly spinning black hole, with \(a^* \approx 0.998\), the efficiency is \(\eta \approx 0.30\).
3.2.3. The magneto-rotational instability and MHD turbulence

Until the early 1990s, there was no clear candidate for the actual angular momentum transport mechanism in black hole (or, indeed, neutron star, white dwarf or protostellar) accretion disks—normal atomic/molecular viscosity turns out to be orders of magnitude too small to drive the accretion that was presumed to cause the prodigious radiative powers observed from such sources. Detailed theoretical studies of accretion disks were plagued by our lack of understanding of the basic angular momentum transport mechanism. However, in 1991, Steve Balbus and John Hawley [118] showed that the presence of a (arbitrarily) weak magnetic field in the accretion disk material will render it susceptible to a powerful local MHD shear instability. Assuming that the disk is in the MHD limit, they showed that the only criterion for linear instability is that the angular velocity of the disk decreases outwards, \( \frac{d\Omega^2}{dr} < 0 \); this condition is easily satisfied for any accretion disk around a compact object. It is now widely accepted that this instability, known as the magneto-rotational instability (MRI), lies at the heart of angular momentum transport in the inner regions of accretion disks around compact objects (for a detailed review of accretion disk MHD, see [108]).

A full exploration of the non-linear evolution of the MRI requires numerical MHD simulations. The initial era of simulations [119–121] examined the dynamics in a local patch of the disk, with an extent much smaller than the radius of the disk. These simulations confirmed the linear MRI analysis, and allows the evolution to be followed into the non-linear regime. It is found that the MRI enters the non-linear regime (some aspects of which have been studied analytically by Goodman and Xu [122]) and MHD turbulence results after a small number of dynamical time scales. The process becomes a self-sustaining dynamo, and the turbulence does not decay away with time. Of course, Cowling’s antidynamo theorem dictates that non-axisymmetric simulations are required in order to produce self-sustained dynamo action. In recent years, simulators have been able to leap beyond these local simulations and tackle complete accretion disks (so-called “global” simulations [123–125]). As in the local simulations, the global models also produce sustained MHD turbulence.

The importance to accretion disks comes from the fact that the MRI has the properties of an exchange instability in which high angular momentum material is exchanged with low angular momentum material, thereby facilitating angular momentum transport. From the simulations, one can extract effective values of the Shakura–Sunyaev \( \alpha \)-parameter. For local simulations, one obtains \( \alpha = 10^{-3}–10^{-2} \) if there is no net vertical magnetic flux threading the shearing box, and \( \alpha = 10^{-2}–10^{-1} \) when there is a significant (but still dynamically weak) net vertical flux [120]. The global simulations (which, to date, have only been performed for the case of zero net vertical flux) show \( \alpha = 10^{-2}–10^{-1} \). There is no inconsistency between these global and local results; large-scale field structures generated within the global models can effectively impose a net magnetic flux on the local level, thereby enhancing the local angular momentum transport.

3.2.4. The role of magnetic fields in the innermost disk

While stating the assumption of zero-torque at the radius of marginal stability, Page and Thorne [115] noted that significant magnetic fields might allow torques to be transmitted some distance within the radius of marginal stability. This idea has been rejuvenated by recent work on MRI driven turbulence.

Significant insight can be gained from analytic considerations, assuming that the inner regions of the accretion disk are time-independent and axisymmetric. In separate treatments by Julian Krolik and Charles Gammie [126–128], it was shown that magnetic fields might be effective at transporting
Fig. 5. Model dissipation profiles for torqued time-independent accretion disks (from [126]). The thick line shows the dissipation profile for a standard non-torqued disk (i.e., one in which the ZTBC applies) around a near-extremal Kerr black hole with spin parameter $a = 0.998$. The thin solid lines show the torque-induced dissipation component $D_{\text{tor}}$ for spin parameters of (from left to right) $a = 0.998, 0.99, 0.9, 0.5, 0.3,$ and $0$. Note how the torque-induced dissipation profile becomes more centrally concentrated as the spin parameter is increased. The dashed lines show the torque-induced component of the emitted flux, including the effects of returning radiation.

Accompanying this angular momentum transport, we would expect either local dissipation/radiation of the liberated binding energy, or energy transport into the radiatively efficient portions of the disk. Either way, the principal consequence is to enhance the efficiency $\eta$ of the disk above the value predicted by the Page and Thorne models. The simplest analytic approach (followed by Agol and Krolik [126]) is to model this effect as a finite torque applied at the radius of marginal stability, i.e., an explicit violation of the ZTBC. In addition to modeling the magnetic connection between the accretion disk and the plunging region, this torque can also model any magnetic connection between the accretion disk and the rotating black hole itself [129]. This torque performs work on the accretion disk, thereby adding to the power ultimately dissipated in the disk and increasing the disk efficiency.

This salient aspects of this model are illustrated in Fig. 5 which displays the dissipation rate in the disk per unit disk face area as a function of Boyer–Lindquist radius, $D(r)$. The thick line in this plot shows the dissipation law for a standard Page and Thorne accretion disk around a black hole with spin parameter $a = 0.998$ ($D_{\text{PT}}(r; a = 0.998M)$ given by Eq. (19)). Also shown are the additional dissipation profiles resulting from non-zero torques applied at the radius of marginal stability $D_{\text{tor}}(r; a^*)$. It can be shown that [126]

$$D_{\text{tor}}(r; a^*) = \frac{3M\dot{r}\sqrt{\frac{3}{2}} C_{\text{ms}}^{1/2} \Delta \eta}{8\pi r^{7/2} C(r)}$$

where $C(r) = 1 - 3M/r + 2a^*M^{3/2}/r^{3/2}$ and $\Delta \eta$ is the additional disk efficiency induced by the torque. Strictly, this is true if the additional torque does not directly affect the dynamics of the disk. In this case, the total dissipation is just the sum $D(r; a^*) = D_{\text{PT}}(r; a^*) + D_{\text{tor}}(r; a^*)$. The torque-induced component to the dissipation is much more centrally concentrated than the standard disk dissipation...
profile. In fact, it is so centrally concentrated that a significant fraction of the radiative flux emitted from the disk surface will return to the disk surface due to strong light bending effects. This “returning radiation” will be absorbed or scattered from the disk surface and, hence, adds to the local energy budget relevant for determining the observed amount of radiative flux from a given radius. Including this effect, the total emitted flux per unit area of the disk can be written as

$$D_{tot}(r; a^*) = D_{PT}(r; a^*) + D_{tor}(r; a^*) + D_{ret}(r; a^*)$$ \tag{27}

where $D_{ret}(r; a^*)$ is the term accounting for returning radiation. From the point of view of an observer riding on the disk at a large radius, the returning radiation produces a “mirage” of the central disk raised to some height above the disk plane. Thus, at large radii, irradiation due to returning radiation scales as $r^{-3}$ while the dissipated flux dominates at small radii. Thus, as shown in [126], the term $D_{ret}$ is approximately given by

$$D_{ret}(r; a^*) \approx \frac{3M \dot{M} \Delta \eta R_{\infty}(a^*)}{8\pi r^3}$$ \tag{28}

where $R_{\infty}(a^*)$ is well described by the fitting formula given by Agol and Krolik [126]. The function $D_{tot}(r; a^*)$, which includes the effects of returning radiation, is also shown in Fig. 5.

The analytic models of Gammie, Agol and Krolik [28,126] suggest that these torqued accretion disks can become super-efficient ($\eta > 1$). In this case, some fraction of the radiative power must be extracted from the spin of the black hole. This is a particular realization of the Penrose process, with the accreting material in the innermost parts of the plunging region being placed on negative energy orbits by magnetic connections with the accretion flow at larger radii.

Simulations are required to move beyond the restrictive geometry and physics of the analytic models. To date, almost all of the relevant MHD simulations have been inherently non-relativistic and have mocked up the effects of General Relativity via the use of the Paczyński–Wiita [130] pseudo-Newtonian potential

$$\phi = \frac{M}{r - 2M}$$ \tag{29}

where, to recall, we have set $G = c = 1$. Newtonian computations using this modified point-source potential serve as remarkably effective toy-models for examining dynamics in the vicinity of a Schwarzschild black hole. Simulations of MHD accretion disks within this potential have been presented in a number of papers [124,125,131–133]. These works support the hypothesis that energy and angular momentum can be transported out of the plunging region, although the universality of these processes is still the subject of debate [131,132].

We note in brief that, at the time of writing, fully relativistic MHD codes capable of studying the evolution of the MRI and MHD turbulence around rapidly rotating black holes are just starting to become available [134]. Simulations performed with such codes will be extremely important in assessing the true nature of the innermost regions of black hole accretion disks.

### 3.3. Accretion disk coronae

An accretion disk described exactly by the standard model produces a relatively soft, quasi-thermal spectrum (dominated by optical/UV radiation in AGN, and soft X-rays in GBHCs). However, accreting black hole systems (both stellar and supermassive) often exhibit power-law components
to their spectra which extend to hard X-ray energies with exponential rollovers at high energy \((E \sim 300\text{ keV})\). A promising mechanism for producing such a spectrum is unsaturated inverse Compton scattering \([135,136]\). Extensive descriptions of inverse Comptonization and its applications to astrophysical sources can be found in the literature \([137,138]\). Here we provide a brief summary of these descriptions.

Compton scattering is the scattering of a photon by an electron. In the rest frame of the electron, the differential scattering cross section for a photon of energy \(E \ll m_e c^2\) is given by

\[
\frac{d\sigma_T}{d\Omega} = \frac{3}{8\pi} \sigma_T \left(1 + \cos^2 \theta \right),
\]

where

\[
\sigma_T = \frac{8\pi}{3} \left( \frac{e^2}{m_e c^2} \right)^2 = 6.67 \times 10^{-25} \text{ cm}^2
\]

is the Thomson cross section and \(\theta\) is the angle between the photon’s initial and final trajectories. In the initial rest frame of the electron, the photon imparts kinetic energy to the recoiling electron, and hence the photon loses energy,

\[
\Delta E = -\frac{E^2}{m_e c^2} (1 - \cos \theta).
\]

In the laboratory frame, however, the electron can impart energy to the photon, up to \(\Delta E = (\gamma - 1) m_e c^2\), where \(\gamma\) is the electron’s relativistic Lorentz factor.

It can be shown \([137,138]\) that for a non-relativistic thermal distribution of electrons with temperature \(T_e\), the average photon energy change for a single scattering is given by

\[
\langle \Delta E \rangle = (4kT_e - E) \frac{E}{m_e c^2}.
\]

If \(E \gg 4kT_e\), the electron’s kinetic energy is negligible and we recover the angle average of the previous equation for photon energy loss. If, however, \(E \ll 4kT_e\), the photon gains energy, with the fractional energy gain being approximately proportional to \(\gamma^2\). This proportionality can be understood as one factor of \(\gamma\) coming from the boosting of the photon into the initial electron rest frame, and another factor of \(\gamma\) coming from a boosting of the scattered photon back into the lab frame \([137,138]\). This gain of photon energy is referred to as inverse Compton scattering.

So long as \(E \ll 4kT_e\), the photon continues to gain energy after each scattering event with \(\Delta E / E \approx 4kT_e / m_e c^2\). Thus, after \(N\) scattering events, the average final photon energy, \(E_f\), compared to the initial photon energy, \(E_i\), is given by

\[
E_f \approx E_i \exp \left( N \frac{4kT_e}{m_e c^2} \right).
\]

For a medium with an optical depth of \(\tau_{es}\), the average number of scatters is roughly \(\max(\tau_{es}, \tau_{es}^2)\). This leads one to define the Compton \(y\) parameter,

\[
y \equiv \max(\tau_{es}, \tau_{es}^2) (4kT_e / m_e c^2),
\]

such that \(E_f \approx E_i \exp(y)\) \([137]\). For \(y > 1\), the average photon energy increases by an ‘amplification factor’ \(A(y) \approx \exp(y)\). For \(y \gg 1\), the average photon energy reaches the thermal energy of the
electrons. This latter case is referred to as saturated Compton scattering, while the former case is referred to as ‘unsaturated inverse Comptonization’. Unsaturated inverse Comptonization is of most interest in the study of black hole systems.

As discussed above, the ‘standard’ optically thick, geometrically thin disk model produces a maximum temperature that scales with the compact object mass as $M^{-1/4}$, so one naturally expects lower disk temperatures in AGN, as opposed to stellar mass black hole systems. However, one can instead define a ‘virial temperature’ which refers to the average accretion energy per particle. Specifically, the gravitational energy released per particle of mass $m$ scales as $GMm/R$, and therefore scales as $mc^2/[R/(GM/c^2)]$. Thus, in terms of geometrical radii, the virial temperature is independent of compact object mass and furthermore can reach a substantial fraction of the rest mass energy of the accreted particles. Electron virial temperatures of tens to hundreds of keV can be readily achieved in the innermost regions of compact object systems. (The proton virial temperature is a factor of $\approx 2000$ greater; therefore, it is conceivable that electrons can achieve even higher temperatures depending upon the temperature of the protons and the coupling between the electrons and protons.)

It is such a distribution of low-to-moderate optical depth, high temperature electrons that is suspected of creating the observed power law spectra in black hole systems. Although multiple photon scatterings in such a corona become exponentially unlikely, multiple scatterings lead to exponential photon energy gain. The two effects balance, to some degree, and produce a power law spectrum. Simple estimates [137] show that for a given Compton $y$ parameter, the photon index of the resulting power law is approximately given by

$$\Gamma = -\frac{1}{2} + \sqrt{\frac{9}{4} + 4y},$$

i.e., $\Gamma \approx 2$ for $y = 1$. Of course, the exact function $\Gamma(y)$ is dependent upon geometry and other assumptions. A power law form only holds for photon energies somewhat less than the electron thermal energy. As photons approach the electron thermal energy, they no longer gain energy from scattering, and a sharp rollover is expected in the spectrum. Thus, in inverse Comptonization models, the observed high energy spectral cutoff yields information about the temperature of the underlying electron distribution.

There are two principle uncertainties in applying coronal models to observed black hole systems, and both are substantial. First, the mechanism for heating the electrons to near virial temperatures is currently unknown. Current hypotheses invoke magnetic processes, perhaps akin to solar flares on the Sun or heating of the solar corona. Contrary to solar models, however, black hole coronae may be energetically dominant. Second, the geometry of the corona is also completely unknown. Thus, models have taken to hypothesizing specific geometries (see Fig. 6), and parameterizing the energy input into the corona.

