Current High-Energy Emission around Black Holes

Proceedings of the 2nd KIAS Astrophysics Workshop

editors
Chang-Hwan Lee
Heon-Young Chang

World Scientific

Current High-Energy Emission around Black Holes

This page is intentionally left blank

Current High-Energy Emission around Black Holes

Proceedings of the 2nd KIAS Astrophysics Workshop

Korea Institute for Advanced Study September 3–8, 2001

editors

Chang-Hwan Lee

Heon-Young Chang

Korea Institute for Advanced Study, Korea

World Scientific

New Jersey • London • Singapore • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

CURRENT HIGH-ENERGY EMISSION AROUND BLACK HOLES
Proceedings of the 2nd KIAS Astrophysics Workshop

ISBN 981-02-4870-9

Printed in Singapore by World Scientific Printers (S) Pte Ltd

Preface

There have recently been intense research activities in black-hole physics among the members of the Astronomy and Physics community in Korea and the 2nd KIAS Astrophysics Workshop aimed at providing a chance for them to share their interests with local as well as foreign research scientists.

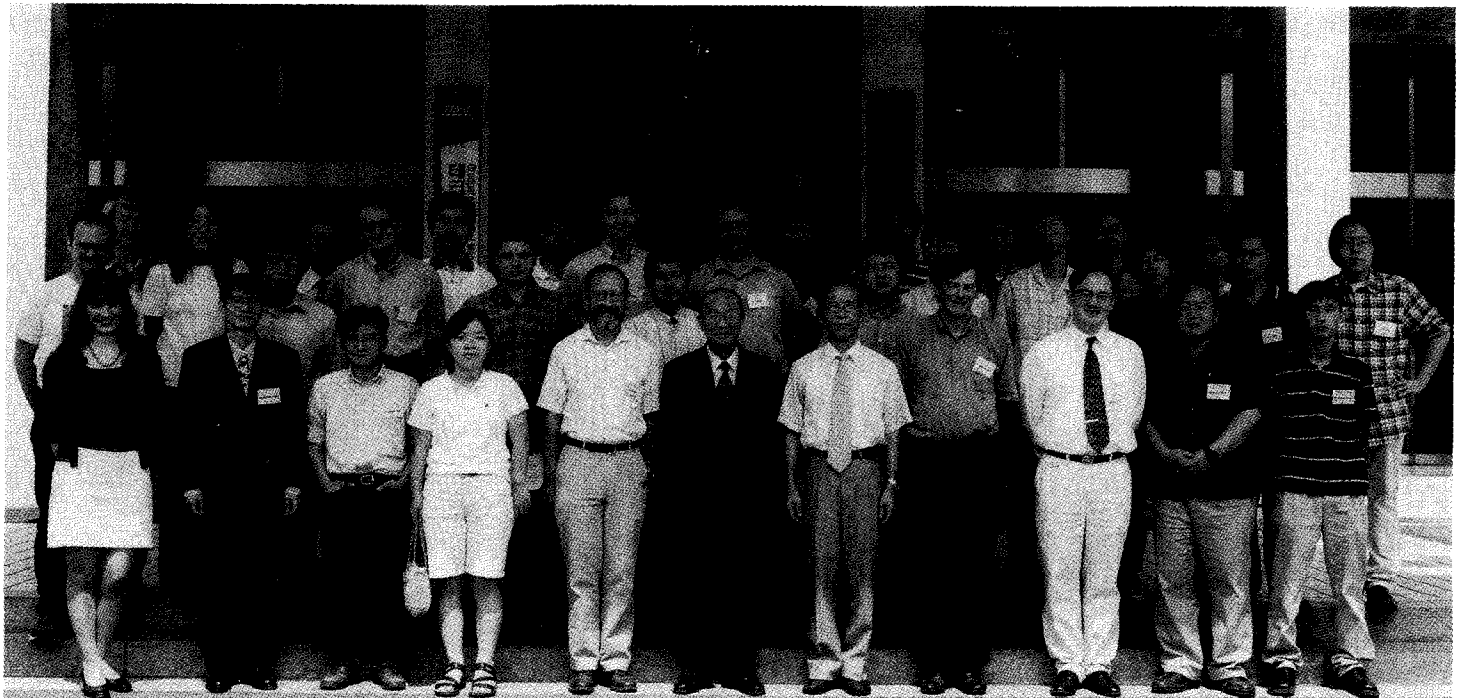
In our workshop poster, we put an artist's conception of a black hole in the middle of a famous Korean folk painting, "Dancing Child," painted by Hong-Do Kim, an 18th century Korean artist. This design exposes our desire that the participants enjoy "dancing around black holes." We believe that all the speakers actively participated in the "dancing," and the audience enjoyed the "performances." A non-scientific lecture given by Dr. Yong-Ok Kim, one of the most celebrated philosophers in Korea, was a special event highlighting the spirit, not only for the foreign attendants but also for the local ones.

One of the most important aspects of holding such a workshop in KIAS was not only to introduce Korean science as it is to the foreign participants, but also to provide opportunities for the foreign participants to be exposed to Korean culture. For this, we were generously helped by Dr. Sang-Hyeon Ahn who organized a city cultural tour and also Mr. Yo Chang Jung who made sure that we arrived in time at the Chang-Deok-Gung palace. The success of these efforts was evident in the article by Prof. Blandford, who discussed "Ondol Dynamics," named after the Korean traditional heating system, and "Seungmu Dynamics," named after a Korean folk dance.

In this volume, various aspects of high-energy emissions around black holes, ranging from observations to pure theory of general relativity, are discussed. We hope that this volume is helpful for workers in the field in general and for the young scientists who are interested in black hole astrophysics in particular. We also hope that this volume provides imaginary fields for the reader to "play around black holes."

We would like to thank all the members of the organizing committee, the speakers, and the participants. We would also like to thank the KIAS staff members, especially Ms. Jung Eun Lee, Ms. Hyung Soon Lee, and Mr. Si-Hyoung Oh for their generous support. Our special thanks go to Prof. Hyun Kyu Lee for his advice and help.

Chang-Hwan Lee



2nd KIAS Astrophysics Workshop

Current High-Energy Emission around Black Holes

Sep. 3-8, 2001 KIAS International Conference Hall, Seoul, Korea

Contents

PART 1

Black Hole Observations

This page is intentionally left blank

BLACK HOLE DEMOGRAPHICS

LAURA FERRARESE
Rutgers, the State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA
E-mail: lff@physics.rutgers.edu

1 Introduction

It is frequently the case that revolutionary scientific ideas are first proposed and then remain dormant for years, or sometimes decades, before their importance is truly appreciated. Often, the reawakening of interest is driven by new technological developments. Such was, for instance, the case of supermassive black holes (SBHs) in galactic nuclei. By the mid 1960s, just a few years after the discovery of QSOs, it was generally recognized that their energy source must be gravitational in nature. Yet for the following three decades the existence of SBHs was destined to be surrounded by skepticism. By the mid 1990s, a few years after the launch of the Hubble Space Telescope, it was widely accepted. Today, it is generally agreed upon that SBHs play a fundamental role in the formation and evolution of their host galaxies. Freed from the burden of having to demonstrate the very existence of supermassive black holes, we can now begin asking more fundamental questions: how are black holes related to their host galaxies, how did they form, how do they accrete, how do they evolve, and what role do they play in the formation of cosmic structure?

The purpose of this contribution is to review the current status of black hole demographics. I will not address the various techniques that are used to measure black hole masses: excellent discussions can be found in the recent literature (e.g. Kormendy & Richstone 1995; Ho 1999). Neither will I discuss the somewhat tumultuous events that lead to a critical reassessment of the "Magorrian relation" (Magorrian *et al.* 1998) and its analog for local AGNs (Wandel 1999) since a full description of such events can be found elsewhere (Merritt & Ferrarese 2001c). Instead, I will revisit the issue of black hole demographics in light of recent advances in the study of high redshift QSOs (section 2), local AGNs (section 3) and local quiescent galaxies (section 4). I will then outline the prospects for future progress (section 5), and discuss what I believe will be the challenges for the years to come.

2 Black Hole Demographics: High Redshift QSOs

The existence of SBHs in the nuclei of nearby galaxies has gained popular consensus only in recent years. That supermassive black holes must power QSO activity has, however, been widely suspected since the mid 1960s (e.g. Robinson *et al.* 1965). It is therefore not surprising that the first studies of black hole demographics were conducted, over two decades ago, using optical counts of high redshift QSOs. In a seminal paper entitled "Masses of Quasars", Andrzej Soltan (1982) proposed a simple argument: QSO optical number counts yield a QSO luminosity function which can be integrated to give a mean comoving energy density in QSO light. After applying the appropriate bolometric corrections and assuming a reasonable conversion factor of mass into energy, Soltan concluded that the SBHs powering high redshift ($z > 0.3$) QSOs comprise a total mass density of $\sim 5 \times 10^4$ M$_\odot$ Mpc^{-3}, each SBH having a mass of $10^8 - 10^9$ M$_\odot$. Soltan's arguments, which have been employed many times in the following years (Chokshi & Turner 1992; Small & Blandford 1992; Salucci *et al.* 1998), lead to the inescapable conclusion that most, if not all, nearby galaxies must host dormant black holes in their nuclei. This finding has been the main driver for SBH searches in nearby quiescent galaxies and has kindled the interest in the accretion crisis in nearby galactic nuclei (Fabian & Canizares 1988), ultimately leading to the revival of accretion mechanisms with low radiative efficiencies (Rees *et al.* 1982, Narayan & Yi 1995).

Armed with recent measurements of the QSO luminosity function from the 2dF QSO Survey ($0.3 < z < 2.3$, Boyle *et al.* 2000) and the Sloan Digital Sky Survey ($3.0 < z < 5.0$, Fan *et al.* 2001), we are in a position to update Soltan's results. If $\Phi(L, z)$ is the QSO luminosity function, the cumulative mass density in SBHs which power QSO activity can be expressed as:

$$\rho_{QSO}(> M) = \frac{K_{bol}}{\epsilon c^2} \int_0^\infty \int_L^\infty \frac{L' \Phi(L', z)}{H_0 (1+z) \sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}} \mathrm{d}L' \mathrm{d}z \quad (1)$$

where the mass accretion rate is simply $= K_{bol} L \epsilon^{-1} c^{-2}$, with K_{bol} the bolometric correction (from Elvis *et al.* 1986), and ϵ the energy conversion coefficient (assumed equal to 0.1). An $\Omega_\Lambda = 0.0$, $\Omega_m = 1.0$, $H_0 = 75$ km s^{-1} Mpc^{-1} cosmology is assumed for consistency in comparing the results with those derived in the following sections. The cumulative mass density due to QSO accretion is shown in Fig. 1. It should be noted that the magnitude limits of the 2dF and Sloan QSO surveys correspond to Eddington limits on the SBHs masses of 4.5×10^7 $M_\odot$ and 7.3×10^8 $M_\odot$ respectively. Cumulative mass densities down to 10^6 $M_\odot$ are calculated on the (unverified) assumption that the QSO luminosity

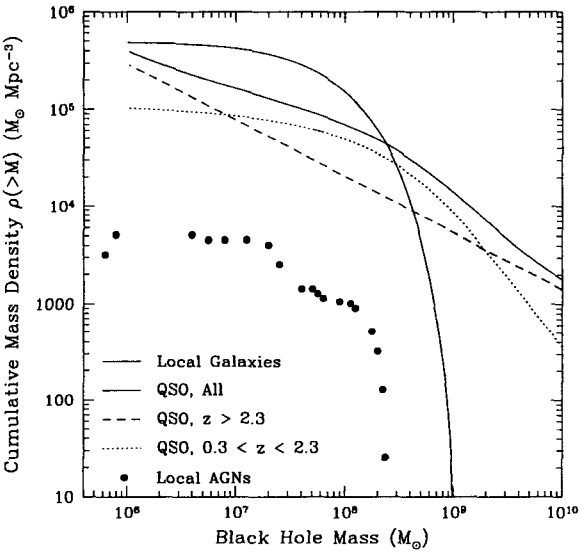


Figure 1: Comparison between the black hole mass function in high redshift QSOs (blue lines - dotted line: 2dF sample; dashed line: SDSS sample; solid line: both samples combined); local AGNs (red) and local quiescent galaxies (black, corresponding to the dotted black line in Fig. 4).

function holds at the corresponding magnitude ($B \sim -19$). Furthermore, the lower redshift limit of integration for the SDSS luminosity function was pushed down to the high redshift boundary of the 2dF survey ($z = 2.3$), although there are no QSO luminosity functions covering the $2.3 < z < 3.0$ range. For masses larger than 10^8 $M_\odot$, the extrapolation from $z = 3$ to $z = 2.3$ of the spatial density (e.g. Fig. 3 of Fan *et al.* 2001) or mass density (Fig. 7 in these proceedings) as a function of redshift from the SDSS joins rather smoothly the curve derived from the 2dF survey, therefore our assumption is likely justified. However, for smaller masses or luminosities, the SDSS mass density, extrapolated to $z \sim 2.3$, overpredicts the QSO mass density (per unit redshift) derived from the 2dF data by an order of magnitude. Thus, it is likely that the linear rise of the SBH cumulative mass density for the high redshift QSOs between 10^8 and 10^6 $M_\odot$ represents an *upper bound* to the real curve, which could have been overestimated by a factor of a few (i.e., up to $\sim$ three).

In short, the cumulative mass density from the optical QSO counts due to accretion onto high redshift QSOs ($0.3 < z < 5.0$) appears to be in the range $(2 - 4) \times 10^5$ $M_\odot$ Mpc^{-3}. Notice that this estimate does not account for the

possibility that sizable black holes might have already been in place *before* the optically bright phase of QSOs. Furthermore, I have neglected the contribution to the SBH mass density from the so called "obscured" or "Type II" QSOs, the existence of which is required to explain the observed properties of the X-ray background. In analogy with local Seyfert 1 and Seyfert 2 galaxies, in Type II QSOs molecular material, with column density in the neighborhood of 10^{23} cm^{-2}, completely hides the nucleus from view at optical wavelengths (e.g. Fabian & Iwasawa 1999). The contribution of Type II QSOs could be significant. For instance, Barger *et al.* (2001) calculate lower and upper limits of 6×10^4 and 9×10^5 $M_\odot$ Mpc^{-3} for the mass density in the SBHs which comprise the X-ray background. Gilli, Salvati & Hasinger (2001) find that the spectral shape of the hard (2−10 Kev) X-ray background can be best explained if obscured AGNs evolve more rapidly as a function of redshift than do their unobscured counterparts. Their model assumes a ratio between absorbed and unabsorbed AGNs increasing from ~ 4 in the local universe to ~ 10 at $z \sim 1.3$, and remaining constant at higher redshifts. Such a model, if correct, would translate into an increase by nearly a factor of 10 in the SBH cumulative mass density derived above.

3 Black Hole Demographics: Local AGNs

In the supermassive black hole business, the masses which are most challenging to measure are those of the black holes powering local AGNs. Compared to QSOs, lower-luminosity AGNs have a small ratio between nuclear non-thermal and stellar luminosity, making it difficult to assess what fraction of the total luminosity is due to accretion onto the central black hole. Furthermore, the history of past activity is not known, so it is not obvious what fraction of the SBH mass, $M_\bullet$, predated the onset of the present nuclear activity (and the assumption that AGNs radiate at the Eddington limit is not justifiable). In other words, Soltan's arguments, which hold rather nicely for high redshift powerful QSOs, are not applicable to their less flamboyant, nearby cousins.

To make matters worse, the techniques that allow us to detect supermassive black holes in quiescent galaxies are seldom applicable to the hosts of AGNs. In Seyfert 1 galaxies, and in the handful of QSOs for which traditional dynamical studies of the gas or stellar kinematics can be performed, the presence of the bright non-thermal nucleus (e.g. Malkan, Gorjian & Tam 1998) overwhelms the very spectral features which are necessary for dynamical studies. The only Seyfert galaxy in which a SBH has been detected by spatially-resolved kinematics is NGC 4258, which is blessed with the presence of an orderly water maser disk (Watson & Wallin 1994; Greenhill *et al.* 1995;

Miyoshi *et al.* 1995). The radius of influence of the black hole at its center, $\sim 0''.15$, can barely be resolved by the Hubble Space Telescope (HST) but is fully sampled by the VLBA at 22.2 GHz. Unfortunately, water masers are rare (Greenhill *et al.* 2002) and of the handful that are known, only in NGC 4258 are the maser clouds distributed in a simple geometrical configuration that exhibits clear Keplerian motion around the central source (Braatz *et al.* 1996; Greenhill *et al.* 1997, 1996; Greenhill, Moran & Hernquist 1997; Trotter *et al.* 1998). A study of black hole demographics in AGNs must therefore proceed through alternative techniques.

To my knowledge, the only attempt at deriving a mass function for local ($z < 0.1$) AGNs was published by Padovani *et al.* (1990) using the CfA magnitude limited sample of Seyfert 1 galaxies. For each galaxy, the mass of the central SBH was derived from the dynamics of the broad line region (BLR) under the virial approximation, $M_\bullet = v^2 r/G$. The radius r of the BLR was calculated by assuming that the ionization parameter U (or more precisely, its product with the electron density) is known and invariant from object to object. U depends on the inverse square of r, and linearly on the number of ionizing photons; the latter quantity can be derived if the spectral energy distribution of each object is known. As noted by Padovani *et al.* , however, the ionization parameter is likely *not* invariant. More recently, Wandel, Peterson & Malkan (1999) have calibrated the "photoionization method" against the more sophisticated technique which has become known as "reverberation mapping" (Blandford & McKee 1892; Peterson 1993; Netzer & Peterson 1997; Koraktar & Gaskell 1991). The latter method relies on the fact that if the non-thermal nuclear continuum is variable, then the responsivity-weighted radius r of the BLR is measured by the light-travel time delay between emission and continuum variations. As in the case of the photoionization method, the mass of the central black hole follows from the virial approximation, *if* the BLR is gravitationally bound. The latter assumption has now received strong support in a few well-studied cases (Koratkar & Gaskell 1991; Wandel, Peterson & Malkan 1999; Peterson & Wandel 2000, but see also Krolik 2001).

Wandel, Peterson & Malkan (1999) concluded that photoionization techniques and reverberation-mapping estimates of the BLR sizes (or central masses) compare well, but only in a *statistical* sense. In other words, $M_\bullet$ estimates from the two methods can differ by up to an order of magnitude for individual objects, yet there does exist a reasonably good linear correlation between the two quantities when large samples are investigated, which bodes well for the Padovani *et al.* analysis. Eight of the Seyfert 1 galaxies in the Padovani *et al.* sample also have reverberation-mapping masses (Wandel *et al.* 1999; Kaspi *et al.* 2000). For these galaxies, the reverberation masses are a factor 3.6±3.4

larger (in the mean) than the photoionization masses. In Fig. 1, the cumulative mass function in local Seyfert 1 galaxies derived by Padovani *et al.* , once corrected for this factor and scaled to $H_0 = 75$ km s^{-1} Mpc^{-1}, is compared to the mass function in QSO black holes.

The total density of SBHs in Seyfert 1 galaxies is $\sim 5000\ M_\odot$ Mpc^{-3}. Despite the upward revision by a factor $\sim$ eight compared to the original estimate, the main conclusion reached by Padovani *et al.* still holds: "the bulk of the mass related to the accretion processes connected with past QSO activity does not reside in Seyfert 1 nuclei. Instead, the remnants of past activity must be present in a much larger number of galaxies". In the local universe, the ratio of Seyfert 2 to Seyfert 1 galaxies is $\sim$ four (Maiolino & Rieke 1995), while LINERs are a factor of a few more numerous than Seyferts (Vila-Vilaro 2000). Yet even after correcting the mass density given above to include these classes of AGNs, the total cumulative mass density in local AGNs falls a factor of several below that estimated for high redshift QSOs. Thus, in the search for SBHs, powerful QSOs and completely quiescent galaxies appear to be equally promising targets.

4 Black Hole Demographics: Local Quiescent Galaxies

4.1 Supermassive Black Holes and their Host Galaxies

Measuring SBH masses in the nuclei of "normal" nearby galaxies has been a staple of the astronomical literature since the late 1970s. It all started in 1978, when Wallace Sargent and collaborators published an investigation of the nuclear dynamics of the Virgo cluster cD, M87 (Sargent *et al.* 1978), claiming the detection of a five billion solar mass black hole. Other famous detections followed: M32 in the mid '80s (Tonry 1984), M31 a few years later (Kormendy 1988). Each claim, however, seemed to have its detractors, beginning with Binney & Mamon (1982) who dismantled Sargent's M87 black hole detection and alerted the community to the perils of the now familiar "mass-to-light ratio – velocity anisotropy degeneracy".

Important ground-based work on SBHs continued through the '90s (Richstone, Bower & Dressler 1990; Kormendy *et al.* 1996a, 1996b; Magorrian *et al.* 1998), producing a series of tantalizing but frustratingly inconclusive results. Indeed, it was not until the launch of HST that dramatic progress was made. It was HST data that firmly established the existence of a SBH in M87 (Harms *et al.* 1994), thereby ending a two-decade controversy. Since then, SBH masses based on HST/FOS and STIS data have been published for ten additional galaxies (Ferrarese, Ford & Jaffe 1996; Bower *et al.* 1998; van der Marel & van den Bosch 1998; Ferrarese & Ford 1999; Emsellem *et al.* 1999; Cretton & van

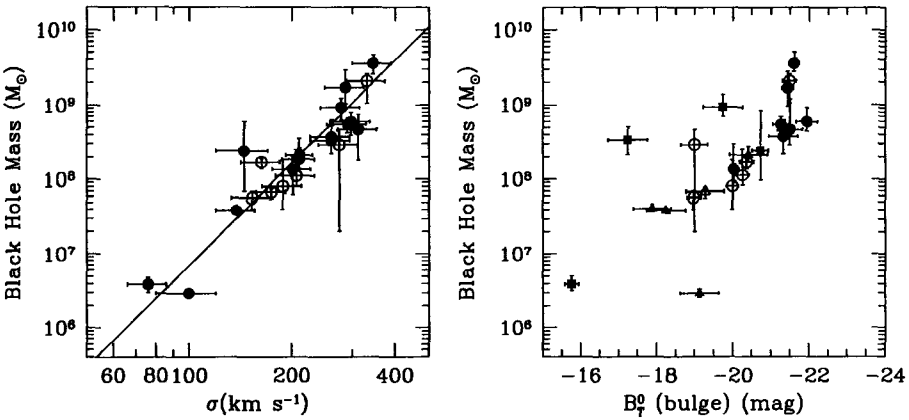


Figure 2: (left) Correlation between central velocity dispersion and black hole mass for all secure SBH detections. Published data are shown as solid symbols, data based on unpublished analyses as open symbols.
Figure 3: (right) Correlation between bulge $B-$band magnitude and black hole mass for the same sample shown in Fig. 2. Elliptical galaxies are shown as circles, lenticulars and compact ellipticals as squares, and spirals as triangles

den Bosh 1999; Verdoes Kleijn *et al.* 2000; Gebhardt *et al.* 2000; Joseph *et al.* 2001; Barth *et al.* 2001; Sarzi *et al.* 2001).