An early suggestion [141] for the geometry of coronal systems was that the accretion disk consisted of two zones—a cool, outer disk, and a hot inner, disk. It was envisioned that energy release was actively occurring within the disk itself, but that a “two-temperature” disk was being formed within

---

6 These particular expressions come from a solution of the Kompaneets equation, which under certain approximations describes the diffusion in phase space of photons undergoing Comptonization. Excellent discussions of the Kompaneets equation can be found in the review by Pozdnyakov et al. [138], the book by Rybicki and Lightman [137], and Chapter 24 of the book by Peebles [139].
the inner regions. Ions would reach temperatures of $\approx 10^{11-12}$ K, while electrons would achieve temperatures of only $\approx 10^9$ K. Hard X-ray emission would be due to Comptonization of soft photons from the outer disk by the hot electrons in the inner disk. It was later shown [142,143] that this specific disk model was thermally unstable; however, the basic geometry, which we will refer to as the ‘sphere+disk’ geometry, has been proposed in a variety of forms since that time. The attraction of such a geometry is that the hot, inner corona sees only a fraction of the soft flux from the outer, cool disk, and is therefore not strongly Compton cooled. This allows the corona to remain very hot and to produce hard spectra [144].

The above disk models postulate a radial separation between the geometrically thin, optically thick disk and the geometrically thick, small optical depth corona. Other sets of coronal models exist, however, wherein both ‘disk’ and ‘corona’ exist at the same radii, but the accretion energy is postulated to be dissipated predominantly within the corona [145–148]. It was thought early on that a promising source for the necessary coronal energy would be dissipation of magnetic flux tubes as they buoyantly rose above the disk into the corona [149]. The recent studies of MHD instabilities in accretion flows have revived interest in such mechanisms [118,108,150].

The simplest such model is one where the corona sandwiches the accretion disk [151–154]. For a uniform and plane parallel corona and disk, the basic energetics of this geometry can be obtained from conservation principles. Let us decompose the radiative flux into that from the corona $F_c$ and
that from the optically thick part of the disk $F_d$. The coronal flux is comprised of two parts, an outward component, $F_c^+$, and an inward component, $F_c^-$, which will be intercepted and reprocessed by the accretion disk. The total flux from the disk, $F_d^t$, is comprised of any intrinsic dissipation within the disk, $F_d^i$, plus the flux intercepted from the corona. Thus $F_d^t = F_d^i + F_c^-$. The total flux from the corona, $F_c^+ + F_c^-$, is comprised of any intrinsic dissipation within the corona, $F_c^i$, plus the total outward flux of the disk. By definition, this is set equal to the outward flux from the disk multiplied by the Compton amplification factor, $A_c$. Specifically,

$$F_c^+ + F_c^- = A_c F_d^t \quad (37)$$

After multiple scatterings within the corona, the radiation field becomes nearly isotropic, therefore $F_c^+ \approx F_c^- + \Theta F_d^t$, where $0 \leq \Theta \leq 1$ is a measure of the anisotropy of the coronal flux. As the optical depth of the corona, $\tau_{es} \rightarrow 0$, $\Theta \rightarrow 0$. If one now assumes that a fraction, $f$, of the total (local) accretion energy is dissipated in the corona, while $1 - f$ is dissipated within the disk, one obtains $F_d^i / F_c^i = (1 - f) / f$, and

$$A_c = \frac{1 + \Theta f/2}{1 - f/2} \quad (38)$$

The above highlights two important considerations of this slab geometry. First, even under the most extreme conditions of $\Theta \approx 1$ and $f \approx 1$ (i.e., all accretion energy is dissipated within the corona), the Compton amplification factor does not exceed 3. This is equivalent to saying that the Compton $y$ parameter does not exceed $\approx 1$. Second, 100% of the soft, seed photons from the disk pass through the corona. Thus, the corona is relatively easily Compton cooled and it is difficult for it to obtain very high temperatures [140,144]. Even if the optical depth of such a corona is made small, pair production within the corona will ensue until the rate of heating is balanced by Compton cooling. With a pure slab geometry, extremely hard spectra (i.e., with very high energy cutoffs) are impossible to achieve [140,155].

There have been several suggestions put forward that allow the above basic geometry to achieve higher coronal temperatures, and thus harder spectra. If the corona is ‘patchy’, such that small coronal regions sit atop the disk (sometimes referred to as a ‘pill box’ geometry), then a large fraction of the coronal flux can be intercepted by the disk, with a relatively small fraction of the (reprocessed) disk flux being intercepted by the corona. Thus, the coronal patches are not very effectively Compton cooled [140]. Another method for varying the amount of soft flux reprocessed by the corona is to invoke relativistic motion within the corona. Downward motion beams the radiation towards the disk and increases the local disk reprocessing, while motion away from the disk leads toward beaming away from the disk, thereby decreasing reprocessing [156,157]. Finally, as we will further discuss below, if the disk surface is highly ionized, then rather than being reprocessed by the disk into soft radiation, the hard radiation will merely reflect back through the corona, without substantially cooling it [158–161].

### 3.4. Radiatively inefficient accretion disks

Here we briefly discuss an important class of accretion disks models that have received much attention in recent years—the radiatively inefficient disks. Such disks, in which the radiative-efficiency
η becomes small, can occur when the accretion rate is either very low or very high. The physics is rather different in these two cases. For low accretion rates, the accreting plasma becomes so tenuous that the electrons and ions may lose thermal contact with each other (assuming that they are principally coupled via Coulomb collisions). The ions are very poor radiators; thus if most of the accretion energy is channeled to the ions rather than the electrons, it will not be radiated and instead remain as thermal energy advected along within the accretion flow [162,163]. In the so-called advection dominated accretion flows (ADAFs), this energy is advected right through the event horizon [164–166]. Such models were in fact first proposed to explain the hard state spectrum of Cyg X-1 [162], but were later applied to AGN [163], and they have been used to model the Galactic Center [167–169].

It was later realized that the situation described by these low luminosity ADAF models was dynamically unlikely—the viscous transport of energy within this flow could readily unbind material further out, possibly leading to a powerful wind [170] or strong convection [171]. These suggestions remain controversial [113] and are the subject of active research [172–174], but in any case, such a flow is likely to be extremely hot (electron temperatures of $\sim 10^9$ K) and optically thin.

Although these models possess the sphere+disk geometry (the inner region is advection dominated while the outer region represents a standard disk), ADAF models postulate that a large fraction of the seed photons for Comptonization come from synchrotron radiation internal to the radiatively inefficient flow [175]. In addition, their low-levels of emission are concentrated towards small radii [164], while their outer radii have been postulated to extend anywhere from $10–10^4 GM/c^2$ [176]. In the context of the disk atmosphere modeling described below, most (but not all) variants of the ADAF model predict weak spectral lines from the innermost regions of the accretion flow.

For very high accretion rates, comparable to that needed to produce the Eddington luminosity, the accretion inflow time scale can become less than the time it takes for radiation to diffuse out of the disk, and hence the photons can become trapped in the accretion flow [177]. The inability of these disks to radiate the gravitational potential energy, together with the viscous transport of energy and strong radiation pressures present, will almost certainly lead to strong outflows. The appearance of such disks is highly uncertain—it is unclear whether an X-ray emitting corona forms, and whether the atmosphere of such a disk is in a state capable of producing X-ray reflection spectral signatures. At the very least, it is highly unlikely that strong line features will emanate from the innermost, highly relativistic regions of the accretion flow.

### 3.5. Disk atmospheres and X-ray reprocessing features

As opposed to the ADAFs described above, standard, radiatively efficient accretion disks are optically thick (i.e., opaque) across all frequencies of interest—thus all of the radiation directly attributed to the accretion disk must be reprocessed through its outer layers. To be slightly more precise, we shall refer to the outer few Thomson depths\(^7\) of the accretion disk surface as the “atmosphere”. It is obviously crucial to study the disk atmosphere if we are to use observations of accretion disks to study the environment of black holes. Due to the scope of this review, we focus on the reprocessing and “reflection” of X-rays by the disk atmosphere.

---
\(^7\) The Thomson depth is the distance over which the electron scattering optical depth is unity.
3.5.1. Simple X-ray reflection models

Initially, we explore a model for the disk atmosphere that is clearly oversimplified but helps elucidate the physics of X-ray reflection. We suppose that the surface of the accretion disk can be modeled as a semi-infinite slab of uniform density gas, irradiated from above by a continuum X-ray spectrum produced in the disk corona via thermal Comptonization. Furthermore, we assume that the hydrogen and helium are fully ionized, but all other elements (collectively referred to as “metals” by astronomers) are neutral. This is a crude approximation to the situation found in a “cold” AGN accretion disk. Now we consider the possible fates of an incident X-ray photon. Firstly, the photon can be Compton scattered by either the free-electrons associated with the ionized hydrogen and helium, or the outer electrons of the other elements. Secondly, the photon can be photoelectrically absorbed by one of the neutral atoms. For this to happen, the photon must possess an energy above the threshold energy for the particular photoelectric transition. The transitions with the largest cross-sections are those associated with the photo-ejection of a K-shell (i.e., \( n = 1 \) shell) electron. Following K-shell photoionization, the resulting ion commonly de-excites in one of two ways, both of which start with an L-shell (\( n = 2 \)) electron dropping into the K-shell. In the first case, the excess energy is radiated as a K\( \alpha \) line photon (i.e., fluorescence). In the second case, the extra energy is carried away via the ejection of a second L-shell electron (i.e., autoionization or the Auger effect). The fluorescent yield of a species gives the probability that the excited ion will de-excite via fluorescence rather than autoionization.

Fig. 7 shows the results of a Monte-Carlo simulation modeling these processes when a power-law X-ray continuum with photon index \( \gamma = 2 \) is incident on a gaseous slab [24,178], assuming the slab has cosmic abundances [179]. At soft X-ray energies, the albedo of the slab is very small due to the photoabsorption by the metals in the slab. However, at hard X-ray energies, this photoabsorption becomes unimportant (i.e., the photoelectric cross section falls to small values) and most of the X-rays incident on the slab are Compton scattered back out of the slab. Associated with the photoionization of metals in the slab, there is a spectrum of fluorescent emission lines. Due to the combination of high fluorescent yield and large cosmic abundance, the most prominent such fluorescence is the K\( \alpha \) line of iron at 6.40 keV. It is interesting to note in Fig. 7 the weak “shoulder” on the low-energy side of the iron–K\( \alpha \) line. This feature corresponds to line photons that have Compton scattered, and hence lost energy due to electron recoil, before escaping the disk. It is often referred to as the “Compton shoulder”.

It is customary to measure the relative strengths of astrophysical emission lines via the use of equivalent widths. The equivalent width of an emission line is the energy (or wavelength) range over which the continuum radiation contains a flux equal to that contained in the emission line. For the case presented in Fig. 7 the equivalent width of the iron line when viewed in the combined direct + reflected spectrum is approximately 180 eV. The other lines are much weaker, with equivalent widths at least an order of magnitude less [180].

---

8 In fact, the K\( \alpha \) emission line is a doublet (comprising of the K\( \alpha_1 \) and K\( \alpha_2 \) lines). Since the splitting between the K\( \alpha_1 \) and K\( \alpha_2 \) lines is only a few eV, it is not relevant for the current discussion of relativistically broadened emission lines. Thus, we will ignore the fine structure of the K\( \alpha \) line for the remainder of this review.
3.5.2. Physical X-ray reflection models

The physical situation with a real disk atmosphere is much more complex. We can consider the atmosphere as a layer of gas that is heated from below by the main body of the accretion disk (which, in turn, is heated by the “viscous” dissipation of gravitational potential energy) and irradiated from above by high-energy radiation from the corona. To determine the structure of the disk atmosphere, we must model the radiative transfer, and solve for the thermal and ionization balance of each fluid element. Furthermore, the atmosphere almost certainly responds dynamically to the heating/irradiation, thereby requiring us to solve the equations of hydrodynamics (or, more generally MHD) at the same time. Finally, we must acknowledge the fact that radiation pressure may well be important in many disks of interest.

This is clearly a very complex problem, and progress to date has been made only by considering rather special cases. The first step towards realistic disk reflection models maintained the assumption of a stationary, time-independent, uniform density structure, and assumed that the atmosphere was in thermal and ionization equilibrium. By solving the radiative transfer problem in this fixed background, the thermal and ionization structure as a function of depth in the atmosphere can be obtained [182–185]. Hence, the resulting X-ray reflection spectrum can be determined. A useful quantity in the discussion of photoionized atmospheres is the ionization parameter,

$$\zeta(r) = \frac{4\pi F_s(r)}{n(r)}$$

(39)
Fig. 8. Theoretical ionized reflection spectra for different ionization parameters, as computed using the constant density models of David Ballantyne and collaborators [181]. From bottom to top, the curves show the reflected X-ray spectrum for ionization parameters of log $\xi = 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5$ and 5.0. The detailed behavior is discussed in the main text.

where $F_x(r)$ is the X-ray flux received per unit area of the disk at a radius $r$, and $n(r)$ is the comoving electron number density: it measures the ratio of the photoionization rate (which is proportional to $n$) to the recombination rate (proportional to $n^2$). The constant density ionization models suggest that the value of $\xi$ delineates four regimes of behavior (see Fig. 8).

1. $\xi < 100$ ergs cm s$^{-1}$: This is the “neutral reflection” regime since the X-ray reflection resembles that from cold gas containing neutral metals (as in Fig. 7). There is a cold iron line at 6.4 keV, and the Compton backscattered continuum only weakly contributes to the observed spectrum at this energy. There is a weak iron K-shell edge at 7.1 keV.

2. $100$ ergs cm s$^{-1} < \xi < 500$ ergs cm s$^{-1}$: In this “intermediate ionization” regime, the iron is in the form of Fe$^{xvii}$–Fe$^{xxiii}$ and there is a vacancy in the L-shell ($n = 2$) of the ion. These ions can resonantly absorb the corresponding K$\alpha$ line photons. Following each absorption, there can be another fluorescent emission (with the consequence that the photon has effectively been scattered) or de-excitation via the Auger effect. A given iron K$\alpha$ photon can be trapped in the slab via the resonant scattering until it is terminated by a Auger event. Thus, only a few line photons can escape the disk leading to a very weak iron line. The reduced opacity below the iron edge due to ionization of the lower-Z elements leads to a moderate iron absorption edge.

3. $500$ ergs cm s$^{-1} < \xi < 5000$ ergs cm s$^{-1}$: In this “high ionization” regime, the ions are too highly ionized to permit the Auger effect (which requires at least 2 L-shell electrons prior to photoionization). While the line photons are still subject to resonant scattering, the lack of a
destruction mechanism ensures that they can escape the disk. The result is Kα iron emission of Fe\textsc{xv} and Fe\textsc{xvi} at 6.67 and 6.97 keV, respectively. Due to the high levels of ionization and subsequent lack of photoelectric opacity characterizing the low-Z elements, the Compton backscattered continuum is a significant contributor to the observed emission at 6 keV. Thus, there is a large iron absorption edge in the observed spectrum.

(4) $\xi > 5000$ ergs cm s$^{-1}$: In this “fully ionized” regime, the disk is too highly ionized to produce any atomic signatures. There is no iron emission line or edge.

More recent variants of these fixed density models have improved the treatment of Compton scattering in the upper atmosphere [181], and have assumed a Gaussian density profile $\rho(z) = \rho(0) \exp[-(z/z_0)^2]$ [158]. The qualitative behavior is, however, unchanged.