The success of HST can be ascribed to the fact that its unprecedented spatial resolution (in the optical regime, at least!) makes it possible to resolve, in favorable cases, the region of space within which the SBH's gravitational potential dominates that of the surrounding stars, i.e., the "SBH sphere of influence". This is more crucial than might at first be realized. It has become obvious (Ferrarese & Merritt 2000) that resolving the sphere of influence does not simply aid the SBH detection: it is a *necessary* condition for a detection to be made. Ground-based observations generally lack the spatial resolution necessary to penetrate the SBH sphere of influence, and this condition inevitably leads to spurious detections and overestimated masses (see Merritt & Ferrarese 2001c for a more thorough discussion of this issue). To date, with a few notable exceptions (the Milky Way, Genzel *et al.* 2000, Ghez *et al.* 2000; NGC 4258, Miyoshi *et al.* 1995; NGC 5128, Marconi *et al.* 2001), all firm SBH detections — *detections based on data which resolve the SBH sphere of influence* — are based on HST data (see Table 1 of Merritt & Ferrarese 2001c).

It was by isolating these secure detections that it became possible to unveil the existence of a fundamental, seemingly perfect correlation between black hole mass, $M_\bullet$, and velocity dispersion, σ, of the host bulge (Fig. 2, Ferrarese

& Merritt 2000; Gebhardt *et al.* 2000): the relation emerged from what had appeared almost as a scatter plot when the sample was restricted only to galaxies in which the SBH sphere of influence had been resolved (Ferrarese & Merritt 2000; Merritt & Ferrarese 2001b; Merritt & Ferrarese 2001c). A regression analysis, accounting for errors in both coordinates, of all published SBH detections (listed in Table 1 of Merritt & Ferrarese 2001c) gives

$$M_\bullet = (1.66 \pm 0.32) \times 10^8 \, M_\odot \left(\frac{\sigma}{200 \text{ km s}^{-1}} \right)^{4.58 \pm 0.52}. \quad (2)$$

Including the few preliminary masses based on unpublished analyses produces an indistinguishable slope of 4.64±0.47. The reduced χ^2 of the fit, 0.74, points to a relation with negligible intrinsic scatter, in agreement with the observations made by Ferrarese & Merritt (2000) based on the smaller sample available at the time. Because of its tightness, the $M_\bullet - \sigma$ relation has largely supplanted the well known correlation between $M_\bullet$ and bulge magnitude M_B (Fig. 3; Kormendy & Richstone 1995; Magorrian 1998), and has emerged as the tool of choice in the study of SBH demographics (Merritt & Ferrarese 2001b). Indeed, the tightness of the $M_\bullet - \sigma$ relation is its most puzzling feature, presenting formidable challenges to theoretical models for the formation and evolution of SBHs (e.g Haehnelt & Kauffmann 2000; Kauffmann & Haehnelt 2000; Adams *et al.* 2001; Ciotti & van Albada 2001; Burkert & Silk 2001). Even if one assumes that a tight relation was imprinted at an early stage of galaxy/SBH formation, it is difficult to understand how it could have survived unaltered in the face of mergers. It is especially remarkable that the relation should hold true for galaxies of disparate Hubble types (from SBs to compact ellipticals to cDs) belonging to wildly different environments (from rich clusters to the field), showing perfectly smooth (e.g. NGC 6251) or highly disturbed (e.g. NGC 5128) morphologies. For instance, it has recently been noted (McLure & Dunlop 2001) that the large scatter in the $M_\bullet - M_B$ relation (a reduced χ^2 of 23, Ferrarese & Merritt 2000) can be significantly reduced, but only at the expense of excluding most lenticular and spiral galaxies (and the odd elliptical, cf. Fig. 3). Even the remarkably tight correlation discovered by Graham *et al.* (2001) between $M_\bullet$ and "concentration parameter" C (the fraction of total bulge light contained within a predetermined radius) is marred by the occasional outlier. However, *every* galaxy, even the ones which do not obey the $M_\bullet - M_B$ or $M_\bullet - C$ relations, seems to conform magically to the $M_\bullet - \sigma$ relation.

4.2 Two Routes to SBH Demographics in Local Quiescent Galaxies

Because of its tightness, the $M_\bullet - \sigma$ relation provides us with a direct and powerful tool to estimate the mass density of SBHs, $\rho_\bullet$, in the local universe. One approach is to combine the known mass density of spheroids (e.g. Fukugita *et al.* 1998) with the mean ratio between the mass of the SBH and that of the host bulge. Merritt & Ferrarese (2001a) used the $M_\bullet - \sigma$ relation to estimate $M_\bullet$ for a sample of 32 galaxies for which a dynamical measurement of the mass of the hot stellar component was available (from Magorrian *et al.* 1998). For this sample, the frequency function $N[\log(M_\bullet/M_{\rm bulge})]$ is well approximated by a Gaussian with $\langle\log(M_\bullet/M_{\rm bulge})\rangle \sim -2.90$ and standard deviation ~ 0.45. This implies $M_\bullet/M_{\rm bulge} \sim 1.3\times10^{-3}$ or, when combined with the mass density in local spheroids from Fukugita *et al.* (1998), $\rho_\bullet \sim 5 \times 10^5$. This estimate is a factor of five smaller than obtained by Magorrian *et al.* (1998) using what we now believe to be inflated values for the masses of the central black holes in many galaxies.

Here I will present an independent derivation of $\rho_\bullet$ which, while not directly leading to $\langle M_\bullet/M_{\rm bulge}\rangle$, has the advantage of producing an analytical representation of the cumulative SBH mass density as a function of $M_\bullet$. The idea is simple: if $M_\bullet$ correlates with the luminosity of the host bulge, the SBH mass density can be calculated once the luminosity function of bulges is known. Black hole masses are related to bulge luminosity directly through the $M_\bullet - M_B$ relation, a representation of which is given by Ferrarese & Merritt (2000) as $\log M_\bullet = -0.36M_B + 1.2$. Unfortunately, the large scatter of the $M_\bullet - M_B$ relation (Fig. 3), combined with the small number of galaxies on which it is based, makes it impossible to establish whether elliptical and spiral galaxies follow a similar relation. Indeed, the observations of McLure & Dunlop (2001) cast doubts on whether spirals and lenticulars follow an $M_\bullet - M_B$ relation at all. This is unfortunate since the galaxy luminosity function does show a dependence on morphology (e.g., Marzke *et al.* 1998), and it is therefore desirable to conduct the analysis independently for different Hubble types. An alternative approach is to derive a relation between $M_\bullet$ and bulge luminosity by combining the $M_\bullet - \sigma$ relation (which, given the present sample, seems independent of the morphology of the host galaxy) with the Faber-Jackson relation for ellipticals and its equivalent for spiral bulges. The drawback here is that the Faber-Jackson relation has large scatter and is ill defined, especially for bulges.

The luminosity function for spheroids can be derived from the luminosity function of galaxies, generally represented as a Schechter function, once a ratio between total and bulge luminosity (which depends on the Hubble type of the

galaxy considered) is assumed. The latter is adopted from Table 1 of Fukugita *et al.* (1998). Here, I will use the galaxy luminosity function derived by Marzke *et al.* (1998) from the Second Southern Sky Redshift Survey (SSRS2), corrected to $H_0 = 75$ km s^{-1} Mpc^{-1} and an Einstein-de Sitter universe. Marzke *et al.* derived luminosity functions separately for E/S0s and spirals, in a photometric band B_{SSRS2}. This band is similar to the Johnson's B-band, where representations of both the $M_\bullet - M_B$ relation and the Faber-Jackson relation exist: $B_{SSRS2} = B + 0.26$ (Alonso *et al.* 1993). A Schechter luminosity function,

$$\Phi(L)\mathrm{d}L = \Phi_0 \left(\frac{L}{L_*}\right)^{\alpha} e^{-L/L_*} \frac{\mathrm{d}L}{L_*}, \tag{3}$$

is then easily transformed into a SBH mass density if $L = AM_\bullet{}^k$,

$$\Psi(M_\bullet)\mathrm{d}M_\bullet = \Psi_0 \left(\frac{M_\bullet}{M_*}\right)^{k(\alpha+1)-1} e^{-(M_\bullet/M_*)^k} \frac{\mathrm{d}M_\bullet}{M_*}, \tag{4}$$

where $\Psi_0 = k\Phi_0$, $M_* = (\beta L_* 10^{0.4\times 0.26}/A)^{1/k}$, and $\beta \equiv L/L_{bulge} = 0.23$ for spirals and 0.76 for E/S0 galaxies. β is the sum of the ratios between bulge to total B-band luminosity for different Hubble types, each weighted by the fraction of the mean luminosity density contributed by each type (from Fukugita *et al.* 1998).

Fig. 4 shows the cumulative SBH mass function separately for the E/S0 and spiral populations, derived from the $M_\bullet - M_B$ relation (from Ferrarese & Merritt 2000, dotted lines) and the $M_\bullet - \sigma$ relation (from this paper) combined with the Faber-Jackson relations for ellipticals and spirals (from Kormendy & Illingworth 1983, corrected to $H_0 = 75$ km s^{-1} Mpc^{-1}). While the two distributions differ in the details, there is little difference in the total mass density, which falls in the range $(4-5)\times 10^5\ M_\odot$ Mpc^3. This is in excellent agreement with the estimate of Merritt & Ferrarese (2001a).

Table 1 summarizes the mass density estimates for SBHs discussed in the preceeding three sections. While a detailed comparison of the distribution of masses remains to be carried out (for instance, Fig. 1 suggests a larger fraction of very massive black holes, $M > 10^9$, in high redshift QSOs than have been found in local galaxies), the overall picture is one of agreement: local studies seem to have recovered the overall mass density inferred from high redshift QSOs. It appears that supermassive black holes are a fundamental component of every large galaxy.

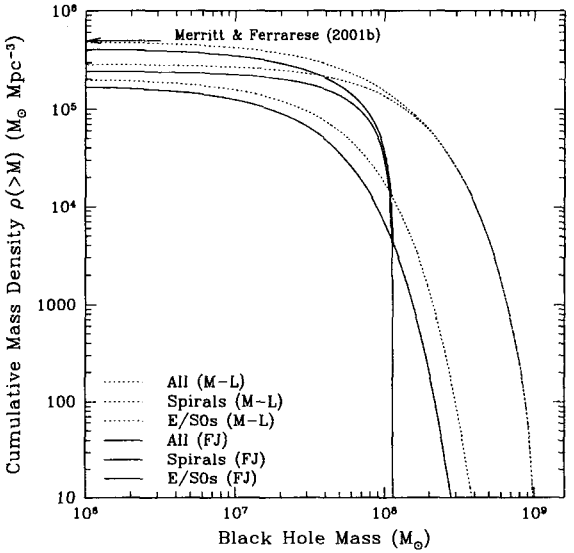


Figure 4: The mass function in local black holes for spirals (blue), E/S0 (red) and a complete sample of galaxies (black). Dotted lines are derived from the $M_\bullet - M_B$ relation, solid lines from the $M_\bullet - \sigma$ combined with the Faber-Jackson relation (as described in the text).