Yet more realistic models require us to relax the assumption of a fixed density structure. The next simplest, yet physical, assumption is to suppose that the disk atmosphere is in hydrostatic (or, more generally, hydromagnetic) equilibrium. The density, temperature and pressure must then adjust so as to achieve hydrostatic equilibrium. When such hydrostatic, irradiated disk-atmosphere models are constructed, a qualitatively new phenomenon occurs—the thermal ionization instability (TII) [186].

The origin of the TII is as follows. The equations defining local thermal and ionization equilibrium for a specified density admit a unique solution. On the other hand, if the pressure is specified, there are certain values of pressure for which there exist more than one equilibrium thermal and ionization state (with the intermediate temperature equilibria often proving to be thermally unstable). This is the thermal ionization instability. The primary consequence of the TII is that, over certain ranges of pressure, a multiphase medium can form with comparatively cool regions co-existing in pressure equilibrium with very hot regions.

Now consider moving to greater and greater depths in a hydrostatic disk atmosphere. As the pressure increases monotonically, one will encounter layers of the atmosphere in which the material is subject to the TII and, hence, multiple temperature and ionization structures are allowed. The exact solution actually realized by a particular fluid element depends upon the history of that element, but there are general arguments suggesting that it will usually exhibit the highest temperature solution [186,159]. As one transits from regions (i.e., pressures) where the TII operates to regions where it does not, there will be abrupt changes in the temperature, density and ionization state of the material. In practice, a highly ionized and well-defined hot “skin” will develop, overlying the colder parts of the disk that are more relevant for producing interesting X-ray reflection features. If this ionized skin has a low optical depth, it will be largely irrelevant to the observed reflection spectrum. However, if the electron scattering optical depth of the skin approaches unity, it will cause a smearing of any sharp features in the reflection spectrum (due to Compton recoil) as well as an apparent dilution of the observed reflection features (since it will Compton scatter X-ray continuum photons back towards the observer without any appreciable spectral change). For strongly irradiated disks, the hot skin will be optically thick and act as a “Compton mirror”, simply reflecting back the irradiating spectrum (with slight modifications caused by electron recoil which only become important for very hard X-rays; see for example [187]). If the irradiating spectrum is sufficiently hard ($I > 2$), the Compton mirror can form before the disk is capable of producing observable H- or He- like Kα iron line emission.

When faced with real data, how is one to make progress given the complexities involved in modeling the reflection spectrum? A major simplification, pointed out by David Ballantyne and
collaborators [181], is that self-consistent reflection spectra can be approximately mapped to more manageable constant-density models, albeit with different inferred reflection fractions (due to dilution by an ionized skin) and ionization parameters. Thus, for all practical purposes, X-ray observers to date have used the constant density models when attempting to fit real data.

4. Emission lines from accretion disks

Having discussed accretion disk theory, the properties of accretion disk atmospheres, and X-ray reflection, we move to the main thrust of our discussion—the use of emission lines from the disk surface as a probe of the inner regions of black hole accretion disks. In this section, we treat the basic physics underlying the energy (frequency) profiles of disk emission lines. We closely follow the discussion presented in [188].

4.1. Formalisms for computing line profiles

Since the spectrum is emitted from an accretion disk close to the black hole, an observer at infinity will see the reflected X-ray spectrum of Fig. 7 distorted by relativistic effects, namely the Doppler effect, relativistic aberration, and gravitational light-bending/red-shift.

The specific flux $F_{v_0}$ at frequency $v_0$ as seen by an observer at infinity is defined as the (weighted) sum of the observed specific intensities $I_{v_0}$ from all parts of the accretion disk,

$$F_{v_0} = \int_{\Omega} I_{v_0} \cos \theta \, d\Omega,$$

where $\Omega$ is the solid angle subtended by the accretion disk as seen from the observer and $\theta$ is the angle between the direction to the disk and the direction of the observed photon. Since the black hole is assumed to be very far away from the observer, we can safely set $\cos \theta = 1$. Thus, we have to compute the specific intensity $I_{v_0}$ at infinity from the spectrum emitted on the surface of the accretion disk, $I_{v_e}$. In an axisymmetric accretion disk, $I_{v_e}$ is a function of the radial distance of the point of emission from the black hole, $r_e$, and of the inclination angle, $i_e$, of the emitted photon, measured with respect to the normal of the accretion disk.

The emitted and observed frequencies differ due to Doppler boosting and gravitational red-shift. We define the frequency shift factor $g$, relating the observed frequency $v_0$ and the emitted frequency $v_e$,

$$g = \frac{v_0}{v_e} = \frac{1}{1 + z},$$

where $z$ is the redshift of the photon. According to Liouville’s theorem, the phase-space density of photons, proportional to $I(v)/v^3$, is constant along the photon path. It is therefore possible to express Eq. (40) in terms of the emitted specific flux on the accretion disk:

$$F_{v_0} = \int_{\Omega} \frac{I_{v_0}}{v_0^3} \, d\Omega = \int_{\Omega} \frac{I_{v_e}}{v_e^3} \, d\Omega = \int_{\Omega} g^3 I_{v_e}(r_e, i_e) \, d\Omega.$$

In other words, the computation of the emerging spectrum breaks down to the computation of $g$. In the weak field limit, when $r \gg 3$, and in the Schwarzschild metric, $g$ and therefore the line profile
emitted by the accretion disk, can be evaluated analytically. Profiles computed this way have been presented, e.g., by Andy Fabian and collaborators [189] for the Schwarzschild case, and by Chen and Halpern [190] in the weak field limit. In general, however, the computation has to be done in the strong-field Kerr metric.

The brute-force approach to the computation of $g$ in the Kerr metric is the direct integration of the trajectory of the photon in the Kerr metric [191,192]. This approach allows the computation of exact line profiles even in the case of very complicated geometries, such as geometrically thick or warped accretion disks [193] or disks with non-axisymmetric (e.g., spiral wave) structures [194,195]. The second way to compute the observed flux was first used by Cunningham in 1975 [196] who noted that the observed flux from can be expressed as

$$F_{\nu_0} = \int T(i_e, r_e, g) I_{\nu_0}(r_e, i_e) \, dg \, r_e \, dr_e,$$

where the integration is carried out over all possible values of $g$ and over the whole surface of the accretion disk. This form is well suited for fast numerical evaluation. All relativistic effects are contained in the transfer-function $T$ which can be evaluated once for a given black hole spin parameter. We refer to [196–198], and the references in these works for the technical details.

4.2. The emerging line profile

In Fig. 9–11 we illustrate the relativistic effects on the emerging line profile. All line profiles have been computed with the code of Roland Speith [198]. In our computations we assumed a
Fig. 10. Iron line profiles from relativistic accretion disk models as a function of emissivity index, \( \beta \). The black hole’s spin parameter is assumed to be \( a = 0.5 \) and, again, it is assumed that the disk extends to the radius of marginal stability.

Fig. 11. Iron line profiles from relativistic accretion disk models as a function of black hole spin parameter. The disk is assumed to extend to the radius of marginal stability with an emissivity index of \( \beta = 3 \).

geometrically thin, but optically thick Keplerian accretion disk to be the source of the line radiation. The local emissivity of the line on the disk was parameterized as

\[
I_{\nu}(r_e, i_e) \propto f(i_e) r_e^{-\beta},
\]

(44)
where $f(i_b)$ is a function that can be chosen to model various limb-darkening scenarios. This parameterization is sufficient for most practical work [199]. For optically thick material, $f = \text{constant}$.

In all line profiles shown here, the emitting part of the disk is assumed to extend from the radius of marginal stability to an outer radius of $r_{\text{out}} = 50 \text{ GM/c}^2$. For $\beta > 2$, the location of this outer radius is unimportant to the line profile since the line emission is dominated by the inner regions. However, for $\beta < 2$, this choice of outer radius is of some importance since the bulk of the line emission comes from these outer regions. In the latter case, the separation between the red and blue peaks of the iron profile diminishes as the outer radius is increased (due to the diminishing line of sight velocity difference between the approaching and receding sides parts of the disk).

Common to all line profiles is a characteristic double-horned shape (Fig. 9). To see the origin of this, consider the emission line profile that arises from a narrow annulus of the disk $r \rightarrow r + dr$. By assumption, the emitting matter in this annulus has a single orbital speed but, clearly, the observed Doppler shift from different parts of the annulus will differ due to the changing angle between the observer’s line-of-sight and the velocity vector of the matter (which is undergoing near-circular motion). Line emitting material on the portion of the orbit moving away from the observer will contribute lower-frequency emission than material on the approaching portion of the orbit. Considering this simple case, it can be seen that the line flux in a given frequency range $dv$ will be at a maximum when the Doppler shifts are at an extremum, i.e., for the minimum and maximum Doppler shifts from the annulus. These “red” and “blue” peaks in the line profile of the ring-element correspond to extrema in the “pathlength” through material at a given line-of-sight velocity—these extrema are found on the receding and approaching parts of the annulus. Once these horns are formed in the line profile, special relativistic aberration (i.e., the beaming of emission in the direction of motion) strongly enhances the blue horn over the red. This effect is clearly visible in Fig. 9.

In addition to the line-of-sight Doppler boosting, the line emission is also redshifted due to the gravitational potential and the transverse Doppler effect (i.e., special relativistic time dilation). The influence of all of these effects on the line profile depends on the observer’s inclination angle $i_0$: For a disk seen almost face on (i.e., $i_0$ close to $0^\circ$), the gravitational/time-dilation redshifts dominate. Even with a face-on disk, integrating the observed emission across all radii including the gravitational/time-dilation effects can still produce a rather broad emission line, unlike the case of a face-on Newtonian disk which would display a very narrow line. With progressively larger inclinations $i_0$, line-of-sight Doppler effects become dominant and the line becomes broader.

The broadest parts of the profile are due to material emitted very close to the black hole, as is evident from Fig. 10, where line profiles for different emissivity coefficients $\beta$ are shown. For large $\beta$, most of the line emission takes place close to the last stable orbit, so that these profiles are the broadest. Note that for values of $\beta < 2$ the red wing of the profile gets weaker until it is almost undetectable.

If one assumes that the line emitting disk extends down to, and is then truncated, at the radius of marginal stability, the strong dependence of $r_{\text{ms}}$ on the black hole spin translates into a strong dependence of line profile on black hole spin. In particular, the red-wing of the iron line can extend to much lower energies in the case of a rapidly rotating black hole as compared with a non-rotating black hole. This is illustrated in Fig. 11. However, as we discuss in more detail in Section 6.1, it

---

9 It is customary in astronomy to call lower energies “red” and higher energies “blue”, even when talking about lines in energy bands other than the optical.
is unclear that the effects of the radius of marginal stability can be modeled so trivially, leading to many subtleties to using these line profiles to probe black hole spin.

5. The X-ray astronomer’s arsenal

5.1. Past and current

Of course, the Earth’s atmosphere is completely opaque to X-rays from space—thus, X-ray astronomy must be conducted from the upper atmosphere or, preferably, space. Early X-ray astronomy utilized detectors on sub-orbital rockets and high-flying balloons. Over the past 30 years, however, X-ray astronomy has been conducted primarily via earth-orbiting satellites, launched by a wide range of nations and international collaborations. Countries participating in these missions have included the United States of America, the United Kingdom, Japan, the European Union, Germany, the Soviet Union and Russia, Italy, the Netherlands, and India.

Below we discuss some of those satellites that have made the greatest contribution to the study of accretion disk features in compact object systems. The list is not meant to be exhaustive or exclusive, but it does provide a fair sampling of the observational capabilities at the disposal of X-ray astronomers in the past, present, and near future. We also discuss some of the future missions being planned by US National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA).

5.2. Past and current X-ray observatories

Table 1 lists the capabilities of the major X-ray observatories relevant to the present discussion. Before the launch of ASCA, the most important observatories for iron line studies were the European’s EXOSAT and Japan’s Ginga satellites. As we will discuss in the following sections, these observatories provided the first clear evidence for iron fluorescence and X-ray reflection in accreting black hole sources [189,200–202]. However, we had to await the major increase in spectral resolution brought about by the launch of ASCA before the relativistic effects could truly be explored.

The fourth Japanese X-ray satellite, initially called Astro-D, was renamed to ASCA shortly after launch in February 1993 ASCA had two major innovations which revolutionized the study of many astrophysical X-ray sources. It was the first observatory to possess (four) focusing X-ray telescopes that worked up to 10 keV (previous focusing X-ray telescopes only operated in the soft X-ray band). But, most importantly, ASCA had X-ray sensitive CCDs placed at the focal plane of two of these telescopes, giving X-ray spectra with an energy resolution $E/\Delta E \approx 50$. This allowed, for the first time, X-ray line widths to be resolved, as well as details of broader features to be discerned. Many of the results of AGN studies discussed below were obtained via ASCA observations. NASA and ESA also had a significant contribution to the ASCA project.

---

10 From Asuka, pronounced ASCA, the name of a region in central Japan, and an era in Japanese history around 600 AD during which Asuka was the capital. It is also an ancient Japanese word meaning “flying bird”.
Table 1
Observatories and instruments that have been important for studies of X-ray reflection from black hole sources

<table>
<thead>
<tr>
<th>Observatory (lifetime)</th>
<th>Instrument</th>
<th>Area (cm²)</th>
<th>Band pass (keV)</th>
<th>Resolution (E/ΔE at 6 keV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXOSAT (ESA) May 1983–April 1986</td>
<td>GS 100</td>
<td>2–20</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Ginga (Japan) February 1987–November 1991</td>
<td>LAC 4000</td>
<td>1.5–37</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>ASCA (Japan+NASA) February 1993–March 2001</td>
<td>GIS 2 × 50 &amp; 1 keV</td>
<td>0.8–11</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>RXTE (NASA) December 1995–present</td>
<td>PCA 6500</td>
<td>2–60</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>BeppoSAX (IT+NL) April 1996–April 2002</td>
<td>LECS 22 &amp; 0.28 keV</td>
<td>0.1–10</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Chandra (NASA) July 1999–present</td>
<td>ACIS 340 &amp; 1 keV</td>
<td>0.2–10</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>XMM-Newton (ESA) December 1999–present</td>
<td>EPIC-MOS 2 × 920 &amp; 1 keV</td>
<td>0.2–12</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

Note that some of these observatories possess other instruments/detectors that we have not listed since they are not of direct relevance to iron line studies. Instrument abbreviations: ACIS: AXAF Charged Coupled Imaging Spectrometer; EPIC: European Photon Imaging Camera; GIS: Gas Imaging Spectrometer; GS: Gas Scintillation Proportional Counter; HETG: High Energy Transmission Grating; HEXTE: High Energy X-ray Timing Experiment; LAC: Large Area Proportional Counter; LECS: Low Energy Concentrator Spectrometer; ME: Medium Energy Proportional Counter; MECS: Medium Energy Concentrator Spectrometer; PCA: Proportional Counter Array; PDS: Phoswich Detection System; SIS: Solid-State Imaging Spectrometer.