Summary of Mass Densities in Supermassive Black Holes

Method	$\rho_\bullet$ ($10^5\ M_\odot\ Mpc^{-3}$)
QSO optical counts, $0.3 < z < 5.0$	$2 - 4$
AGN X-ray counts, $z > 0.3$	$0.6 - 9$
Spectral fit to the X-ray background, z unknown	$2 - 30$
Local AGNs, $z < 0.1$	$0.05 - 0.6$
Local Quiescent Galaxies, $z < 0.0003$	$4 - 5$

5 Moving Forward: Open Issues

5.1 What More Can Be Learned about the $M_\bullet - \sigma$ Relation?

With so much progress in the past few years, it is only natural to be optimistic about what the near future might bring. Indeed, a considerable amount of effort will be devoted to the study of supermassive black holes in nearby galaxies, with HST remaining the instrument of choice for the investigation. Roughly 130 galaxies have, or will be, observed with HST/STIS within the next year. While only a fraction of these observations are likely to lead to secure SBH detections (Merritt & Ferrarese 2001c), these results are highly anticipated, and will help to better define the slope and scatter of the $M_\bullet - \sigma$ relation.

Nevertheless, one important section of parameter space will remain unex-

plored. Now that the existence of SBHs is as well established as that of the galaxies in which they reside, the most pressing need has become, in my opinion, an exploration of the low mass end of the $M_\bullet - \sigma$ relation. However, the vast majority of the galaxies in the HST pipeline are expected to host SBHs with $M_\bullet \sim 10^8 M_\odot$, a range already well-sampled by the current data. None of the ongoing programs is likely to measure a SBH of $M_\bullet < 10^7\ M_\odot$ (Merritt & Ferrarese 2001c).

This is unfortunate since determining how far the $M_\bullet - \sigma$ relation extends is key for discriminating between different scenarios for the formation of SBHs. The smallest nuclear SBHs whose masses have been established dynamically are in the Milky Way (Genzel *et al.* 2000) and M32 (Joseph *et al.* 2001), both with $M_\bullet \approx 3 \times 10^6 M_\odot$ (Fig. 2). Evidence for black holes with $10^3 < M_\bullet < 10^6$ $M_\odot$ (dubbed "intermediate" mass black holes, or IBHs) is so far circumstantial, the most likely candidates being the super-luminous *off-nuclear* X-ray sources (ULXs) detected by Chandra in a number of starburst galaxies (Fabbiano *et al.* 2001; Matsumoto *et al.* 2001).

The link between IBHs and SBHs is unclear. If Chandra's off-nuclear ULXs are indeed IBHs, they could sink slowly to the galaxy center through dynamical friction and provide the seeds for nuclear SBHs (Ebisuzaki *et al.* 2001). Or the latter might be born *in situ*, through collapse of a protogalactic cloud, possibly before the onset of star formation in the bulge (Loeb 1993; Silk & Rees 1998; Haehnelt, Natarajan & Rees 1997). Deciding between these and competing formation scenarios will undoubtedly keep theorists busy for many years. However, different theories would almost certainly make different predictions about the form of the $M_\bullet - \sigma$ relation, and this is the most promising route for distinguishing between them. For instance, *in situ* formation in nuclei is unlikely to result in black holes less massive than $\sim 10^6 M_\odot$ (e.g Haehnelt, Natarajan & Rees 1998), while accumulation of IBHs would probably not result in as tight a correlation between $M_\bullet$ and σ unless some additional feedback mechanism were invoked (e.g. Burkert & Silk 2001). But little progress is likely to be made until we know whether IBHs are present in galaxy nuclei and if so, where they lie relative to the $M_\bullet - \sigma$ relation defined by SBHs. Therefore, exploring the $M_\bullet - \sigma$ relation in the $M < 10^6\ M_\odot$ range will be an important challenge in the years to come.

A first step in this direction has been taken recently with the derivation of an upper limit, of a few thousand solar masses, for the putative black hole inhabiting the nucleus of the nearby spiral M33 (Merritt, Ferrarese & Joseph 2001; Gebhardt *et al.* 2001; Valluri *et al.* 2002). As small as this upper limit might seem, it is still consistent with the $M_\bullet - \sigma$ relation as characterized in this paper, when extrapolated (by three orders of magnitude!) to the thousand

solar mass range. Unfortunately, until the next technological leap, there is little hope of significantly tightening this upper limit: at the distance of M33, the black hole's sphere of influence is well below (by at least an order of magnitude) the resolution capabilities of HST. Indeed, with one notable exception, there are *no* galaxies expected to contain a black hole below the $10^6\ M_\odot$ mark that are close enough, and have high enough central surface brightness, to allow HST to measure $M_\bullet$. The one exception, the Local Group spheroidal galaxy NGC 205, is scheduled to be observed by HST as part of program 9448 (P.I. L. Ferrarese). NGC 205 is expected to host a $\sim 7.5 \times 10^5\ M_\odot$ black hole; at a distance of 740 kpc, a black hole as small as $6 \times 10^5\ M_\odot$ can be detected. Even so, it seems inevitable that, to fully characterize the low mass range of the $M_\bullet - \sigma$ relation, we must look beyond HST.

In my opinion, the answer is reverberation mapping. Although the obvious drawback is that it is only applicable to the 1% of galaxies with Type 1 AGNs, reverberation mapping is intrinsically unbiased with respect to black hole mass, provided the galaxies can be monitored with the appropriate time resolution. Furthermore, reverberation mapping can probe galaxies at high redshifts and with a wide range of nuclear activity, opening an avenue for the exploration of possible dependences of the $M_\bullet - \sigma$ relation on cosmic time and activity level.

The stage is being set to embark upon this new endeavor. In the past few years, the reliability of reverberation-mapping masses has been called into question on both observational (e.g. Ho 1999; Richstone *et al.* 1998) and theoretical (Krolik 2001) grounds. However, on the observational side, the doubts appear to be dissipating. The observation that SBHs in AGNs appeared to be undermassive, by a factor ~ 50, compared to SBHs in quiescent galaxies (Wandel 1999), was apparently the result of two erroneous assumptions: the overestimate (by a factor $\sim$ six) of SBH masses in quiescent galaxies derived from the $M_\bullet - M_B$ relation of Magorrian *et al.* (1998); and an overestimate of the AGN host bulge magnitudes (by up to ~ 3.5 mag) adopted by Wandel (McLure & Dunlop 2000; Merritt & Ferrarese 2001c; Wandel 2002). Indeed, Merritt & Ferrarese (2001c) conclude that the ratio of SBH to bulge mass in Seyfert, QSO and quiescent galaxy samples are all consistent: $\langle M_\bullet/M_{\rm bulge}\rangle = 0.09\%$ (QSOs) and 0.12% (Seyferts), $\langle M_\bullet/M_{\rm bulge}\rangle = 0.13\%$ for quiescent galaxies.

On the theoretical side, Krolik (2001) argues that the unknown BLR geometry, radial emissivity distribution, and angular radiation pattern of the line emission, coupled with the often less than optimal temporal sampling of the data, can lead to systematic errors in the reverberation masses of a factor $\sim$ three or more. While there is little doubt that Krolik's objections are all well-justified, my collaborators and I have taken an observational approach to this issue. Since there are no independent measurements of $M_\bullet$ for any of

the reverberation-mapped AGNs, we have opted for an indirect comparison by placing these galaxies onto the $M_\bullet - \sigma$ plane. Initial results (Ferrarese *et al.* 2001) suggest that the AGN sample follows the same $M_\bullet - \sigma$ relation as the quiescent galaxies on which the relation is defined. More secure conclusions should be reached within the next year, once the AGN sample is doubled (Pogge *et al.* 2002). At the moment, the evidence suggests that reverberation mapping works, in spite of the theoretically motivated concerns.

5.2 Beyond the $M_\bullet - \sigma$ Relation: Exploring the Dark Side of Galaxies

The $M_\bullet - \sigma$ relation probes a direct connection between SBHs and galactic bulges. The velocity dispersion, σ, is measured within a region which, though large compared to the black hole sphere of influence, is at least an order of magnitude smaller than the optical radius of the galaxy, and is likely dominated by luminous matter (Faber & Gallagher 1979). Therefore, σ is unable to tell us about the connection between SBHs and other fundamental baryonic structures, such as the galactic disk or halo, while the link to the dark matter (DM) component also remains utterly unexplored.

That this issue has not yet been addressed is somewhat surprising, since it is not the mass of the bulge but rather, the *total mass* of the galaxy (or of the DM halo), which is the key ingredient of most theoretical models proposed for the formation of SBHs (Adams, Graff & Richstone 2000; Monaco *et al.* 2000; Haehnelt, Natarajan & Rees 1998; Silk & Rees 1998; Haehnelt & Kauffmann 2000; Cattaneo, Haehnelt & Rees 1999; Loeb & Rasio 1994). Once the models predict a correlation with total mass (or DM halo mass), the correlation with bulge mass is implicit because, in standard CDM scenarios, the bulge mass is loosely determined by the halo properties (e.g. van den Bosch 2000; Haehnelt, Natarajan & Rees 1998; Zhang & Wyse 2000).

It is natural to ask whether the $M_\bullet - \sigma$ relation might just be the by-product of an even more fundamental relation between $M_\bullet$ and the total gravitational mass of the galaxy. As it turns out, such a fundamental relation is likely to exist (Ferrarese 2002). Fig. 5 demonstrates the existence of a tight correlation between the bulge velocity dispersion (the same quantity used in defining the $M_\bullet - \sigma$ relation, typically measured within an aperture of size $R \lesssim 0.5$ kpc) and the circular velocity v_c, measured at radii $R \sim 20 - 80$ kpc, for a sample of 16 spiral galaxies. A regression analysis, accounting for errors in both variables, gives

$$\log v_c = (0.88 \pm 0.17) \log \sigma + (0.47 \pm 0.35) \tag{5}$$

with a reduced χ^2 of 0.64.

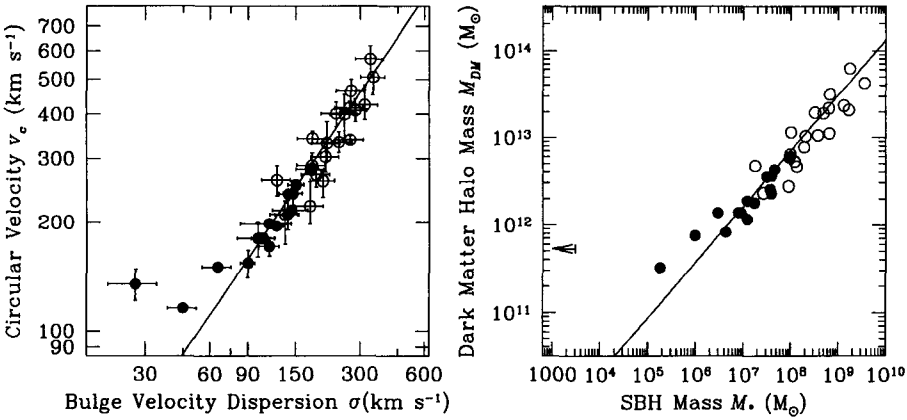


Figure 5: (left) Correlation between the rotational velocity and bulge velocity dispersion for a sample of 16 spiral galaxies (solid circles) and 21 ellipticals (open circles; plot adapted from Ferrarese 2002).
Figure 6: (right) Same as Fig. 5, but with v_c and σ converted to halo mass and black hole mass respectively (see text for further details). The upper limit on the SBH mass in M33 (Merritt *et al.* 2001) is shown by the arrow.