The two premier X-ray observatories currently operating are the **Chandra X-ray Observatory** (NASA) and **XMM-Newton** (ESA). Before discussing future planned missions, we briefly describe these two observatories in a little more detail.

5.2.1. The Chandra X-ray observatory

The Advanced X-ray Astrophysics Facility (AXAF) was launched in July 1999 as one of NASA’s “Great Observatories”, and was subsequently renamed Chandra, in honor of Subrahmanyan Chandrasekhar. Its prime characteristic is its extremely high spatial resolution. Its X-ray mirrors (consisting of nested shells of grazing incidence optics) are able to resolve angular scales of only 0.2 arcsec. Even with all of the practicalities of observing, resolutions of 0.5 arcsec can be achieved. The primary detector array (ACIS) comprises of an array of CCDs at the focal plane of the telescope. When operated in “imaging” mode, the X-rays strike these CCDs directly, producing a high-quality image in which the energy of each photon is tagged with moderate energy resolution (E/ΔE ~ 50).
In “spectroscopy” mode, one of a set of (transmission) diffraction gratings is placed in the X-ray beam, allowing high resolution spectra (up to $E/\Delta E \sim 1000$ at soft X-rays; $E/\Delta E \sim 200$ at 6 keV) to be obtained at the expense of one spatial dimension.

5.2.2. XMM-Newton

The X-ray Multi-Mirror Mission (now called XMM-Newton) was launched by the European Space Agency in December 1999. It possesses three grazing incidence X-ray telescopes with CCD detectors (the European Photon Imaging Camera; EPIC) in each focal plane. XMM-Newton is complementary to Chandra in the sense that, while not possessing the high spatial resolution of Chandra (it can achieve spatial resolutions of $\sim 5$ arcsec), it has a significantly larger collecting area thereby allowing high-throughput spectroscopy. Since many studies of X-ray reflection in accreting black holes are signal-to-noise limited, XMM-Newton has become an extremely important observatory for these purposes. Two of the telescopes also have (reflection) diffraction gratings that permanently intercept half of the X-ray beam in two of the telescopes, providing simultaneous high-resolution (up to $E/\Delta E \sim 500$) X-ray spectra in the soft X-ray band (below 2 keV).

5.3. Future missions

5.3.1. Astro-E II

Astro E was supposed to be the fifth in the series of Japanese X-ray satellites. Its overall soft X-ray effective area and mirror design were similar to those of ASCA, except that its mirrors had a factor of two improvement in spatial resolution. In addition to soft X-ray CCD detectors, Astro E had a hard X-ray detector that was sensitive to energies as high as 700 keV. Most importantly, however, Astro E was to utilize the first space-based X-ray micro-calorimeters. Rather than achieve high spectral resolution via dispersion gratings, X-ray photons would deposit heat into calorimeter elements with eV energy resolution, thus achieving $E/\Delta E \approx 1000$ or more at 6 keV. The calorimeters required cryogenic cooling, limiting their useful lifetime to approximately 2 years. Thereafter, however, the remaining Astro E detectors would have had comparable energy resolution to that of the ASCA CCDs, and broader energy coverage (0.3–700 keV) than either RXTE or BeppoSAX. The first Astro E satellite was launched in early 2000. Unfortunately, attitude control of the launch vehicle was lost, and Astro E was placed in a shallow, elliptic orbit, from which it fell to Earth a short while later. Following much the original design plan, however, the Astro E2 mission is currently being built. It will have all of the originally intended capabilities of Astro E, and a successful launch is expected in 2005. NASA is a major collaborator in this endeavor.

5.3.2. Constellation-X

NASA’s Constellation-X is an observatory that is in the early stages of design and planning, with a proposed for launch date of c.2009. Similar to Astro E, the prime detector instruments are expected to be X-ray micro-calorimeters, albeit with total effective area greater than any current or previous X-ray satellite ($A \sim 3$ m$^2$). The planned energy range will be 0.2–40 keV, with a spatial resolution of 6 arcsec. The large effective area is to be achieved by “mass-producing” and launching four separate satellite systems with their own independent mirror and detector assemblies. The combination of high spectral resolution with large effective area has raised the prospects that Constellation-X can be
used for detailed iron line variability studies, including “reverberation mapping” between continuum and line features, as we further discuss below.

5.3.3. XEUS

ESA’s XEUS, the X-ray Evolving Universe Mission, is a potential successor to XMM-Newton. The current plan envisions two phases for the mission. The mirrors and detectors are planned to be housed in two separate spacecraft assemblies, with both being in close proximity to the International Space Station. The detectors will cover the 0.05–30 keV range, with a collecting area of 6 m² (i.e., approximately 10 times the peak effective area of RXTE) at energies of 1 keV in the first phase of the mission. This is planned to be expanded to 30 m² at 1 keV and 6 m² at 8 keV in the second phase of the mission. The instruments will have energy resolution of 1–10 eV (i.e., \( \Delta E/E \approx 1000 \) at 6 keV) and spatial resolution of 2 arcsec. Like Constellation-X, this would be a very powerful combination of large effective area and high spectral resolution, making XEUS ideal for reverberation mapping studies, as we further describe below.

5.3.4. MAXIM

The proposed Micro-arcsecond X-ray Interferometry Mission (MAXIM) aims to obtain sub-microarcsecond resolution X-ray images of nearby galactic nuclei (especially our Galactic Center, M87, and nearby Seyfert galaxies). At this resolution, one can resolve X-ray emitting structures on the scale of the event horizon of the central massive black hole. Furthermore, one may be able to map, in both spatial and velocity coordinates, the iron line fluorescence across the surface of the accretion disk. This would provide unprecedented and direct constraints on the dynamics of the disk. The technological developments required to construct this ambitious observatory are challenging, requiring one to perform X-ray interferometry in deep space across baselines of hundreds of meters. However, these difficulties are not insurmountable, and it is reasonable to suppose that such an observatory could be constructed within 20 years, with a less able “pathfinder” mission in 10 years or so.

6. Iron lines from active galactic nuclei

The cleanest examples of observed relativistic iron lines have been found in AGN. There are three principal reasons why AGN tend to present cleaner disk reflection signatures than GBHCs. Firstly, the primary X-ray continuum is often found to be very featureless, being well modeled by a simple power-law form. This makes it straightforward to subtract the continuum in order to study the underlying disk reflection signatures. The physical reason underlying the simplicity of the continuum form in AGN is the wide separation of temperatures between the accretion disk surface (with \( kT \sim 10 \text{ eV} \)) and the disk corona (with \( kT \sim 100 \text{ keV} \))—thermal Comptonization of the optical/UV disk emission by the corona produces a featureless power-law spectrum across much of the X-ray band. The accretion disks in GBHCs, on the other hand, can get so hot (\( kT \sim 1 \text{ keV} \)) that the Wien tail of the thermal disk emission can overlap with the red-wing of any iron line, making continuum subtraction much more challenging. Secondly, we expect there to be a wide range of parameter space over which AGN accretion disks are not strongly ionized, thereby simplifying the study of the X-ray reflection signatures. On the other hand, GBHCs may well be generically ionized.
due to the high temperature of the disk. Finally, since GBHCs typically lie in the plane of the Galaxy, they are usually much more heavily absorbed than type-I AGN, further complicating the modeling of the continuum radiation.

6.1. A case study of MCG–6-30-15

There are a small number of AGN that have become test-beds for the use of X-ray observations to probe black hole and accretion disk physics. Here, we shall describe studies of one such object—the Seyfert 1 galaxy MCG–6-30-15. This is a somewhat unremarkable S0-type galaxy in the constellation of Centaurus, with a redshift of $z=0.008$, placing it at a distance of $D=37$ Mpc. As we will describe, X-ray studies of this object have allowed us access to some of the most exotic black hole physics yet observed.

Unfortunately, the mass of the supermassive black hole in this AGN remains rather poorly determined—some general arguments place it in the range $10^6–2 \times 10^7 M_{\odot}$ [203]. Most of the astrophysical conclusions drawn below are robust to this uncertainty since, to first order, the mass of the black hole merely scales the size of the system rather than changing any of the underlying physics. However, constraints on the mass are important if we are to physically interpret observed temporal variability, since all of the fundamental time scales of the accretion disk scale linearly with black hole mass.

6.1.1. The first detection of a relativistic accretion disk

MCG–6-30-15 was the first AGN where X-ray reflection from fairly neutral material was clearly detected using the EXOSAT [201] and Ginga [204] observatories. This demonstrated that there was cold and optically thick material in the vicinity of the SMBH that was being irradiated by the power-law X-ray continuum. It was suspected by many that this material was indeed the accretion disk. However, direct confirmation of this had to await the superior spectral capabilities of ASCA.

As has already been described, ASCA was the first observatory capable of obtaining medium resolution 0.5–10 keV X-ray spectra of cosmic sources. Thus, it must be appreciated that the demands for observing time on ASCA were intense, with many different sub-communities of the astrophysical world wishing to utilize this new resource to explore the universe (indeed, ASCA has made major contributions in fields as diverse as stellar coronae, the diffuse interstellar and intracluster medium, and gamma-ray bursts). Furthermore, at the time when ASCA started operating, there was a certain level of skepticism within the astrophysical community concerning the possibility of detecting strong relativistic effects via spectroscopic studies of AGN—thus, since the majority of observing time on ASCA was allocated in open peer-reviewed competitions, it was not a foregone conclusion that ASCA would spend significant time taking high-quality spectra of AGN. The fact such spectra were obtained, thereby allowing the direct detection of relativistic effects around supermassive black holes, is largely due to the perseverance of Yasuo Tanaka (the Project Scientist) and Andy Fabian, for which they were awarded the Rossi Prize of the American Astronomical Society in 2001.

Returning to the history of MCG–6-30-15, a rather short (half-day) ASCA observation of MCG–6-30-15 showed the iron line to be significantly broadened in energy space, being well modeled as a Gaussian profile with centroid energy $E_0=6.2$ keV and a standard deviation of $\sigma=0.7$ keV
Fig. 12. Time-averaged (upper panels) and peculiar line profiles (lower panels) of the iron K emission from MCG–6-30-15 seen in the two long ASCA observations in 1994 (left) and 1997 (right). In the 1994 observation, a very broad profile with a pronounced red-wing is seen during a period of Deep Minimum of the light curve (lower left), compared to the time-averaged line profile shown in the upper panel. In contrast, during a sharp flare in the 1997 observation, whole line emission is shifted to energies below 6 keV and there is no significant emission at the rest line-energy of 6.4 keV (lower right). Both peculiar line shapes can be explained by large gravitational redshift in small radii on the accretion disk.

[205,206]. The improved signal-to-noise from a much longer (4.5 days\textsuperscript{11}) ASCA observation was required to determine that the iron line did, indeed, possess the profile expected from the surface of a relativistic accretion disk [20]. As shown in Fig. 1, this line profile is well fit with line emission from a disk around a Schwarzschild black hole with a disk inclination of $\theta \approx 27^\circ$ (measured away from the normal to the disk plane), with line flux distributed across the disk proportional to $r^{-3}$, extending down to $r = 6M$.

In fact, these ASCA data were a great deal richer than just described. As is common with type-1 AGN, the X-ray flux of this AGN varied rapidly as a function of time. Via a detailed analysis of this and subsequent ASCA observations, Kazushi Iwasawa found that the iron line flux and profile also undergoes dramatic changes with time [207,208]. The detailed iron line changes are complex and defy simple characterization (see Fig. 12). In the July 1994 ASCA observation, it was found that, during a large flaring event, the iron line became quite narrow with an energy close to the intrinsic

\textsuperscript{11} During the early stages of the ASCA mission, this was an observation of unprecedented length. The typical ASCA observation rarely exceeded one day.
value. This suggests that the observed fluorescence was dominated by fairly large radii in the disk during the flare. On the other hand, during another flare observed in a 1997 ASCA observation, the line emission redshifted entirely to energies below 6 keV suggesting dominance by small radii in which gravitational redshifts are large.

Of particular interest was the identification of the so-called “deep minimum” (DM) state. During the DM state, the continuum flux falls below its average value by a factor of 2–3, and the fluorescent iron line appears to become much broader and stronger. In fact, the redshifts needed to explain the observed line profile are so great as to require line emission from within \( r = 6M \). As we shall see, we believe that the DM state gives us a window into the more exotic astrophysics of rapidly rotating black holes.

Of course, the claim that iron line studies are probing the region within a few gravitational radii of the black hole is a bold one, and should be tested against other models at every opportunity. Given the quality of data, the July 1994 MCG–6-30-15 line profile has become a traditional test bed for such comparisons. The first issue to assess is the correctness of the continuum subtraction that underlies the determination of the iron line profile. Andy Fabian and collaborators systematically examined the effect of unmodeled spectral components (e.g., absorption edges) and found that the line profile was robust to all physically reasonable possibilities [209]. Given the reality of the spectral features, one must then critically assess its interpretation as an emission line arising from a relativistic accretion disk. The most serious challenger is a model in which a source of intrinsically narrow iron line emission is surrounded by a rather optically thick (\( \tau \sim 5 \)), cool (\( kT \sim 0.2 \) keV) and yet highly ionized plasma. Repeated Compton recoil suffered when the photons scatter off the electrons can then broaden and redshift the observed line [210]. However, such a scenario predicts that the continuum spectrum passes through the same Compton cloud, filtering out high-frequency temporal variability and imposing a continuum spectral break at \( E \sim m_e c^2 / \tau^2 \). The presence of high-frequency temporal variability and the absence of a continuum spectral break strongly argues against such a scenario [211,212].

In another alternative model, it has been has proposed that energetic protons transform iron in the surface of the disk into chromium and lower \( Z \) metals via spallation which then enhances their fluorescent emission [213]. With limited spectral resolution, such a line blend might appear as a broad skewed iron line. This model suffers both theoretical and observational difficulties. On the theoretical side, high-energy protons have to be produced and slam into the inner accretion disk with a very high efficiency (one requires 10% of the binding energy of the accreted material to be channeled into this process alone). On the observational side, it should be noted that the broad line in MCG–6-30-15 [214] is well resolved by the ASCA SIS (the instrumental resolution is about 150 eV at these energies) and it would be obvious if it were due to several separate and well-spaced lines spread over 2 keV. There can of course be Doppler-blurring of all the lines, but it will still be considerable and require that the redward tails be at least 1 keV long.

6.1.2. Is the black hole rapidly spinning?

The very broad line seen in the ASCA DM state of MCG–6-30-15 leads us to consider rotating black holes. If, for now, we make the assumption that there is no fluorescent emission line from within the radius of marginal stability, the DM spectrum requires a rapidly rotating black hole in order to have \( r_{\text{ms}} \) appreciably less than \( 6M \) [207]. In fact, if one assumes that the iron line flux is distributed across the disk according to Eq. (19), the formal limit on the dimensionless spin parameter of the black hole is \( a > 0.94 \) [215].
There has been a great temptation within the X-ray astrophysics community to interpret very broad iron lines, with $r_{\text{in}} < 6M$, as direct evidence for black hole rotation. However, when one attempts to make these arguments rigorous, a loophole is rapidly discovered. By its basic nature, one can see that it is possible to produce arbitrarily redshifted spectral features from around any black hole if any region of the disk (even those within the plunging region) is permitted to produce observable line emission [216]. As the emitting region tends to the event horizon, the redshifts grow arbitrarily large. Thus, an attempt to determine the space–time geometry (and, in particular, the spin parameter of the black hole) from the time-averaged iron line profile necessarily involves considerations of the astrophysics of the inner accretion disk and, in particular, the extent to which emission/reflection from the plunging region of the disk might be important.

Within the standard model for black hole accretion (i.e., that of Page et al. [115,114]), there is no angular momentum transport and no dissipation of energy within the plunging region. Hence, this region of the disk cannot support an X-ray active corona. Furthermore, if the corona of the dissipative part of the disk has a small scale-height, the surface of the plunging region will not be subjected to X-ray irradiation for purely geometric reasons, and will not produce X-ray reflection signatures. These considerations would validate the assumption of no X-ray reflection signatures from within the plunging region thereby allowing the spin of the black hole to be determined in the manner discussed above. However, there are two scenarios in which appreciable X-ray reflection might occur from part of the plunging region. Firstly, it is possible that the outer disk corona might be geometrically thick, thereby allowing it to irradiate a dissipationless plunging region. Given this scenario, the ASCA data for the DM-state of MCG–6-30-15 data could be consistent with even a non-rotating (Schwarzschild) black hole [216]. Secondly, as discussed in Section 3.2.4, the accretion flow within the plunging region may well be able to dissipate a significant fraction of its binding energy. One could easily imagine the formation of a powerful X-ray emitting corona sandwiching the plunging region of the disk. Although this scenario has not been worked out in detail, such an inner corona might be responsible for locally produced X-ray reflection within the plunging region. It should be noted that the plunging region is likely to be highly ionized, and so any detailed study of X-ray reflection from this region must use self-consistent ionized reflection models as were discussed in Section 3.5.2 [217].

The above history made MCG–6-30-15 a prime target for study with the high-throughput EPIC instruments on board the XMM-Newton observatory. The first XMM-Newton observation of this source happened to catch it in a prolonged DM state. The superior sensitivity of these instruments allowed a much more detailed study of this enigmatic state to be conducted [218].

In agreement with the ASCA-based conclusions of Iwasawa, the X-ray reflection features were found to be very broad and very strong (see Fig. 13). The detailed modeling of the XMM-Newton/EPIC spectrum required the modeling of the full reflection spectrum, not just the fluorescent iron line, with relativistic shifting/smearing applied to both the fluorescent line and the reflection continuum. Again, the reflection features were found to be so broad and redshifted as to require reflection from within $r = 6M$. Assuming a background space–time corresponding to a near-extremal Kerr black hole (with $a^* = 0.998$, giving $r_{\text{in}} = 1.23M$), the data constrain the inner edge of the reflecting region to lie closer than $r_{\text{in}} < 2M$, with a steep emissivity profile outside of that radius of $4.5 < \beta < 6.0$. This line profile can be successfully modeled assuming a Schwarzschild (non-rotating black hole) background (see discussion in Section 6.1.2), but very extreme parameters are demanded—the inner edge of the reflection region is constrained to be $r_{\text{in}} < 3M$, very deep within the plunging region,
Fig. 13. XMM-Newton data from the first observation of MCG–6-30-15. Panel (a) shows the ratio between the data and a simple model consisting of a power-law fitted to the 0.5–11 keV data. Panel (b) shows the ratio of data to a more detailed model which also accounts for line-of-sight absorption towards the accreting black hole. The residuals in this plot are attributed to X-ray reflection from the relativistic disk. (c) Deconvolved spectrum of the iron K band portion of the spectrum, showing a very broadened and redshifted iron line. Figure from [218].

with an extremely large emissivity index $\beta > 10$. Thus, in the Schwarzschild scenario, essentially all of the fluorescence has to originate from an annulus in the innermost parts of the plunging region, $2M–3M$.

So, the possibility of X-ray reflection signatures from within the radius of marginal stability remains a wild-card affecting the detailed interpretation of these data. While there are no models that treat such emission in detail, the extreme parameters required to explain the XMM-Newton data within the context of a Schwarzschild black hole ($r_{\text{in}} < 3M$ and $\beta > 10$) seem entirely unreasonable. At the very least, one would expect this material to be in an extremely high ionization state and incapable of producing iron fluorescence and other X-ray reflection features. Thus, the case for a rapidly spinning black hole in this object is now very strong, even accounting for the possibility that there may be some emission from the plunging region.
6.1.3. Is the spinning black hole energizing the disk?

To explore the implications of the MCG–6-30-15 XMM-Newton data further, we employ the standard assumption that there is little or no X-ray reflection from within the plunging region. To make a connection with theoretical accretion disk models, we must relate the distribution of observed X-ray reflection to the distribution of primary energy dissipation in the accretion disk (since it is the latter that is predicted by the models). The simplest assumption is that the local intensity of X-ray reflection from the disk surface is proportional to the local dissipation in the underlying disk. This would be the case if the ionization of the disk surface was reasonably uniform with radius, and a fixed fraction of the locally dissipated energy was transported into a geometrically thin X-ray emitting corona. In reasonable disk models, any appreciable ionization changes would be in the sense of the innermost regions being more highly ionized, therefore producing weaker X-ray reflection signatures and requiring an even steeper dissipation profile than inferred from the X-ray reflection studies. We also note that, in MCG–6-30-15, a large fraction (25–50%) of the electromagnetic luminosity is emitted in the X-ray band by the Comptonizing corona. Thus, it cannot be grossly wrong to assume that the X-ray activity traces the underlying energetics of the disk.

Taking the X-ray reflection as a proxy for the underlying disk dissipation, we see that these data are clearly discrepant with the standard models for black hole accretion disks for any spin parameter. As discussed in Section 3.2.2, the flux emitted per unit proper area of a standard disk around a near-extremal Kerr black hole, $D_{PT}(r)$, is zero at $r = r_{ms}$ by virtue of the ZTBC, peaks at $r \sim 1.6r_g$ and then gradually steepens to approach $D(r) \propto r^{-3}$. At no point does $D_{PT}(r)$ become as steep as $r^{-4.5}$, as required by the XMM-Newton observations. This is true for any assumed BH spin. However, as discussed in Section 3.2.4, magnetically induced torques at the radius of marginal stability can significantly enhance the amount of dissipation in the innermost regions of the accretion disk. The ‘best’ possible case (in terms of producing steep dissipation profiles) is where the magnetic torque is applied entirely at the radius of marginal stability. In Fig. 14, we plot the logarithmic derivative of the dissipation profiles shown in Fig. 5. One finds that it is still impossible to achieve the required steep dissipation profiles in the case of non-rotating or slowly rotating black holes. Only in the case of a rapidly rotating black hole does the dissipation profile become sufficiently steep. In fact, one finds that sufficiently steep dissipation profiles can only be obtained in the regions of parameter space where an appreciable fraction of the extra energy is derived directly from the spin of the black hole. As discussed in Section 3.2.4, the physical origin of this torque is likely a magnetic connection between the accretion disk and either the plunging region or the rotating event horizon itself.

Relativistic effects may also reveal themselves through the strength of the reflection features. These XMM-Newton data show that, during the DM state, the relative strength of the reflection is about twice that expected from a plane-parallel coronal geometry. While there may be geometrical explanations, such an enhancement is in fact a natural consequence of returning radiation [196,218].

It must be reiterated that, while the case for a rapidly rotating hole in MCG–6-30-15 is solid, the argument for a torqued accretion disk is vulnerable to the possibility of powerful fluorescence/X-ray reflection from within the plunging region. More theoretical work on the plunging region is required to assess the impact it might have on these arguments. There is also another way of alleviating the need for a torqued accretion disk. Martocchia and Matt have recently suggested that the DM spectrum of MCG–6-30-15 could be explained via the gravitational focusing of X-rays emitted by a source that is $r \sim 3M$ from the black hole at high latitudes (i.e., close to the spin axis of the
black hole) [219]. One problem suffered by this scenario is that it overpredicts the relative strength of the X-ray reflection by a factor of two. Thus, for this picture to be correct, only half of the disk surface area can be in a state capable of producing X-ray reflection signatures. Even if this picture is correct, the presence of such a powerful X-ray source on the spin axis of the black hole (presumably corresponding to the base of an MHD jet) is itself a strong circumstantial indicator of black hole spin-energy extraction.

The first results from a later and very long (400 ks) XMM-Newton observation of MCG–6-30-15 were reported recently by Andy Fabian and collaborators [220]. This observation has produced the best iron line profile yet obtained from any object (Fig. 15). There is strong evidence for line profile variability between the two XMM-Newton observations, with the latter observation showing a prominent blue horn to the line profile. This horn indicates the presence of emission from larger radii in the disk. However, the extreme red-wing to the line profile, requiring a very steep central emissivity profile, was also present in these data. We await further analysis of this tremendously important dataset.

6.1.4. Relativistically broadened soft X-ray emission lines

Before leaving MCG–6-30-15, we will briefly discuss a fascinating debate that is currently raging about the soft X-ray spectrum of this object. The soft X-ray spectrum of this, and many other type-1 AGN, is rather complex. Prior to the launch of XMM-Newton, it was widely accepted that this complexity was due to absorption by photoionized material along the line of sight to the central engine (the so-called warm absorber; [221,222]), with the principal observables being the photoelectric absorption edges of OVII and OVIII. In many objects, this hypothesis has been confirmed by high-resolution grating spectroscopy with Chandra and XMM-Newton; such spectra clearly display the resonant absorption lines of oxygen that are expected to accompany the photoelectric edges
Fig. 15. Continuum subtracted iron line profile from the second XMM-Newton observation of MCG–6-30-15 [220]. This is the highest signal to noise relativistic line profile yet measured.

(e.g., see the fabulous X-ray spectrum of NGC 3783 obtained by Kaspi and collaborators [223]). However, in MCG–6-30-15, an XMM-Newton grating spectrum showed that the simple warm absorber model fails—there are anomalous “edges” and an apparent lack of strong oxygen resonance absorption lines [224]. On the basis of this, Branduardi-Raymont and collaborators suggested that the soft X-ray spectral complexity is due to strong, relativistically broadened oxygen, nitrogen and carbon recombination lines rather than a warm absorber [224]. The team led by Julia Lee analysing the non-simultaneous Chandra grating data can, however, explain the soft X-ray spectrum in terms of a warm absorber which contains embedded iron-rich dust [225]. It remains unclear whether the dusty warm absorber model can explain the XMM-Newton grating data.

Given the rapid and on-going development of these arguments (both for and against the idea of soft X-ray recombination lines that possess relativistic disk profiles), we refrain from drawing conclusions at this time. The interested reader is pointed to the astrophysical literature on X-ray spectroscopy of MCG–6-30-15.

6.2. Other “normal” Seyfert nuclei

While MCG–6-30-15 has been subjected to particularly detailed study, there have also been intensive studies of other Seyfert galaxies. Given the limited collecting area of ASCA, multi-day observations of bright AGN were required in order to obtain reasonable signal-to-noise in the relativistic iron line. Prior to the existence of many such long datasets, Paul Nandra and co-workers showed that the relativistic iron line is present in the “average” ASCA spectrum of a sample of two dozen Seyfert-1 galaxies [226]. However, in the past 5 years, deep observations of Seyfert-1 galaxies with ASCA and, more recently, Chandra and XMM-Newton have been performed. Here we summarize the results of these studies.
In addition to MCG–6-30-15, relativistically broad iron lines with very high signal-to-noise have been found by ASCA in the Seyfert galaxies NGC 3516 and NGC 4151. In NGC 3516, the red-wing of the iron line appears to track changes in the continuum flux, as expected by the simple reprocessing model. The blue-wing of the iron line, however, displays higher amplitude variability that is uncorrelated with the continuum flux, suggesting more complex changes in the pattern of fluorescence across the disk [227,228]. More interestingly, the ASCA data suggested the presence of an absorption line at $\sim 5.9$ keV (absorbing flux from the broad iron emission line). If real, this feature may arise from redshifted resonance absorption by ionized iron in tenuous plasma above the disk, with the energy shift due to either gravitational redshift [229] or infall of the material [228]. If the latter interpretation is true, this may be a rare detection of material actually in the process of accreting into the black hole. Recent simultaneous Chandra grating and XMM-Newton observations of NGC 3516 have also revealed fascinating substructure within the broad line profile in the form of narrow emission spikes at 5.6 and 6.2 keV [230]. These may be the first indications of the expected transient non-axisymmetric illumination.

The high-quality ASCA data on NGC 4151 also reveals interesting variability and complexity within the line profile [231–233]. During the 1995-May ASCA observation, the iron line profile displayed significant variability, especially in the strength of the red-wing, despite relatively small changes in the observed continuum [233]. On the other hand, during the 2000-May observation, the opposite conclusion appears to hold, with the continuum undergoing significantly larger changes than the iron line [234]. Thus, it appears that the iron line in a given AGN can change its variability behavior, although the non-uniformity of the data analysis methods employed in these (independent) studies prevents one from making a straightforward conclusion.

In comparison with the very broad accretion disks iron lines in MCG–6-30-15, NGC 3516 and NGC 4151, the well-known Seyfert galaxy NGC 5548 has developed a reputation for displaying a rather narrow iron line. Fitting accretion disk models to ASCA data for this object suggested that the line emitting region of the disk was truncated at an inner radius of about $r = 10M$ [235,236]. However, we now know that this source displays a composite iron line [237]. The Chandra-HETG has revealed a “narrow” core to the line (with a velocity dispersion of $4500^{+3500}_{-2600}$ km s$^{-1}$) originating via fluorescence of material a substantial distance from the black hole. The large equivalent width of this narrow core ($130^{+60}_{-50}$ eV) suggests that this material subtends a large part of the sky as seen by the primary X-ray source. Once this component is accounted for in medium-resolution spectra, the Chandra data are entirely consistent with a relativistic accretion disk extending down to the radius of marginal stability [237]. However, the last word has yet to be spoken about NGC 5548. Very recently (in fact, after the initial submission of this review), there was a convincing non-detection of the broad line in NGC 5548 by XMM-Newton [238]. In order to make sense of results from these different observatories, we are forced to conclude that broad iron lines, in at least some objects, are transitory.