For spiral galaxies, v_c is measured directly from HI or optical rotation curves. In elliptical galaxies, v_c can be derived from dynamical models of the observed stellar absorption line profiles, velocity dispersion and surface brightness profiles. Fig. 5 shows that the spirals naturally blend with a sample of 21 elliptical galaxies (from Kronawitter et al. 2000) in the $v_c - \sigma$ plane; both samples obey the relation given in equation (5).

The implications of equation (5) are exciting. The circular velocity v_c is a measure of gravitational mass through the virial theorem, and can be related to the DM halo mass (Navarro & Steinmetz 2000; Bullock et al. 2001). Keeping in mind that, as discussed in section 5.1, the $M_\bullet - \sigma$ relation is not well defined below $10^7\ M_\odot$, and not defined at all below $10^6\ M_\odot$, the $v_c - \sigma$ relation can be translated into a relation between the mass of the central black hole (related to σ through equation 2) and that of the DM halo (Fig. 6):

$$\frac{M_\bullet}{10^8\ M_\odot} \sim 0.046 \left(\frac{M_{DM}}{10^{12}\ M_\odot} \right)^{1.6} \tag{6}$$

(Ferrarese 2002). The existence of this relation seems to conflict with recent claims that SBHs do not relate to any other galactic structure but the bulge (Richstone 1998; Kormendy & Gebhardt 2001; Gebhardt *et al.* 2001).

The relation between $M_\bullet$ and M_{DM} is non-linear, with the ratio $M_\bullet/M_{DM}$

decreasing from 6×10^{-5} for $M_{DM} \sim 10^{14}\ M_{\odot}$, to 5×10^{-6} for $M_{DM} \sim 10^{12}\ M_{\odot}$. Haehnelt, Natarajan & Rees (1998) advocated a nonlinear relation between SBH and DM halo mass in order to reproduce the luminosity function of QSOs, noting that a linear relation would translate into too low a value for the QSO duty cycle, $t_{QSO} \sim 3 \times 10^5$ yr. Increasing the QSOs lifetime to values more in line with current observational constraints (e.g. Martini & Weinberg 2001) produces an increasingly steeper relation between $M_\bullet$ and M_{DM}. If $t_{QSO} \sim 1.5 \times 10^7$ yr (equal to the Salpeter time), then the slope of the $M_\bullet - M_{DM}$ relation must be increased to ~ 2 to provide a reasonable fit to the QSO luminosity function. The empirical correlation shown in Fig. 6 seems to support such claims. Furthermore, Fig. 6 indicates that the tendency of massive halos to become less efficient in forming SBH as M_{DM} decreases, is even more pronounced for halos with $M_{DM} < 10^{12}\ M_{\odot}$, and breaks down completely in the case of M33. Such halos might indeed be *unable* to form SBH, as proposed on theoretical grounds by Haehnelt, Natarajan & Rees (1998) and Silk & Rees (1998).

5.3 Additional Clues to SBH and Galaxy Formation and Further Challenges

I conclude this review with some general comments about the early stages of galaxy and black hole formation. The $M_\bullet - \sigma$ relation (and, even more so, the $M_\bullet - M_{DM}$ relation) implies a causal connection between the evolution of black holes and their host galaxies. But what came first: the stars or the black holes? And was the $M_\bullet - \sigma$ relation imprinted during the early stages of galaxy formation? The answer to the latter question is generally assumed to be affirmative, but in fact we have no direct proof of it. The most distant galaxy in the $M_\bullet - \sigma$ plot (NGC 6251, Ferrarese & Ford 1999) is at ~ 100 Mpc. Studies of reverberation-mapped galaxies (Ferrarese *et al.* 2001, Pogge *et al.* 2002) have reached two times farther, and it might be technically possible to push the envelope up to $z \sim 1$. It seems unlikely that we will ever be able to build an $M_\bullet - \sigma$ relation at the redshift corresponding to the optically bright phase of the QSOs ($z \sim 2 - 3$), let alone at redshifts at which the first protogalactic fragments are believed to have formed, $z > 5$. Present day dwarf galaxies might very well be relics from such an era (Mateo 1998; Carraro *et al.* 2001); however, detecting SBHs in these systems requires a spatial resolution well beyond the capabilities of present instrumentation. In fact, the $M_\bullet - \sigma$ relation is defined primarily by bright giant ellipticals which are likely to have an extensive history of merging. In other words, we have no direct information about the "primordial" connection between supermassive black holes and their hosts: what we see is the result of gigayears of evolution.

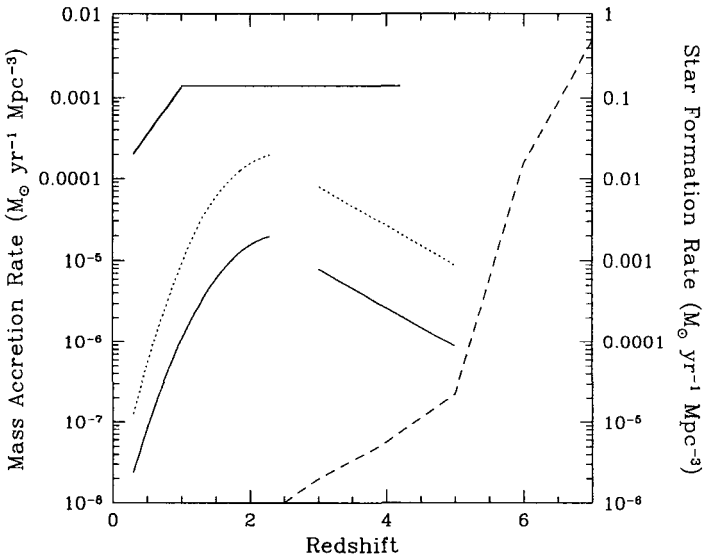


Figure 7: The mass accretion rate onto supermassive black holes with $M_\bullet > 7 \times 10^8\ M_\odot$ (thin solid line, units shown on the left axis), compared to the star formation rate from Steidel *et al.* 1999 (thick solid line, units on the right axis) and the epoch of formation of the DM halos which host such SBHs according to Fig. 5 and 7 (dashed line, from Gottlöber, Klypin & Kravtsov 2001). Eighty percent of the halos which (presumably) host QSO engines are capable of forming stars at redshifts $z > 5$. The dotted line shows the mass accretion rate onto high redshift QSOs corrected for the contribution of obscured objects, as in Gilli *et al.* 2001.

A scenario in which galaxy formation precedes the formation of supermassive black holes seems to fit more naturally within the current paradigm of hierarchical structure formation (e.g. Miralda-Escude & Rees 1997). For instance, star formation can proceed in halos with virial temperature as low as 10^4 K, which can form at redshifts $z > 10$ (e.g. Ostriker & Gnedin 1996). Subsequent stellar evolution in these systems would produce enough energy through stellar winds or supernovae explosions to expel most of the remaining gas from the shallow potential wells (Couchman & Rees 1986; Dekel & Silk 1986), likely inhibiting the formation of supermassive black holes. Deeper potential wells, which are more conducive to SBH formation (e.g. Haehnelt, Natarajan & Rees 1998) would only form at later times. Studies of elemental abundances in high redshift ($z > 3$) QSOs support this view: most of the metal enrichment and star formation seem to have taken place at least 1 Gyr before the luminous phase of the QSO (Hamann & Ferlan 1999 and references

therein; Dietrich, *et al.* 2001).

Fig. 7 shows a comparison between the mass accretion rate onto optically luminous QSOs with $M_\bullet > 7\times10^8$ (corresponding to the magnitude limit of the SDSS QSO Survey for objects radiating at the Eddington limit), and the star formation rate from Steidel *et al.* 1999 (see also Abraham *et al.* 1999; Cowie *et al.* 1997). Similarities between the two curves, which have been noted many times (e.g. Boyle & Terlevich 1998) are diminished by these recent results, even after the QSO results are corrected for the possible contamination of obscured objects (Barger *et al.* 2001; Gilli *et al.* 2001; Salucci *et al.* 1998). If anything, Fig. 7 supports the conclusion that star formation was well underway by the time the QSOs started shining.

The connection between QSO activity and merging rate is also not readily apparent: observations show that the merging rate depends on redshift as $(1+z)^\alpha$ with $\alpha = 2-4$ (Le Fevre *et al.* 1999; Burkey *et al.* 1994; Carlberg *et al.* 1994; Yee & Ellingson 1995; Abraham 1999). Even in the $z < 2.3$ range, where both curves decline, the number of mergers declines by at most a factor 30, while the comoving density of QSO declines by three orders of magnitude. Perhaps more telling is the comparison with the merging history of DM halos and the ensuing formation of galaxies. Fig. 7 also shows the distribution of formation redshifts for present day halos with virial velocities > 300 km s^{-1} taken from the N-body simulation of Gottlöber, Klypin & Kravtsov (2001). According to Figures 5 and 6, these are the halos associated with the black holes sampled by the SDSS, also shown in Fig. 7. Virtually all such halos are able to host a luminous galaxy (a condition reached when the halo progenitor first reaches a virial velocity > 50 km s^{-1}) before a redshift ~ 2.5, i.e. before the optically bright phase of the QSOs.

In the midst of all this, one thing is certain: SHBs can no longer be studied in isolation. Understanding how they form, and how they shape their surroundings, requires a good deal more information from seemingly unrelated fields than could have been anticipated just a few years ago.

References

1. Abraham, R., in *Galaxy Interactions at Low and High Redshift*, IAU Symposium 186, eds J. E. Barnes, and D. B. Sanders, p.11 (1999)
2. Abraham, R., *et al.* , *MNRAS* **308**, 569 (1999)
3. Adams, F.C., Graff, D.S., & Richstone, D., *ApJ* **551**, L31 (2000)
4. Alonso, M.V., *et al.* , *AJ* **106**, 676 (1993)
5. Barger, A. J., *et al.* , *AJ* **121**, 662 (2001)
6. Barth, A. J., *et al.* , *ApJ* **555**, 685 (2001)

7. Binney, J., & Mamon, G.A., *MNRAS* **200**, 361 (1982)
8. Blandford, R. D. & McKee, C. F., *ApJ* **255**, 419 (1982)
9. Boyle, B.J., & Terlevich, R.J., *MNRAS* **293**, L49 (1998)
10. Boyle, B.J., *et al.* , *MNRAS* **317**, 1014 (2000)
11. Bower, G. A., *et al.* , *ApJ* **492**, L111 (1998)
12. Braatz, J. A., Wilson, A. S. & Henkel, C., *ApJS* **106**, 51 (1996)
13. Bullock, J.S., *et al.* , *MNRAS* **321**, 559 (2001)
14. Burkert, A. & Silk, J., *ApJ* **554**, 151 (2001)
15. Burkey, J.M., *et al.* , *ApJ* **429**, L13 (1994)
16. Carlberg, R.G., Pritchet, C.J., & Infante, L., *ApJ* **435**, 540 (1994)
17. Cattaneo, A., Haehnelt, M. G. & Rees, M. J., *MNRAS* **308**, 77 (1999)
18. Carraro, G., *et al.* , *MNRAS* **327**, 69 (2001)
19. Chokshi, A. & Turner, E. L., *MNRAS* **259**, 421 (1992)
20. Ciotti, L., & van Albada, T. S. , *ApJ*, in press (astro-ph/0103336) (2001)
21. Couchman, H. M. P., & Rees, M. J., *MNRAS* **221**, 53 (1986)
22. Cowie, L.L., *et al.* , *ApJ* **481**, L9 (1997)
23. Cretton, N. & van den Bosch, F.C., *ApJ* **514**, 704 (1999)
24. Dekel, A., & Silk, J., *ApJ* **303**, 39 (1986)
25. Dietrich, M., *et al.* , *A&A* **351**, 31 (1999)
26. Ebisuzaki, T., *et al.* , *ApJ* **562**, L19 (2001)
27. Elvis, M., *et al.* , *ApJ* **310**, 291 (1986)
28. Emsellem, E., Dejonghe, H. & Bacon, R., *MNRAS* **303**, 495 (1999)
29. Fabbiano, G., Zezas, A., Murray, S.S., *ApJ*, in press (astro-ph/0102256) (2001)
30. Faber, S., & Gallagher, J.S., *ARAA* **17**, 135 (1979)
31. Fabian, A.C., & Canizares, C.R., *Nature* **333**, 829 (1988)
32. Fabian, A.C., & Iwasawa, *MNRAS* **303**, L34 (1999)
33. Fan, X., *et al.* , *AJ* **121**, 54 (2001)
34. Ferrarese, L. & Ford, H.C., *ApJ* **515**, 583 (1999)
35. Ferrarese, L., Ford, H.C. & Jaffe, W., *ApJ* **470**, 444 (1996)
36. Ferrarese, L. & Merritt, D., *ApJ* **539**, L9 (2000)
37. Ferrarese, L., Pogge, R. W., Peterson, B. M., Merritt, D., Wandel, A. & Joseph, C. M., *ApJ* **555**, L79 (2001)
38. Ferrarese, L., *ApJ*, submitted (2002)
39. Fukugita, M., Hogan, C.J., & Peebles, P.J.E., *ApJ* **503**, 518 (1998)
40. Gebhardt, K., *et al.* , *AJ* **119**, 1157 (2000)
41. Gebhardt, K., *et al.* , *AJ* **122**, 2469 (2001)
42. Genzel, R., Pichon, C., Eckart, A., Gerhard, O. E. & Ott, T., *MNRAS* **317**, 348 (2000)
43. Ghez, A. M., Klein, B. L., Morris, M. & Becklin, E. E., *ApJ* **509**, 678

(1998)
44. Gilli, G.M., Salviati, M., & Hasinger, G., *AA*, in press (astro-ph/0011341) (2001)
45. Gottlöber, S., Klypin, A., & Kravtsov, A.V., *ApJ* **546**, 223 (2001)
46. Graham, A.W., Erwin, P., Caon, N., & Trujillo, I., *ApJ* **563**, L11 (2001)
47. Greenhill, L.J., *et al.* , *ApJ* **440**, 619 (1995)
48. Greenhill, L.J., Gwinn, C. R., Antonucci, R. & Barvainis, R., *ApJ* **472**, L21 (1996)
49. Greenhill, L.J., Moran, J. M. & Herrnstein, J. R., *ApJ* **481**, L23 (1997)
50. Greenhill, L.J., *et al.* , *ApJ* **486**, L15 (1997)
51. Greenhill, L.J., *et al.* , *ApJ*, in press (astro-ph/0111493) (2002)
52. Haehnelt, M. G. & Kauffmann, G., *MNRAS* **318**, L35 (2000)
53. Haehnelt, M. G., Natarajan, P. & Rees, M. J., *MNRAS* **300**, 817 (1998)
54. Hamann, F., & Ferland, G., *ARAA* **37**, 487 (1999)
55. Harms, R.J., *et al.* , *ApJ* **435**, L35 (1994)
56. Ho, L. C., in *Observational Evidence for Black Holes in the Universe*, eds. S. K. Chakrabarti (Dordrecht: Reidel), p. 157 (1999)
57. Joseph, C., *et al.* , *ApJ* **550**, 668 (2001)
58. Kaspi, S., Smith, P.S., Netzer, H., Maoz, D., Jannuzi, B.T., & Giveon, U., *ApJ* **533**, 631 (2000)
59. Kauffmann, G. & Haehnelt, M., *MNRAS* **311**, 576 (2000)
60. Koratkar, A. P. & Gaskell, C. M., *ApJS* **75**, 719 (1991)
61. Kormendy, J., & Richstone, *ARAA* **27**, 235 (1995)
62. Kormendy, J., *ApJ* **325**, 128 (1988)
63. Kormendy, J., & Illingworth, G., *ApJ* **265**, 632 (1983)
64. Kormendy, J., *et al.* , *ApJ* **459**, L57 (1996a)
65. Kormendy, J., *et al.* , *ApJ* **473**, L91 (1996b)
66. Kormendy, J. & Gebhardt, K., astro-ph/0105230 (2001)
67. Krolik, J., *ApJ* **555**, 72 (2001)
68. Kronawitter, A., *et al.* , *A&AS* **144**, 53 (2000)
69. Le Fevre, O., *et al.* , *MNRAS* **311**, 565 (2000)
70. Loeb, A., *ApJ* **403**, L542 (1993)
71. Loeb, A., & Rasio, F. *ApJ* **432**, L52 (1994)
72. Macchetto, F., *et al.* , *ApJ* **489**, 579 (1997)
73. Magorrian, J., *et al.* , *AJ* **115**, 2285 (1998)
74. Malkan, M. A., Gorjian, V. & Tam, R., *ApJS* **117**, 25 (1998)
75. Maiolino, R., & Rieke, G.H., *ApJ* **454**, 95 (1995)
76. Marzke, R.O., *et al.* , *ApJ* **503**, 617 (1998)
77. Mateo, M., *ARAA* **36**, 435 (1998)
78. Matsumoto, H., *et al.* , *ApJ* **547**, L25 (2001)

79. Marconi, A., *et al.* , *ApJ* **549**, 915 (2001)
80. Martini, P., & Weinberg, D.H., *ApJ* **547**, 12 (2001)
81. McLure, R.J., & Dunlop, J.S., *MNRAS* **327**, 199 (2001)
82. McLure, R.J., & Dunlop, J.S., *MNRAS*, in press, (astro-ph/0108417) (2002)
83. Merrifield, M.R., *AJ* **103**, 1552 (1992)
84. Merritt, D. & Ferrarese, L., *MNRAS* **320**, L30 (2001a)
85. Merritt, D. & Ferrarese, L., *ApJ* **547**, 140 (2001b)
86. Merritt, D., & Ferrarese, L., to appear in *The Central Kpc of Starbursts and AGNs*, eds. J.H. Knapen, J.E. Beckman, I. Shlosman, & T.J. Mahoney (astro-ph/0107134) (2001c)
87. Merritt, D., Ferrarese, L. & Joseph, C., *Science* **293**, 1116 (2001)
88. Miralda-Escude, J., & Rees, M.J., *ApJ* **478**, L57 (1997)
89. Miyoshi, M., *et al.* , *Nature* **373**, 127 (1995)
90. Monaco, P., Salucci, P., & Danese, L. , *MNRAS* **311**, 279 (2000)
91. Narayan, R., & Yi, I., *ApJ* **428**, L13 (1994)
92. Navarro, J.F., & Steinmetz, M., *ApJ* **538**, 477 (2000)
93. Netzer, H. & Peterson, B. M., in Astronomical Time Series, ed. D. Maoz, A. Sternberg, & E.M. Leibowitz, (Dordrecht: Kluwer), 85 (1997)
94. Ostriker, J.P., & Gnedin, N.Y., *ApJ* **472**, L63 (1996)
95. Padovani, P., Burg, R., & Edelson R.A., *ApJ* **353**, 438 (1990)
96. Peterson, B. M., *PASP* **105**, 247 (1993)
97. Peterson, B. M. & Wandel, A., *ApJ* **540**, L13 (2000)
98. Pogge, R., *et al.* , in preparation (2002)
99. Rees, M., *et al.* , *Nature* **295**, 17 (1982)
100. Richstone, D.O., Bower, G. & Dressler, A., *ApJ* **353**, 118 (1990)
101. Richstone, D.O., *et al.* , *Nature* **395**, 14 (1998)
102. Robinson, I., in *Quasi-Stellar Source and Gravitational Collapse*, University of Chicago press (1965)
103. Rohlfs, K., & Kreitschmann, J., *A&A* **201**, 51 (1988)
104. Salucci, P., *et al.* , *MNRAS* **307**, 637 (1999)
105. Sargent, W.L.W., *et al.* , *ApJ* **221**, 731 (1978)
106. Sarzi, M., *et al.* , *ApJ* **550**, 65 (2001)
107. Silk, J. & Rees, M. J., *A&A* **331**, L4 (1998)
108. Small, T., & Blandford, R.D., *MNRAS* **259**, 725 (1992)
109. Soltan, A., *MNRAS* **200**, 115 (1982)
110. Steidel, C.C., *et al.* , *ApJ* **519**, 1 (1999)
111. Tonry, J. L., *ApJ* **283**, L27 (1984)
112. Trotter, A.S., *et al.* , *ApJ* **495**, 740 (1998)
113. Valluri, M., *et al.* , in preparation (2002)

114. van den Bosch, F., *ApJ* **530**, 177 (2000)
115. van der Marel, R. P. & van den Bosch, F. C., *AJ* **116**, 2220 (1998)
116. Verdoes Kleijn, G.A., van der Marel, R., Carollo, C. M. & de Zeeuw, P. T., *AJ* **120**, 1221 (2000)
117. Vila-Vilaro, B., *PASPJ*, **52**, 305, (2000)
118. Wandel A., *ApJ* **519**, 39 (1999)
119. Wandel, A., Peterson, B. M. & Malkan, M.A., *ApJ* **526**, 579 (1999)
120. Wandel, A., *ApJ*, in press, (astro-ph/0108461) (2002)
121. Watson, W. D. & Wallin, B. K., *ApJ* **432**, 35 (1994)
122. Yee, H.K.C., & Ellingson, E., *ApJ* **445**, 37 (1995)
123. Zhang, B., & Wyse, R., *MNRAS* **313**, 310 (2000)

KILOPARSEC JETS FROM MASSIVE BLACK HOLES IN RADIO-LOUD AGN

ANDRÉ B. FLETCHER
Korea Astronomy Observatory
61-1 Hwaam-Dong, Yuseong-Gu, Daejeon 305-348, Korea
E-mail: abfletch@kao.re.kr

A small fraction of active galactic nuclei (AGN) emit powerful radio waves from component structures (cores, jets, lobes) with sizes ranging from pc to Mpc. The ultimate power source of these radio-loud AGN (RLAGN) is most probably the gravitational accretion of galactic stars, dust and gas onto central massive black holes (MBHs) imprisoning $10^6 M_\odot$ to $10^{10} M_\odot$; such objects are now believed to naturally form in galactic centers during their dynamical evolution. It should be possible to investigate the cosmological behaviour of RLAGN, and of their resident MBHs, by comparing global Monte-Carlo population models with the many publicly available radio catalogs and imaging databases. A simple model is applied to the jet angular size distribution of $\sim 4 \times 10^3$ RLAGN imaged in the 19-year, MITVLA gravitational lens snapshot survey; if these distant kpc-scale jets are powered by Eddington-accreting MBHs, a statistical upper limit ($v/c < \sim 0.1$) can be set on their head propagation speeds. This is a small, first step in the confrontation of statistically large RLAGN sample observations with the theory of their central MBH engines. The time evolution of accreting Kerr MBHs is briefly addressed, and a simple model is sketched for cosmologically evolving MBHs in radio-loud AGN.