However, our experiences with NGC 5548 does illustrates a point of some general practical importance—at medium spectral resolution, narrow (i.e., low-velocity) iron line components can be blended together with the relativistic line profile and, unless account for, produce significant systematic errors in measurements of, for example, the disk inclination, emissivity profile and inner emitting radius. For example, it has been claimed on the basis of ASCA measured iron line profiles that the rather absorbed “intermediate” Seyfert galaxies are viewed face-on [240], a result that runs counter to our current understanding of AGN geometry (see Fig. 16). However, this result is likely to be an artifact of blending between relativistic iron line profiles and narrow lines—including narrow
Fig. 16. Model line profiles for a narrow iron line superposed on an iron line from an accretion disk viewed at 46° (solid line), compared with a line from a disk viewed at an inclination of 20°. These two cases would be difficult to distinguish in data of limited resolution and signal-to-noise. Figure from [239].

lines with reasonable strengths in the spectral models shows that the data for these systems are consistent with fairly edge-on disks [239]. Chandra has made a major contribution in this respect—high resolution grating spectra with the Chandra/HETG can unambiguously identify the narrow iron line component, thereby allowing it to be subtracted from the medium resolution spectra of the relativistic line profiles. This is another sense in which MCG–6-30-15 is a particularly good system for studying relativistic line profiles—Chandra/HETG spectra show that it has, at most, a very weak narrow line component [241] (which, interestingly, suggests that the larger-scale environment of this SMBH is very “clean” with little evidence for an obscuring torus).

NGC 5548 was also the first source for which a particular mystery associated with spectral variability was discovered. Jim Chiang and collaborators performed contemporaneous observations of NGC 5548 with ASCA and RXTE, thereby allowing details of the iron line and the reflection continuum to be constrained simultaneously [236]. As the continuum source varied, it was found that the iron line flux was fairly constant or, stated another way, the equivalent width was inversely proportional to the continuum flux. This is partially explained by the narrow component to the iron emission line noted above, which occurs in material light weeks or more from the black hole. The light travel time eliminates any line variability over the course of the month-long ASCA/RXTE monitoring campaign. However, at the same time, it was found that the continuum reflection fraction as measured by RXTE was positively correlated with the continuum flux. In other words, the relative strength of the fluorescent iron line appeared to be anti-correlated with the relative strength of the reflection continuum (see Fig. 17). Since the iron line and reflection continuum are different facets of the same phenomenon (i.e., X-ray reprocessing in the disk atmosphere), this result is most puzzling. In principle, unmodeled flux correlated changes in the ionization structure of the accretion disk atmosphere might decouple the iron line and apparent reflection continuum strength in this way [203], although the success of detailed ionized disk models in explaining this observation has yet to be demonstrated. Very similar behavior is also found in RXTE studies of MCG–6-30-15 by Julia Lee and collaborators [242].
To summarize the above discussion, iron lines from relativistic accretion disks appear to be generic features in the X-ray spectra of normal Seyfert-1 galaxies, although NGC 5548 and the intermediate Seyferts demonstrate that some care might be needed to extract the relativistic line profile from the data. In one source, MCG–6-30-15, the strength of the relativistic effects argues strongly for a rapidly rotating black hole. Furthermore, the data suggest that the central regions of the accretion disk may be extracting the black hole spin energy. However, substantial uncertainties remain. In particular, we do not possess even a phenomenological characterization of iron line variability that can predict how the line changes as a function of continuum flux or other source properties. Line variability is of particular interest since it undoubtedly is the key to substantial understanding of accretion disk physics.

The lack of a simple correlation between the iron line strength and the observed X-ray continuum has lead some researchers to throw out the whole picture of a relativistically broadened fluorescence iron line, leading them to favour rather complex and seemingly contrived X-ray continua as an explanation for the observed X-ray spectrum (e.g., [234]). However, in the opinion of one of us (CSR), this is an extreme over-reaction to the failure of a particularly simple-minded realization of the fluorescent iron line model. A realistic MHD turbulent disk with have a “surface” that, at a local
level, may be highly complex both in terms of geometry and ionization state. For example, it is easy to image that the (turbulent) surface is comprised of filaments and/or sheets which have a variety of ionization states. The resulting response of the iron line flux from such a complex surface to a change in the irradiating coronal flux (even assuming no direct thermal/dynamical coupling between the corona and the disk surface) is very unlikely to be linear.

The on-going operation of XMM-Newton promises to make significant progress in the near future. In addition to MCG–6-30-15 and NGC 3516, there are a small number of other normal Seyferts for which XMM-Newton data has already been reported [243–245,238]. On the basis of these early studies, it has been suggested that the ASCA studies may have over-estimated the generality of relativistic iron lines in the spectra of Seyfert galaxies. However, robust conclusions must await a more systematic analysis of current and future XMM-Newton datasets, part of which would include a reconciliation with the older yet still valid data from ASCA and RXTE.

6.3. Very-high and very-low luminosity AGN

From the observations discussed above, we can deduce that the accretion disks in most Seyfert-1 nuclei are only modestly ionized and remain optically thick (and capable of producing iron line features) down to the radius of marginal stability. Seyfert-1 nuclei provide a useful reference against which we can compare and contrast the accretion disks of other types of AGN. By making such comparisons, we can hope to learn how the properties and physics of accretion disks change as a function of the gross observables of the system. Here, we examine constraints on the dependence of the accretion disk properties on the overall radiative luminosity of the AGN.

We begin by discussing low-luminosity AGN (LLAGN), roughly defined as sources with a total radiative luminosity of $L < 10^{42}$ erg s$^{-1}$. Being low-luminosity, even the nearest examples are comparatively faint X-ray sources. Furthermore, stellar processes in the galaxy (such as X-ray binaries and vigorous star-formation) can produce a comparable X-ray flux. This galactic emission can be spatially blended with the LLAGN emission and somewhat hamper studies.

One of the best studied LLAGN is the nearby galaxy NGC 4258 (M 106)\textsuperscript{12} — this is the same object for which masers give an excellent constraint on the black hole mass. The first good constraints on the iron line in this source were obtained with a deep ASCA observation [246] (although previous shorter observations had hinted at such a line [247]). It was found that the iron line was narrow, implying that the bulk of the fluorescent emission originated from $r > 100M$. Given the fairly large equivalent width of the line ($\sim 100$ eV) and the lack of any evidence (from any waveband) for cold, optically thick material away from the plane of the accretion disk, this emission line very likely originates from the accretion disk itself. Thus, to produce such a narrow line, the X-ray emitting corona must be extended, producing appreciable X-ray illumination of the disk across a range of radii exceeding $r > 100M$. This is in stark contrast to the higher-luminosity Seyfert galaxies whose broad lines require a very compact X-ray emitting corona ($r < 10M$). This result is consistent with a sphere+disk model for this LLAGN with a transition radius of $r_{tr} > 100M$. However, due to the limited signal-to-noise, the presence of a weaker relativistic iron line in addition to the narrow component could not be ruled out. Recently it has been reported that a rather short XMM-Newton observation of NGC 4258 failed to detect any iron line, setting an upper limit on the line flux to

\textsuperscript{12} More precisely, this source is classified as a low-ionization nuclear emission region (LINER).
be lower than that detected by ASCA [248]. Such variability suggests that the iron line does indeed originate from the accretion disk (rather than more distant material) and, furthermore, implies that the disk/corona system undergoes significant changes in structure and/or ionization state on year time scales.

X-ray studies of the general population of LLAGN have been hampered by their comparative faintness. The best such study to date, based upon ASCA data, detects significant iron line emission in several LLAGN [249,250]. In no object, however, is the line emission found to be significantly broadened. Unfortunately, the signal-to-noise of these datasets is insufficient to draw robust astrophysical conclusions—relativistic iron lines such as those seen in higher-luminosity Seyfert nuclei could readily be hiding in the noise of these data. However, there are excellent prospects for characterizing any broad line component with current and future XMM-Newton and Chandra observations. Such observations will allow us to compare the properties of accretion disks (especially ionization state and truncation radii) in LLAGN with those in normal Seyfert nuclei.

We now turn to a brief discussion of high-luminosity AGN (HLAGN). Again, due to the maturity of the data, much of our understanding is still largely based on ASCA observations. In an important study, Nandra and co-workers studied the average ASCA spectrum of a large sample of AGN as a function of source luminosity ranging from normal Seyfert nuclei to powerful quasars [251]. They found that broad iron line becomes appreciably weaker once one considers sources with an X-ray luminosity greater than $L_X \sim 10^{44} - 10^{45}$ erg s$^{-1}$. The most reasonable interpretation for this trend is that the HLAGN possess more highly ionized accretion disks—this is expected if the luminosity of an AGN is determined primarily by the Eddington fraction rather than the black hole mass. Support for this picture is growing with subsequent XMM-Newton observations. XMM-Newton studies of the high-luminosity Seyfert galaxies Mrk 205 and Mrk 509 (which possess an X-ray band luminosity of $\sim 10^{45}$ erg s$^{-1}$) find significant broadened K$\alpha$ iron lines, with profiles well described with accretion disks models, and energies corresponding to helium- or hydrogen-like iron [252,253]. These observations likely represent direct detections of ionized accretion disks.

Direct evidence for ionized disks have also been found in the X-ray spectra of the so-called narrow-line Seyfert-1 galaxies (NLS1s [254,255]). NLS1s are a peculiar sub-class of Seyfert-1 galaxies which are defined by their rather narrow optical emission lines, together with very soft and highly variable X-ray emissions [256]. While they do not possess a high luminosity in absolute terms, it is thought that NLS1s may have rather small black hole masses and luminosities rather close to their Eddington limit. The detection of highly ionized disks in these objects nicely fits in with this scenario.

### 6.4. Iron lines from radio-loud AGN: a clue to the dichotomy?

One of the greatest mysteries in the field of AGN research is the physical mechanism underlying the radio-quiet/radio-loud dichotomy. Why do some black hole systems choose to launch, accelerate and collimate powerful relativistic jets, whereas other systems do not? The natural environment in which to form an energetic relativistic jet is the relativistic region close to an accreting black hole. A first step in addressing this problem from an observational standpoint is to compare and contrast the central engine structures of radio-quiet and radio-loud AGN. Since they are direct probes of the inner accretion disk, broad iron lines provide a powerful way of facilitating such a comparison.

Radio-loud AGN are rarer, and hence typically fainter, than their radio-quiet counterparts. Furthermore, many of the best candidates for study are found in clusters of galaxies and it can be difficult
to observationally distinguish AGN emission from thermal X-rays emitted by $10^7$–$10^8$ K gas trapped in the cluster’s gravitational potential well. For these reasons, the quality of the observational constraints are rather poorer than for Seyfert-1 galaxies. There does, however, appear to be a difference between the iron line properties of radio-loud nuclei and radio-quiet nuclei. Broad iron lines, and the associated Compton reflection continua, are generally weak or absent in the radio-loud counterparts [257–263]. This effect might be due to the swamping of a normal ‘Seyfert-like’ X-ray spectrum by a beamed jet component (similar to the swamping of optical emission lines in a blazar spectrum). Alternatively, the inner disk might be in the form of a very hot and optically thin radiatively inefficient accretion flow. Finally, the inner disk might be radiatively efficient and optically thick, but not produce prominent X-ray reflection features due to a very high ionization [264]. Again, future observations with XMM-Newton should be able to distinguish these possibilities by searching for very weak broad components to the iron line and iron edges with high signal-to-noise.

7. Iron lines from galactic black hole candidates

A crucial difference between GBHCs and AGN are their characteristic time scales. Except for the least massive AGN, the viscous time scales characterizing the inner disk can be many tens of years. Since the viscous time is the timescale on which the mass accretion rate can change appreciably, we see that it is impossible to sample a range of mass accretion rates within a particular AGN on a human timescale. On the other hand, the viscous time scales in GBHCs can range from days to weeks, so individual observations can, and likely do, represent different intrinsic accretion rates. The same is true for dynamical and thermal time scales. A typical X-ray observation lasting several tens of thousands of seconds represents only a few thermal and dynamical time scales in an AGN, but many thousands of such time scales in GBHCs. In fact, a reasonable estimate of the thermal time scale in a stellar mass black hole corresponds to a few seconds, which often also corresponds to the time scale of the peak variability power in the X-ray lightcurve of such sources.

In many ways, because of these differences in characteristic time scales, the study of stellar mass black hole candidates and AGN are complementary. One figure of merit is the square root (since one is typically dealing with Poisson statistics) of the flux received over a characteristic time scale, i.e., $\sqrt{\left(\frac{L}{D^2}\right)(\frac{GM}{c^3})}$, where $L$ is the source luminosity and $D$ is its distance. For sources with the same fractional Eddington luminosity, the luminosity scales as mass, so this figure of merit scales as $(M/D)$. Thus, although AGN are one thousand or more times further away than stellar mass black holes, they are also up to one hundred million times more luminous and therefore are much more effectively observed over their characteristic dynamical time scales. Stellar mass black holes, however, are much more readily observed over multiple viscous time scales. Ten thousand seconds in the life of a GBHC can be equivalent to as much as a thousand years in the life of an AGN.

7.1. A case study of Cygnus X-1

The GBHC Cyg X-1 has been extensively studied by every major X-ray satellite of the past 30 years, and it has been the focus of extensive theoretical modeling. It spends a large fraction of its time in a spectrally hard, highly variable X-ray state; however, it occasionally transits to a softer, less variable state [265–267]. Several such transitions to a soft state have been observed with
RXTE [61,62,268]. The geometry within each of these accretion states, as well as the causes of the transitions among the different states, has been the focus of much theoretical speculation and observational study. It is precisely these issues that studies of X-ray lines and reflection features hope to illuminate.

Before discussing the specifics of the line studies, it is worthwhile to review some of the suggested models for the Cyg X-1 system, and to briefly consider their implications for predicted line properties. As Cyg X-1 spends most of its time in the hard state, models have focused on explaining this state. Soon after the discovery of hard state black hole candidates, it was suggested that these spectra were due to Comptonization of soft photons in a hot, low optical depth plasma [135], and such spectra were soon successfully applied to the hard state of Cyg X-1 [136,148]. Comptonization models have invoked, variously, slab geometries (and their variations; see Fig. 6), sphere+disk geometries, and have posited soft seed photons coming from either the optically thick, geometrically thin disk or from synchrotron photons generated within the corona itself.

For the slab or sandwich geometries, it is expected that the generated reflection fraction will be of order unity (half the coronal flux is intercepted by the disk), and thus iron line equivalent widths should also be large (∼ 200 eV). If the corona extends inwards toward small disk radii, then relativistic effects will also be prominent. The ‘pill box’ geometry preserves the reflection and line features of this geometry, but reduces the amount of Compton cooling of the corona. ‘Sphere+disk’ geometries were in part specifically considered since, like the pill box, these coronae are less Compton cooled than slab models. In contrast to the pill box models, they also have low reflection fractions (f ≈ 0.3) and produce weak iron line features. The degree to which relativistic effects are prominent in such line features is dependent upon the transition radius between corona and disk. ADAF models essentially adopt the ‘sphere+disk’ geometry; however, their hard spectra tend to be more centrally concentrated, and they often posit large transition radii between inner corona and outer disk. Thus, although not a necessary feature of ADAF models in general, many specific realizations of ADAF models produce very small reflection fractions, as well as very weak and narrow fluorescent iron line features.