1 Introduction

There is now much observational evidence for astrophysical black holes[6] (BHs), both stellar and supermassive. Massive black holes (MBHs) are now believed to develop naturally in galactic centers during their normal dynamical evolution [11 14 24 32 33]. We should then expect to find MBH engines in the centers of active galactic nuclei (AGN), and, in particular, radio-loud AGN (RLAGN). Over the past ~40 years, abundant data of various types, and from many independent surveys, have been collected on many thousands of bright, extragalactic radio sources (EGRS); almost all of which are RLAGN. In the near future, high-resolution imaging data, broad-band spectra, identifications, and redshifts, will become publicly available for these many thousands of RLAGN. All of this information needs to be correlated, and compared with numerical models based on physical theories for the time evolution of MBHs and their derivative jets, disks, and lobes. RLAGN form a very large sub-family of AGN, and their intrinsic luminosities and proper sizes span very broad ranges. Their number density as a function of flux density strongly indicates that they evolve on

cosmological timescales, but the details of this long-term evolution are still not fully understood. To compile basic parameters (e.g. luminosities and projected sizes) for a statistically large sample, redshifts need to be secured. This is a difficult task, as the optical counterparts of RLAGN are usually faint.

The magnetospheric physics of MBHs, and the relativistic plasma, gas, radiation and particle dynamics involved in RLAGN evolution, are all very complex, and so theories have advanced much slower than observations. It is only recently that computer simulations have attacked the general-relativistic MHD (GRMHD) problem of an accreting Kerr BH, which is very difficult even in the axisymmetric (2.5D) case. However, a convergence of theory and observation, for the case of evolving MBHs within RLAGN, is foreseen for the future.

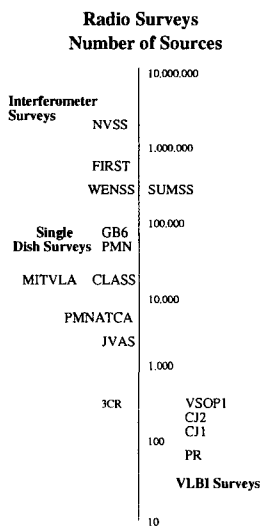


Figure 1: Numbers of sources found in large radio surveys (on a log scale).

2 Current Status

Based on Einstein's general theory of relativity, we expect that astrophysical BHs should exist; they are almost certainly formed as a result of normal evolutionary processes in stars, as well as from stellar dynamical evolution within sufficiently dense systems, such as the centers of galactic bulges. There is clear observational evidence (extreme luminosities; broad, non-thermal spectra; op-

Table 1: Basic parameters of selected large radio source surveys.

Survey	Freq ν GHz	Flux Lim. mJy	Sky Region	θ_{resol} $''$	Radiotelescope	Ref	N_{src}
SKY SCANNING SURVEYS							
3CR	0.18	9Jy	$-0.5 < \delta < +90$	670	Cambridge790m	45	328
PMN	4.85	20	$-87.5 < \delta < +10$	252	Parkes 64m	51	50814
GB6	4.85	18	$0 < \delta < +75$	210	NRAO-GB 91m	15	75162
SKY IMAGING SURVEYS							
WENSS	0.33	18	$+29 < \delta < +90$	54	Westerbork4km	43	0.3×10^6
SUMSS	0.84	5	$-90 < \delta < -30$	43	Molonglo1.6km	4	0.4×10^6
FIRST	1.40	1	$\|b\| > +30$	5	VLA-B 12km	49	0.9×10^6
NVSS	1.40	2.5	$-40 < \delta < +90$	45	VLA-D 1.3km	9	2.0×10^6
SOURCE IMAGING SURVEYS (ARRAYS)							
JVAS	8.44	60	$-30 < \delta < +90, \|b\| > 2.5$	0.20	VLA-A 36km	22	~3000
CLASS	8.44	20	$0 < \delta < +75, \|b\| > 10$	0.20	VLA-A 36km	36	13500
MITVLA	8.4,4.9	60	$-40 < \delta < +40, \|b\| > 10$	0.2-0.4	VLA-A 36km	12	12400
PMNATC	8.6,4.8	50	$-87 < \delta < -39, \|b\| > 2$	0.8-1.5	ATCA 6km	52	8068
SOURCE IMAGING SURVEYS (VLBI)							
PR	5.01	1300	$+35 < \delta < +70, \|b\| > 10$	1-2mas	GVLB 9000km	41	46
CJ1	4.9,1.7	700	$+35 < \delta < +90, \|b\| > 10$	1-10mas	GVLB 9000km	53	135
CJ2	4.99	350	$+35 < \delta < +90, \|b\| > 10$	1mas	GVLB 9000km	17	193
VSOP	4.90	950	$-90 < \delta < +90, \|b\| > 10$	0.3mas	VSOP30000km	18	289

tical, radio and X-ray spectral signatures indicating orbital motion around a massive central object) for massive dark objects (MDOs) in the centers of nearby galaxies, and these are very likely to be MBHs. AGN & RLAGN are thus expected to develop MDOs during their evolution; during the AGN's 'luminous' phase, these MDOs should actually be MBHs. In the case of radio-loud AGN, there is further evidence for MBHs (persistent, large-scale jets with constant spatial orientation and superluminal cores; huge energy stores in extended lobes; strong, linearly polarized synchrotron radiation; radio variability from days to years). Based on all these observational 'smoking guns', we are led to suspect that all RLAGN harbor not only a MDO, but indeed a MBH.

Where can we find the RLAGN which house these putative MBHs? This is where large-area, statistical, surveys of RLAGN come into play. These surveys come in 4 basic types, and the basic parameters of some of the more important ones are given in Table 1. Depending on the instrument, technique and selection criteria used, the resulting sample sizes range from ~50 to 2 million detected radio sources. Almost all of the surveys listed here are publicly available, with the exception of the interferometer array imaging surveys. Some of these latter large campaign results should become publicly available within the next few years or so.

A potentially powerful key to investigating RLAGN astrophysics is the confrontation of these abundant survey statistics with global models of cosmologically evolving populations of powerful EGRS. On the hypothesis that each one of these AGN in fact houses an evolving MBH engine that is the

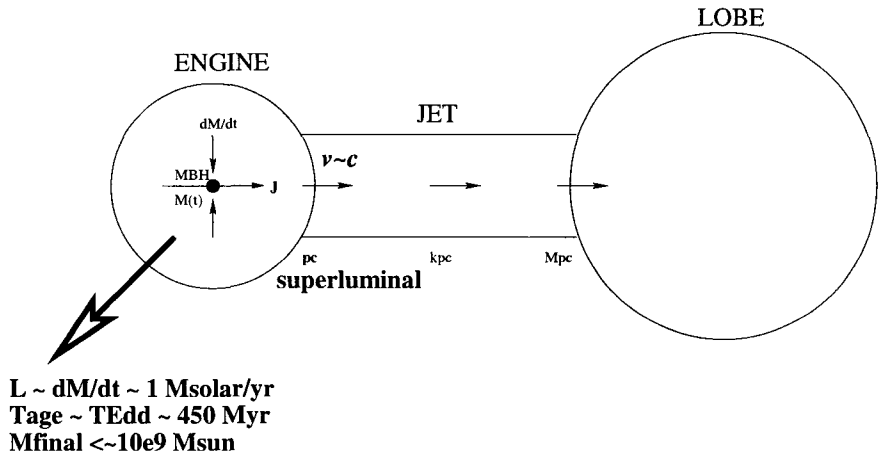


Figure 2: Schematic picture of the basic physics in a radio-loud AGN.

ultimate power source for the distinctive phenomena observed, we may ask the basic question: can we understand all the RLAGN phenomena in terms of models for physically evolving MBHs? The physics of the MBH, of its plasma magnetosphere, and of its attendant accretion disk, will all be crucial links connecting the growth and evolution of the MBH to the observed consequences (jets, lobes, radio spectra) that it drives. Despite the serious lack of redshift information for RLAGN, what is clear so far is that RLAGN are usually found in massive elliptical galaxies (e.g. cD, gE) in the centers of galaxy clusters, and that they undergo strong cosmological evolution [8]. In the high redshift range around $1<z<5$, there was an 'AGN Era' where all AGN were more luminous (luminosity evolution) or more numerous (density evolution), or a mixture of both. As a corollary, we expect to find non-luminous 'dead quasar' remnants of former AGN in the centers of nearby galaxies. There also seems to be a peak in the space density of quasars around $2<z<4$ [44], and a redshift cutoff in the space density of the radio luminosity function (RLF) for powerful sources, somewhere beyond $z>3$ [10]. Detailed identification studies of subsets of various complete radio samples has revealed 2 basic types of EGRS: the 'starbursts' and the 'monsters' [8 19]. We also expect to see size and morphology evolution: jets growing from pc to kpc to Mpc scales; and 'spectral ageing' evolution: the steepening of radio spectra with cosmic time, as the electron population loses its more energetic members. Doppler boosting and inclination effects can complicate the interpretation of the data. In more recent times, the observed

effects of linear size evolution[3,20] and orientation[29] have been studied in detail. The evolving morphologies of powerful, extended, double-lobed radio galaxies [48] have been modelled for the purpose of estimating both intrinsic source parameters (e.g. jet power, density) and also (loose) constraints on the Big Bang (FRW) cosmological parameters. Current work typically focuses on interpreting yet larger and deeper samples of radio sources; there is a great need to compile statistically complete samples with essentially complete redshift information[3], and this remains a critical "bottleneck" problem that must be solved before further understanding of RLAGN becomes possible.

What is the physics behind the engines driving AGN phenomena? An up-to-date review on the theories and observations of AGN may be found in [25]. Kerr (rotating) BH spacetime is reasonably well understood, at least in the classical limit. However, rotating magnetospheres, accretion disks, fluid and plasma jet dynamics, and radiative transfer are much more complicated, and hence less well understood. Further progress in sorting out this dense web of interwoven physics may only be possible via increasingly sophisticated numerical simulations, which themselves are already very technically complex, though slowly improving[23]. Analytic formulations of the BH magnetospheric plasma physics are highly idealized, and there are very few specific solutions, even in these simplified scenarios[1]. Despite these difficulties, a broad understanding has emerged, and this is summarized schematically in Figure 2. The extreme luminosities of AGN, and RLAGN in particular, imply the liberation of a significant fraction of the mass-energy of matter infalling into a central location within the galactic nucleus, at a rate of order $\sim 1 M_{\odot} yr^{-1}$, consistent with Eddington-limited accretion onto supermassive black holes (SMBHs) with M$>\sim 10^6 M_{\odot}$. The Eddington accretion timescale t_{Edd} ~450 Myr is set purely by fundamental constants, and hence all such AGN engines are expected to evolve over similar cosmological timescales; this may be why there is a definite 'AGN Era' ($1<z<5$). Eddington-limited accretion, continued for just a small number of Eddington times, would leave behind fossil SMBHs as large as $10^8 M_{\odot}$ to $10^9 M_{\odot}$; and, indeed, that is what we are finding in most, if not all, nearby bright galaxies at the current epoch. A very interesting and difficult pair of questions to answer, in future, are: (1) What makes some galaxy centers develop AGN, and not others? and (2) What makes some AGN develop into RLAGN, and not others? Regarding question (1), it seems that there may be a continuum scale in BH mass between active and inactive galactic centers [33]. It has long been known that only ~10% of optically bright AGN are also radio-loud. Why are the other ~90% radio-quiet? A possible answer to this key question, involving the fraction of MBHs which are rapidly spinning as a result of galactic mergers, has been suggested in[50]. Piecing together these ini-

tial observations and theories, it seems that normal galaxies, radio-quiet AGN (RQAGN) and RLAGN can all be placed on a rough scale based on the central MBH mass M, which seems to be a major determining factor in the resulting physics and observed phenomena. A specific prediction of this speculative theory is the existence of intermediate mass BHs (IMBHs) in galaxy centers; indeed, candidate IMBHs have already been found.