The observational nature of the reflection and line features observed in Cyg X-1 has been controversial. In order to illustrate how technology, theoretical prejudice and sociology influence our understanding of such a complex system, we will present a detailed historical account of the iron line in Cyg X-1. The first reported detection of a broad Fe line feature in Cyg X-1 was made by Paul Barr and collaborators using data from the EXOSAT Observatory [200]. The line was seen to be of moderate strength (an equivalent width of 120 eV), have a slightly redshifted peak from that of neutral iron (6.2 keV, as opposed to 6.4 keV), and be fairly broad (Gaussian width of σ = 1.2 keV). This line, however, was not interpreted as indicating a relativistic profile. Instead, it was interpreted as due to Compton scattering in a moderately optically thick (τes ∼ 5), low temperature (∼ 5 keV) corona. Scattering in such a corona would indeed both broaden and slightly redshift an intrinsically narrow line. It was further hypothesized that the line was predominantly due to recombination radiation in a highly ionized corona, as opposed to being due to fluorescence radiation.

Several years later, however, these exact same observations were reinterpreted as relativistic broadening of a fluorescent line by Andy Fabian and collaborators [189]. In fact, this was the first attempt to claim such a line in any X-ray source. It was noted that, due to the poor intrinsic resolution of EXOSAT in the iron line region (∆E/E > 10%), the characteristic “double horned” features of relativistic lines would be smeared out, regardless of the inclination, and yield the broad, Gaussian
feature previously fit \[200\]. This interpretation was challenged, however, by prior observations performed with the Tenma X-ray satellite \[269\]. These observations, carried out in 1983 and spanning several phases of the Cyg X-1 binary orbit, revealed a weaker line (equivalent width 60–80 eV) that furthermore was consistent with being somewhat narrower (\(\sigma < 1 \text{ keV}\)) and being dependent upon orbital phase. This suggested, instead, fluorescence from the surface of the companion (donor) star, not the black hole accretion disk. A broader component emanating from the inner disk regions, however, could not be ruled out by the data.

As we have discussed extensively, a fluorescent iron line is produced along with other spectral signatures—one also expects to observe the iron K edge, as well as Compton recoiled photons at higher energy. When these features were added to the models of EXOSAT data (as well as to models of earlier HEAO 1-A2 data), Christine Done and collaborators obtained fits that yielded reflection fractions of \(\Omega/2\pi \sim 0.5\) and narrow Gaussian lines with equivalent widths \(\sim 40\) eV \[270\]. These results were confirmed by observations performed by BBXRT, which was an X-ray telescope with a resolution \(E/\Delta E \sim 30\) that was flown aboard the space shuttle in 1990. Specifically, it was found that the inclusion of a reflection component from material with twice the solar abundance of iron obviated the need for anything other than a weak, narrow iron line \[271\]. This model was then also applied to ASCA data by Ken Ebisawa and collaborators, and, again, when including an iron edge due to reflection, a narrow, weak iron line was found \[272,\] although see the further discussion below.

Unambiguously constraining spectral models that include line features (at 6 keV), reflection features (at 10–30 keV), and thermal Comptonization features (especially the high energy cutoff at \(\approx 100\) keV), requires broad band data. Such a data set for Cyg X-1 was obtained in 1991 with a simultaneous Ginga (\(\approx 2–30\) keV) and OSSE (\(\approx 50–1000\) keV) observation. Ginga, similar to EXOSAT, had poor spectral resolution (\(E/\Delta E \approx 10\)). Marek Gierlinski and collaborators described these observations \[273\] with models comprised of low reflection fractions (\(\Omega/2\pi = 0.2–0.5\)), and narrow iron lines with moderate equivalent widths (\(EW = 90–140\) eV). As described in Section 3.3, there was a growing realization that in order for coronal models to yield very hard spectra with thermal Comptonization cutoffs at very high energy, photon starved geometries were required \[144,155,274\]. A sphere+disk geometry, for example, could describe the Cyg X-1 data, and furthermore would imply a narrow iron line if the transition between the coronal sphere and the outer thin disk occurred at sufficiently large radius \[273,155,274\].

It has been the application of more sophisticated coronal and reflection models (with the latter including such effects as ionization of the disk atmosphere and relativistic smearing of the reflection features), coupled with the advent of the new generation of X-ray telescopes (RXTE, BeppoSAX, Chandra, XMM-Newton), that has led researchers to reconsider the possibility of a relativistically broadened iron line in the spectrum of Cyg X-1. An early observation of Cyg X-1 with RXTE by James Dove and collaborators gave ambiguous results in this regard \[275\]. The spectrum above \(\approx 10\) keV was extremely well fit by a pure power law with an exponential cutoff, without any signs of spectral curvature associated with Compton recoil photons from a reflection model. Extrapolating the exponentially cutoff power law to energies below 10 keV, however, yielded a strong, power law-like excess. This excess could have been produced by three effects: a mismatch in the calibration of spectral slopes between the two instruments that comprise RXTE (PCA and HEXTE; PCA typically yields softer fits than HEXTE to the power-law spectrum of the Crab nebula and pulsar \[276\]), reflection of a continuum spectrum more complicated than the presumed exponentially cutoff power
law, or a combination of an unmodeled soft excess and a strong, broad line. In retrospect, it was likely that all three effects were playing a role in these initial results. The same data, modeled with a sphere+disk coronal model (which included reflection of the Comptonized spectrum with an effective covering fraction of $\Omega/2\pi \approx 0.3$), fit the PCA and HETE spectra simultaneously; however, such models yielded strong, broad residuals in the iron line region [275]. At that time, these line-like residuals could not be attributed definitively to a relativistically broadened feature owing to the then poorly known RXTE calibration near energies of $\approx 6$ keV [275].

With further improvements to the calibration of the RXTE response in the 6 keV region of the spectrum, numerous fits to RXTE observations of Cyg X-1 have continued to indicate strong, broad line features. For one set of (predominantly hard state) RXTE observations of Cyg X-1, Marat Gilfanov and collaborators used a crude Gaussian convolution of a reflection model to simulate relativistic effects [277]. It was found that both the line (equivalent widths $\approx 70–160$ eV) and reflection features ($\Omega/2\pi \approx 0.3–0.7$) were well-described by a Gaussian smearing width $m_{ESC} \approx 0.3–0.9$ keV. Such large smearing widths would be consistent with relativistic motions and gravitational redshifts in the inner disk.

The EXOSAT, Ginga, and ASCA data were later re-examined by Christine Done and Piotr Zycki with a more complex set of reflection models [278]. Similar to prior analyses [270], these models included reflector ionization using the constant density assumption; however, they also included the effects of relativistic distortions. Furthermore, the models considered a non-relativistic (narrow) set of line/reflection features in conjunction with the relativistically smeared line and reflection features. It was found that all the datasets could be well-fit by models wherein the disk ionization was not very high, and the relativistically smeared line and edge features dominated ($\Omega/2\pi \approx 0.1–0.2$) over the non-relativistic features ($\Omega/2\pi < 0.05$) [278]. Andy Young and collaborators reanalyzed these datasets with an independent model and affirmed the description with a large degree of relativistic smearing of the line and reflection features; however, this latter model postulated a larger reflection fraction ($\Omega/2\pi \approx 1$), albeit with a higher degree of ionization [279].

Analyses of BeppoSAX observations of the Cyg X-1 hard state also were well described by broad line features [280,281]. BeppoSAX had extremely broad energy coverage from $\approx 0.5–200$ keV; therefore, it was capable of simultaneously constraining models of the distribution of seed photons for Comptonization (e.g., the disk with temperatures $kT \approx 0.1–2$ keV), the line/reflection region ($6–30$ keV), and the hard, Comptonized tail ($10–200$ keV). As for the re-analyses of EXOSAT, Ginga, and ASCA data [278], analyses of BeppoSAX by Frontera and collaborators revealed a relativistically smeared reflected power law ($\Omega/2\pi \approx 0.1–0.3$), with the degree of smearing being indicative of the reflection being dominated by the innermost regions of the accretion disk [280]. Similarly, sophisticated Comptonization models, including the effects of (non-relativistic) reflection ($\Omega/2\pi \approx 0.25$), also required the presence of a strong (equivalent width $\approx 350$ eV), relativistically broadened line [281].

One can see that there has been something of a split, with the prevailing wisdom prior to 1997 being dominated by the “narrow” features point of view, while the prevailing wisdom post-1997 seems to be dominated by the “broad”, i.e., relativistic, features point of view (with some researchers having been on both sides of the issue). Has this shift been one of sociology, improved theoretical modeling, improved observational data, or some other effects? Improved models have certainly played an important role. For example, in order to explain the ASCA data at the lowest energies, Ken Ebisawa and collaborators invoked a broken power law (with a softer, i.e., steeper, low energy
slope and a break energy at \( \approx 4 \text{ keV} \) in addition to the disk component with peak temperature of \( kT \approx 150 \text{ eV} \) [272]. The presence of such a low energy “soft excess” (above the requirements of the disk component) in the spectrum of Cyg X-1, as well as other hard state black hole spectra, continues to be debated [278,280,281]. It has been argued, however, that the apparent hardening at energies above 4 keV in fact might be due to the red tail of the relativistically broadened iron line [52]. The combination of simultaneous broad band data with models that attempt to (at least somewhat self consistently) model this entire range leaves less leeway to introduce phenomenological model components (e.g., a spectral break at 4 keV) that might remove evidence of a relativistically broadened line. It is worthwhile noting that Comptonization models [275,281] seem to require the broad line, even when including the other effects of reflection.

The above discussion highlights the complexity inherent in stellar mass black hole spectra. Unlike the case of AGN, it is clear that a simple, exponentially cutoff power law is insufficient to adequately describe the underlying X-ray continuum. This is especially true at soft energies, as the hottest GBHC accretion disks \((kT \approx 2 \text{ keV})\) can have significant spectral contributions into the red tail region of any broadened iron line. A further complication arises from the fact that stellar mass black hole disks likely represent a wide range of disk inclinations. It is typically thought that Seyfert 1 disks are viewed reasonably close to face-on (see Fig. 4); therefore, the gravitational redshift and transverse Doppler shifts typically cause the blue wing of a broadened K line to lie at energies below its associated ionization edge. This is not the case with potentially more highly inclined stellar mass black hole systems. There, especially if one misidentifies the continuum power law slope, it is much easier to mis-model the blue wing of such a line with a reflection edge.

Most importantly, due to their extremely short viscous and thermal time scales (compared to AGN), the environment of the accretion flow in a stellar mass black hole system can be very different from observation to observation. The existence of “state changes” clearly points toward this; however, even within a given state large changes in the spectral properties are observed [282,277,283, for example]. Thus, some of the seemingly different results quoted above are undoubtedly due to changes in the structure of the accretion flow between observations (e.g., changes in the ionization state, the disk inclination to our line of site, the disk/corona geometry, etc.). These changes, in fact, can be sometimes associated solely with the line profile. As an example, in Fig. 18 we show two RXTE-PCA observations of the stellar mass black hole candidate GX 339-4 in its hard state [52]. To minimize features due to uncertainties in the calibration and response of RXTE, the data are presented as a ratio to RXTE observations of the Crab nebula and pulsar (which are thought to possess a synchrotron hard X-ray spectrum well approximated by a featureless power-law). Even though these two observations represent nearly the same flux, power law slope, and (low) fitted reflection fraction, one exhibits a broad line residual, while the other one exhibits a line residual consistent with being narrow.

A final important example of spectral variability in stellar mass black hole systems is that due to orbital variation. A binary orbital dependence for the strength of the iron line has been previously reported for Cyg X-1, which was taken as evidence that the line was due to fluorescence by the secondary, not by the relativistic regions of the inner disk [269]. Recent Chandra observations, however, have dramatically demonstrated that the line region in fact is comprised of both narrow and broad components [284,285]. For the first time, instruments exist with high enough spectral resolution to definitively resolve the narrow component of the line; however, as shown by Jon Miller and collaborators, the Chandra sensitivity and energy coverage also convincingly reveal the broad component
Fig. 18. Ratio of RXTE observations of GX 339-4 to RXTE observations of the Crab nebula and pulsar, normalized to unity at 10 keV. The spectrum of the Crab nebula is widely believed to be a featureless power-law continuum. Showing the observations this way gives an indication of the intrinsic resolution of the detectors, as well as minimizes the worry that the line features are artifact of a poorly determined detector response. Note that despite the fact that these two observations have nearly identical fluxes and average spectral slopes, their line profiles are different.

of the line. In Fig. 19 we show the profile, comprised of both narrow and broad features, obtained with the Chandra observations performed by Miller et al. Further Chandra observations reveal that the narrow line component is not always present in the data (H. Marshall 2001, private comm., see also [286]). The narrow line has an equivalent width of only 80 eV, while the parameters of the broad line component (equivalent width $\approx 140$ eV) agree well with those previously found with, e.g., RXTE observations [284]. Thus, whereas the presence of a (variable) narrow line component complicates analyses performed with broader resolution instruments, it apparently does not obviate the need for a broad line to be present.

All of the above observations refer to spectrally hard states (often called “low” to “intermediate” states); however, Cyg X-1 occasionally transits into a much softer state (often called a “high state”). Although rarer, observations of this soft state have been performed with ASCA, RXTE, BeppoSAX [62,287,288,277,281], and more recently Chandra. All of the above caveats about determining line profiles in the hard state apply to the soft state as well. We further note that due to the increased strength and temperature of the soft “disk” component (with maximum temperature $kT \approx 300$ eV), continuum modeling near the red wing of the line region is even more difficult. Given these caveats,
however, several researchers have claimed increased reflection fractions ($\Omega/2\pi \approx 0.6–1.5$), and increased line widths ($\sigma \approx 1–2$ keV), for moderately strong lines (equivalent widths $\approx 80–190$ eV) \cite{288,281}. There have even been claims for evidence that the line energy increases to $\approx 6.7$ keV, i.e., consistent with heavily ionized Fe, in the soft state \cite{62}.

### 7.2. Other GBHC and spectral correlations

Evidence continues to accumulate that many GBHC systems exhibit broad Fe lines, both in their hard and soft spectral states. As discussed above, and as shown in Fig. 18, the hard state of GX 339-4 can show a broadened line with a substantial red wing. Additionally, during a transition from its soft state into its hard state, GX 339-4 exhibited a broad line that was as skewed redward as the line observed in MCG–6-30-15 during its DM state \cite{52}. Similarly strong, broad, very redward-skewed lines have been detected by Jörn Wilms and collaborators in RXTE observations of the soft states of LMC X-1 and LMC X-3; however, as these objects are somewhat faint (and given some of the systematic uncertainties described above) these authors could not unambiguously describe the line properties \cite{66,290}. Observing the somewhat brighter soft state of GRS 1650–500 with XMM-Newton (which has lower background and greater spectral resolution than RXTE), Jon Miller and collaborators found very good evidence for an iron line qualitatively and quantitatively (in terms of physical width, equivalent width, emissivity index, etc.) similar to the DM line of MCG–6-30-15 \cite{289}. This line was also taken as evidence for spin-energy extraction from a rotating black hole. In Fig. 20, we show further observations by Miller and collaborators, performed simultaneously with Chandra and RXTE (J. Miller, private comm.), which yield independent evidence for such a strong, broad line. If, as discussed previously, soft state spectra represent disks extending down to, or even within, their marginally stable orbit, barring the optical-thick material being completely ionized, one naturally expects to find the most broadened lines within these states.
Fig. 20. Simultaneous Chandra (rebinned to show broad features) and RXTE observation of GRS 1650−500 (courtesy J. Miller, see also [289]) fit with a power-law and a soft disk component. The residuals of this fit show a broad component in the Fe line region, and an upturn at high energy, possibly due to reflection.

Broadened iron lines have been detected with BeppoSAX observations of V4641 Sgr [291] and GRS 1915+105 [219]. Both of these sources have exhibited strong radio “jet ejection events” with apparent super-luminal motion and inferred velocities of 0.9c [292,77]. As radio and X-ray studies continue to explore the physics of jets in these sources, and as X-ray spectroscopy continues to probe the innermost regions of the accretion flow near the event horizon, there is the hope that these observations can be combined to reveal answers to such fundamental physical issues such as the connection between black hole spin and jet production. In this regard, GBHCs provide a unique advantage over AGN systems. In a number of these systems, X-ray variability reveals high-frequency, stable quasi-periodic oscillations (QPO), such as the 67 Hz oscillation in GRS 1915+105 which is stable in frequency over a factor of a few in source flux [293], or the 450 Hz QPO observed in GRO J1655−40 [294]. Although there is great debate as to the correct model for these oscillations, there is a general consensus that the high, stable frequencies point towards a direct relationship to the gravity of the black hole in the “strong field” regime of GR. The oscillations therefore may act as independent probes of black hole mass and spin [295–299]. It is hoped that models of variability can be combined with models of broadened lines to yield even tighter constraints on black hole mass and spin.

To what extent are GBHC and SMBH line and reflection profiles the same? Can observations of one class lead to inferences about the other? Given that the two systems complement each other in terms of spectrally resolvable time scales (i.e., the dynamical time scales of AGN are readily observed, while multiple viscous time scales can be observed in GBHC), commonalities between the two could prove very powerful in constraining the underlying physical mechanisms. Recently, Andrzej Zdziarski and collaborators have claimed such a correspondence between the two classes. Specifically, they describe a correlation between the spectral slope of a fitted power law with the amplitude of reflection, with softer spectra yielding greater reflection fractions [282]. The observations that made up this correlation consisted of both AGN and GBHC, mostly observed in relatively hard spectral states. (Ueda and collaborators had previously described such a correlation with four Ginga
observations of GX 339-4 [300].) Considering solely a sample of Seyfert galaxies, Piotr Lubinski and Zdziarski claimed that the correlation could also be carried over to fits of the line region as well, with Seyfert lines being composed of a fixed equivalent width, narrow core line, plus a broad, variable line whose equivalent width increased with spectral softening of the continuum [301]. As discussed previously for the case of NGC 5548, the correlations between line strength and spectra can be quite complex, with that particular case showing a decreasing line equivalent width (i.e., a constant line flux) with a softening and brightening of the spectrum.

Further questions about this possible correlation arose over two issues. First, possible systematic correlations exist between the fitted power laws and reflection fraction. An increased reflection fraction phenomenologically hardens a soft power law, therefore, error bars for the two parameters naturally show such a correlation for a fixed spectrum. Second, the effect had been shown convincingly only by combining a large number of observations of different objects. Both of these objections have been alleviated to some extent by considering multiple observations of single objects that represent a wide range of flux and spectral slopes. Furthermore, these observations have been fit with a variety of models (that bear in mind possible systematic errors), all of which yield the basic correlation [277,302,52]. The ability to observe GBHC over a very wide variety of accretion rate states has proven crucial for such studies.

One suggestion for how the correlation is produced was described by Andrzej Zdziarski and collaborators (elaborating upon earlier suggestions of Juri Poutanen and collaborators [303]), and it bears upon the coronal models described in Section 3.3. It was suggested that there is a certain amount of overlap between the quasi-spherical corona and the outer, thin accretion disk. As this disk moves inward, the corona becomes more effectively Compton cooled and produces softer spectra, while the increase of overlap between disk and corona yields stronger lines and reflection features [282]. As for the case of NGC 5548, however, the details of the overall correlation are more complicated than any simple model. Analyzing RXTE observations of Cyg X-1 and GX 339-4, Marat Gilfanov, Michael Revnivtsev, and Eugene Churazov used a crude Gaussian convolution of a reflection model (to mimic relativistic effects) to deduce a very strong correlation between spectral slope and amplitude of reflection [277,303]. Furthermore, they found that the width of the Gaussian smearing (or equivalently, the amplitude of the relativistic distortions) increased as the spectrum softened and the reflection amplitude increased. This would be consistent with the corona-disk overlap model described above.

Other analyses of the same observations of GX 339-4, however, yielded different results [52]. Instead of fitting a power law, a more realistic Comptonized spectrum was employed (hardness was instead described by a “coronal compactness parameter”, $\ell_c$, with larger compactnesses corresponding to harder spectra). As shown in Fig. 21, although the overall correlation was found, it was not as strong as described by the models of Revnivtsev et al. [302]. Furthermore, a slightly better correlation was found between amplitude of reflection and soft X-ray flux [52, and Fig. 21]. The reflection parameters were also found to be fairly “flat” in a regime of soft X-ray flux corresponding to hysteresis in the source. This regime corresponds to fluxes where the source appears spectrally hard if it had already been spectrally hard, and the source appears soft if it had already been spectrally soft [304]. Additionally, although a weak correlation between the width of the line and amplitude of reflection was found, this correlation was weakened further if ionized reflection models were instead considered [52].

Clearly, the exact correlations between broad continuum spectrum, reflection, and line properties, are extremely complicated, both physically and observationally. The existence of such correlations,
however, are now more widely accepted as being both real and very important. Our understanding of these correlations are likely to grow as both our theoretical models improve (especially as regards our ability to describe highly ionized atmospheres) and as the observations continue to be refined with ever more sensitive instruments.

8. Future prospects for iron line studies

This is a technology-driven field whose advances will parallel those in astrophysical X-ray instrumentation. While we will avoid the temptation for lengthy speculation of what we will find, it is important to say a few words about some of the science issues that will be opened up by further observations.

For the next few years, most of the progress is likely to be made by the continuing operation of XMM-Newton and Chandra (with these instruments likely often being assisted by simultaneous RXTE observations to constrain the high energy spectra). The accumulated body of data from these observatories will allow a detailed categorization of the X-ray reflection properties of all classes of accreting black hole in the nearby universe. By applying detailed considerations similar to those described in Section 6.1, such data will allow the first observational investigation of black hole spin in a sample of both GBHCs and AGN. Investigation of variability in the iron line profile or strength can also be studied in more detail than ever possible before. In the case of AGN, line variability can be studied on time scales down to $\sim 10$ orbital periods (at which point the spectra become too photon-starved to provide useful constraints)—line variability seen on these time scales most likely corresponds to spatial changes in the X-ray emission from the disk corona. These observatories will also provide the first constraints on the astrophysical environment of accreting SMBHs at very high redshift.
After launch in 2005, Astro-E2 will complement XMM-Newton and Chandra. As well as being able to study broad iron lines at high signal-to-noise even when observing alone, its ability to provide very high resolution spectra in the iron line band (thereby completely characterizing any low-velocity components to the iron line) will be a powerful complement to any Chandra and XMM-Newton observation. Even with high-quality data in the iron line band, good constraints on ionized disks greatly benefit from having simultaneous hard X-ray (> 10 keV) observations of the Compton reflection continuum. At present, only the Proportional Counter Array (PCA) on RXTE can obtain such data. In the near future, our ability to constrain the continuum shape above 15 keV will be greatly enhanced by the recent successful launch of ESA’s Integral (October 2002) and the future launch of NASA’s Swift (September 2003).

At the end of this decade or the beginning of the next, it is hoped that the Constellation-X and XEUS observatories will come on-line. With their very large collecting areas, they will represent a major leap in sensitivity. These observatories will make at least two major advances in AGN research. Firstly, we will be able to produce good X-ray spectra of the highest redshift AGN currently known. This opens the real prospect of being able to constrain the cosmic history of SMBH properties as well as the chemical history of the matter in the vicinity of the SMBH. Secondly, we will be able to study spectral variability on time scales of the order of the light crossing time of the black hole. By searching for detailed changes in the iron line profile, we will be able to search for the “reverberation” of X-ray flares, i.e., the X-ray light echo that propagates across the accretion disk due to the finite speed of light [305]. These reverberations signatures encode detailed information about the space-time geometry, and might allow for a quantitative test of General Relativity in the very strong field limit [306,309]. The proposed timing capabilities of XEUS could also make a major impact in our understanding of GBHCs. In particular, we will be able to study the temporal power-spectrum of the X-ray variability in exquisite detail. The ultimate goal is to use the form of these power spectra and, in particular, the families of quasi-periodic oscillations that are known to exist [307,294,308], in order to study accretion disk physics and the space–time geometry.

On a longer time scale (maybe two decades), we look forward to ultra-high resolution imaging of black hole systems. The MAXIM concept (Section 5.3.4) will allow us to image nearby accreting SMBHs (e.g., the SMBHs in our Galactic Center, M87 or nearby Seyfert galaxies) on spatial scales comparable to, or smaller than, that of the event horizon. In some versions of this concept, one will also be able to map out the spectrum of the system (including disk reflection features) across the image. This will clearly lead to major advances—the geometry and physical state of the disk, corona and jet can be inferred directly from such an image. Combining the spectral information, the dynamics of the relativistic disk can be measured and compared with that expected from material orbiting in a Kerr metric. In particularly clean systems (e.g., Seyfert galaxies) these measurements could be used as a test of the Kerr metric itself—such a study would be complementary to the Kerr-testing experiments that will be possible using gravitational wave signatures with the Laser Interferometry Space Antenna (LISA).

9. Conclusions

The past few years have seen a shift in the mind-set of many observationally oriented black hole researchers. A small number of important observations (the mass of the compact object in the binary
star system Cyg X-1, stellar motions in our Galactic Center, the gas disk in M87, and the maser
disk in NGC 4258) have led most astronomers to the conclusions that black holes do indeed exist
beyond any reasonable doubt. With this established, the interest has shifted to the demographics of
black holes in the Universe, and the detailed astrophysics of black hole systems.

As we have described, X-ray spectroscopy currently provides the best understood method of
exploring the astrophysical environment in the immediate vicinity of an accreting black hole. The
accretion disk is the engine that drives all observable phenomena from accreting black hole systems.
In many systems, the surface layers of the accretion disk are expected to produce X-ray spectral
features (so-called X-ray reflection signatures) in response to external hard X-ray illumination from
a disk-corona—the most prominent spectral feature is often the fluorescent K\textsubscript{α} emission line of iron.
This line, which has an intrinsically small energy width, will be dramatically broadened and skewed
by the rapid orbital motions of the accreting material and strong gravity of the black hole. These
relativistic spectral features can be identified unambiguously in many of the AGN for which X-ray
data of sufficient quality exist. Similar spectral features can also be found in many AGN and GBHCs
for which poorer quality data exist—while it may not be possible to rule out other forms of spectral
complexity in these objects, relativistic disk features are often the most compelling and physical of
the possibilities. Given how generic these features are, they provide a powerful way to study the
near environment of accreting black holes.

Moderate luminosity radio-quiet AGN, the Seyfert galaxies, present the cleanest examples of
these spectral features. Analysis of X-ray data from ASCA and, more recently, XMM-Newton finds
evidence for a rather cold accretion disk extending all of the way down to the radius of marginal
stability around the black hole. In one case (MCG–6-30-15), the data argue strongly that the black
hole is rapidly rotating. Furthermore, there may be suggestions that the accretion disk is torqued
by processes associated with the spinning black hole and, in particular, may be tapping into the
rotational energy of the spinning black hole. The case of MCG–6-30-15, which we discussed in
some detail, illustrates the exotic black hole physics that can be addressed via these techniques.

The difference in iron line properties between Seyfert galaxies and other types of AGN allows us
to probe the physical state of the accretion flow as a function of, in particular, the mass accretion
rate. The fact that higher-luminosity AGN display weaker spectral signatures is very likely due to
increasing ionization of the disk surface as the accretion rate (relative to that rate which would
produce the Eddington luminosity) is increased. The weakness of the X-ray reflection displayed by
radio-loud AGN and low-luminosity AGN is much less certain (primarily due to the paucity of high
quality datasets). Possibilities include ionization of the disk surface, dilution of the disk spectrum
by X-ray emission from a relativistic jet, and/or the transition of the disk into a radiatively ineffi-
cient state (which would be extremely hot and optically thin thereby producing no X-ray reflection
features).

Although it is somewhat counter-intuitive given that they can be three orders of magnitude brighter
than AGN, it has been much more difficult to study relativistic disk signatures in GBHCs. In addi-
tion to the fact that the X-ray spectra of GBHC are inherently more complex (since the X-ray band
contains the thermal disk emission and the disk is almost always strongly ionized), the high X-ray
fluxes from these sources can readily overwhelm sensitive photon counting spectrometers. How-
ever, with Chandra and XMM-Newton we can now study relativistic X-ray reflection in GBHCs
free of these instrumental constraints, at medium-to-high spectral resolution, and high signal-to-
oise. Consequently, relativistic iron lines have been found in several GBHCs with a variety of
“spectral states”. With further study of these features over the coming few years, we should be able to constrain the gross geometry of the accretion flow as a function of spectral state.

The future is bright. Until just a few years ago, the physics of relativistic accretion disks and the astrophysics of black hole spin was firmly in the realm of pure theory. Modern X-ray spectroscopy has given us a powerful tool with which we can peer into this exotic world. There is still much to be learnt—much of the physics determining the properties of accretion disks, the demographics of black hole spin (which is closely related to black hole formation and history), and the prevalence of spin-energy extraction (which is of fundamental interest) are just a few of the open questions. Continued investigation with increasingly capable instruments promise to answer these, and many more, questions.

Acknowledgements

The authors thank Philip Armitage, Mitchell Begelman, Andy Fabian, David Garofalo, Kazushi Iwasawa, Julia Lee, Herman Marshall, Cole Miller, Jon Miller, Norbert Schulz, Joern Wilms and Andrew Young for their careful reading, and insightful comments, of various drafts of this article. We are particularly grateful to the referee, Eric Agol, whose report led to a substantial improvement of the manuscript. In addition, we also express our appreciation to Tomaso Belloni, Omer Blaes, Jim Chiang, Paolo Coppi, Stephane Corbel, James Dove, Rob Fender, Julian Krolik, Tom Maccarone, Phil Maloney, and Katja Pottschmidt for stimulating discussions and help in preparing some of the figures for this article. We acknowledge and thank NASA for financial support under grants NAG5-10083 (CSR) and SV-61010 (MAN).

References