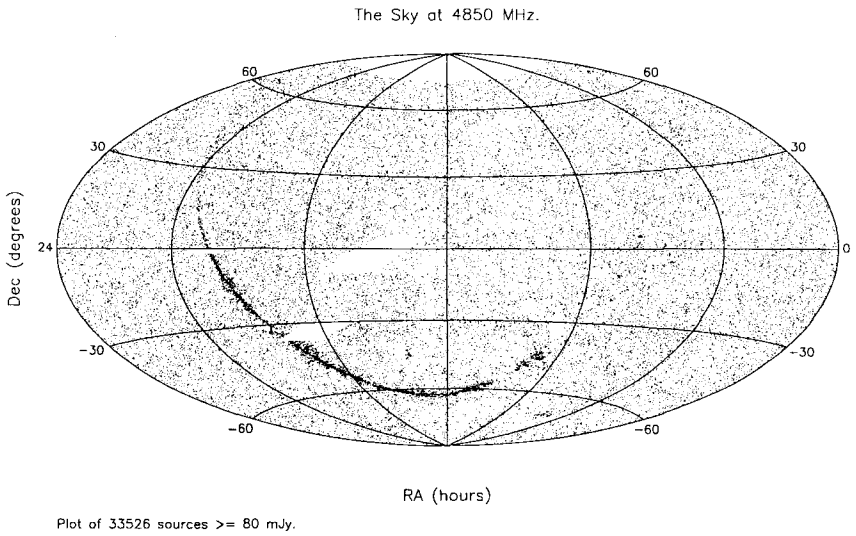


Figure 3: All-sky survey plot of ~33,500 discrete radio sources at 4.8 GHz.

3 MIT Surveys

The MITVLA snapshot survey (1981-1999) was started with the goal of searching for new radio-loud examples of strong gravitational lenses. From 1979 to 1991, the MIT radioastronomy group made 8 all-sky mapping surveys of radio sources at 4.85 GHz (C-Band), using 2 single-dish radiotelescopes: the NRAO Greenbank 91m to scan the northern hemisphere sky, and the ATNF Parkes 64m to scan the southern. These are known as the MIT-Greenbank (MG) [16], and the Parkes-MIT-NRAO (PMN) [51 52] surveys, which resulted in 8 published catalogs. An all-sky 4.85 GHz plot of ~33,500 radio sources, drawn from the northern NRAO 1987 Greenbank (87GB) survey, and from the southern PMN

surveys, is shown in Figure 3. It is clear that the vast majority of radio sources are isotropically distributed in the sky, and hence probably at cosmological distances. In other words, most of the discrete sources in the 4.85 GHz sky map are RLAGN. Such surveys are then very efficient in finding the 'galactic homes' of a vast population of distant MBHs in our universe.

Based on these relatively unbiased all-sky radio source finding surveys, the MIT group started to hunt for radio-loud gravitational lenses; this was the initial scientific motivation for embarking on the MIT Very Large Array (VLA) snapshot campaigns, for which there have been 4 distinct phases over the 19-yr period 1981-1999: (1) 4,100 northern sky MG sources were imaged at 4.86 GHz in 1981-1986; (2) 3,700 MG sources at 8.44 GHz in 1989-1993; (3) 2,100 southern sky PMN sources at 8.44 GHz in 1991-1995; and (4) 2,600 PMN sources at 8.46 GHz in 1998-1999; making an accumulated total of ~12,500 images. The VLA in its largest configuration, A-array, was used in snapshot mode, with typical integration times of 60–120 seconds. Images were obtained with rms noise of ~0.2–0.4 mJy/bm (S/N>~100) and angular resolution ~0.2-0.4 arcseconds. Typical limiting flux densities of selected sources were 50-80 mJy at 4.85 GHz, and the MITVLA samples are statistically representative, but not complete; the individual snapshot target source lists were chosen roughly in order from bright progressing to faint sources, though many had to be omitted due to time and calibration constraints. The resulting accumulated MITVLA sample cannot be easily defined in terms of complete samples, but since the large majority of radio sources were chosen from initially complete samples, the ~12,500 MITVLA snapshots are expected to be statistically representative, at least for the brighter radio sources. These ~12,500 images form ~30% of the estimated total of ~39,000 sources imaged at arc-second resolution in dedicated VLA and ATCA campaigns over the past 20 years. Only ~1,000 of the MITVLA snapshots have been published so far [26 27]. Currently, ~5,000 MITVLA snapshots are readily retrievable from FITS files on DAT tape. Raw VLA archival data for ~9,900 sources are stored on DAT tapes in KAO. In 2002, these 9,900 snapshots will be re-reduced in uniform fashion, so that they may be published and made publicly available. For further details on the MIT single-dish and VLA surveys, as well as on the scientific results which have emerged from the efforts of many MIT students over the past 2 decades, refer to [12 13].

From the finding surveys, the first obvious scientific result that grabs the eye is the remarkably isotropic distribution of EGRS; this has been independently confirmed in many other large-area radio surveys, e.g. NVSS [9]. From the known redshifts of a small fraction of identified MG-VLA sources, there is evidence for a cosmological distribution of radio emitters (0<z<4), which

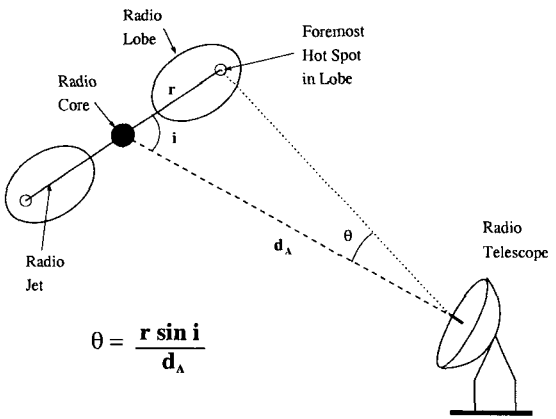


Figure 4: A simple model for kpc-jets in extended radio-loud AGN.

is no surprise. Given this distant population of luminous radio sources, we would theoretically expect ~25 lenses from a complete analysis and follow-up of the MITVLA source sample; so far, 11 lenses have been confirmed, with many more lens candidates still requiring detailed, systematic, follow-up at higher radio resolution, and with ground-based deep optical imaging in good seeing conditions. There is a wide variety of radio morphologies seen in the VLA snapshots; and a systematic classification system, based on the first 1000 MITVLA sources reduced, was published in 1984 [26]. There is also a broad distribution of 'largest angular sizes' (LAS) for the extended radio sources, with a clear majority of relatively small angular size sources ($\theta < 5''$). The MITVLA LAS histogram may be used to place limits on their jet advance (i.e. lobe propagation) speeds (Section 4).

Optical identifications have been published for the first ~1,000 MITVLA sources, and attempted for a further ~2,500. Typically, ~40% of MITVLA sources can be detected to a limiting magnitude of ~20 in the optical POSS plates or the Digitized Sky Survey (DSS). As a general rule, bright, flat-spectrum sources are the easiest to identify optically. A sample of ~120 compact, double-lobed radio galaxies (with angular sizes $0.2'' < \theta_{LAS} < 2.0''$) were extracted from ~4,200 snapshots. These were found to be mostly steep-spectrum radio sources, and hence members of the Compact Steep Spectrum (CSS) class. Deep BVRI imaging to a limiting magnitude R~22, at the MDM 1.3m and CTIO 1.5m, shows these to be red in color; a few spectra reveal low

to intermediate redshifts. These MITVLA CSS sources are probably young galaxies, intermediate in size, and probably in evolutionary development, between the ultra-compact VLBI sources and the 'classic' 3C/4C double-lobed radio galaxies.

There are numerous potential uses and discoveries made possible by the MITVLA database, as well as similar RLAGN image databases (e.g. JVAS, CLASS & PMNATCA), once they are publicly released: new radio calibration sources; new gravitational lenses; new additions to specific classes of RLAGN: such as CSS, GHz-Peaked Spectrum (GPS), high-redshift as well as 'distorted' radio galaxies and quasars. The ~39,000 snapshot images accumulated to date form an indispensable scientific complement to the lower resolution data in the single-dish and sky scanning interferometer surveys (FIRST, NVSS, WENSS & SUMSS), as well as to the ultra-high-resolution VLBI surveys aimed at the central pc-scale cores and jets of RLAGN.

4 Kiloparsec-jet Advance Speeds

A unique contribution from the MITVLA snapshot surveys is the compilation of the 'largest angular size' (LAS) histogram for a spectrally unbiased sample of radio sources (this is not possible for the larger JVAS/CLASS sample of compact, flat-spectrum selected sources). This statistical distribution may be used to set (model-dependent) limits on the growth rate of extended radio structures on kpc-scales, as detailed in [12] [13]. Here, we should be careful to distinguish between the jet head propagation speed $v_{head} = \frac{dR}{dt}$, and the jet fluid flow speed v_{jet} (these are related by the momentum transfer between the jet fluid and the ambient medium). In general, we expect $v_{head} \leq v_{jet}$. A simple geometrical model may be assumed to apply to the extended radio sources in the sample, as shown in Fig. 4.

Various assumptions were made in order to put limits on the jet propagation speed for the MITVLA radio sources: (1) continuous jet outflow (not episodic); (2) uniform distribution of jet ages, with maximum age set by $\sim t_{Edd}$; (3) isotropic distribution of kpc-scale jet axis inclination angle i; (4) Doppler beaming effects are small for pole-on sources (small i); (5) strong cosmological evolution (most sources at z~1); (6) various RW cosmologies ($q_0, H_0, \Lambda = 0$); (7) no instrumental angular bias between $0.''3 - 120''$; and (8) statistically representative sample of MITVLA sources. The major theoretical assumption is that kpc-jets are continuously driven by MBHs accreting at close to the Eddington rate, with maximum age set by t_{Edd} ~450 Myr, *independent of mass.* Statistically, we expect the jets to have an average length of about the average lobe propagation speed $< v_{head} >$ multiplied by the average jet lifetime $< t >$.

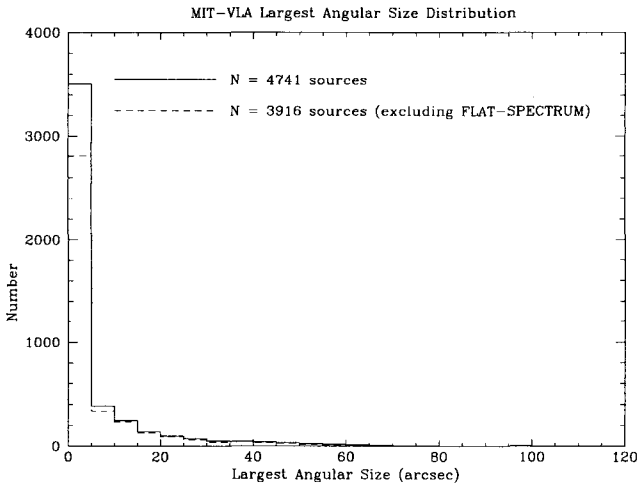


Figure 5: Largest angular size distribution for ~4000 MITVLA survey sources.

A jet advancing at speed v~c, for half an Eddington lifetime t~225 Myr, placed at a typical cosmological distance D~1 Gpc, will subtend an angle of about ~4 degrees. However, the median angular size from Fig. 5 is about 4″, which is 3–4 orders of magnitude smaller! This strongly constrains the average jet advance speed to be in the range $0.0002 < v_{head}/c < 0.1$, with >90% confidence. The main caveat is that if the isotropy and jet continuity assumptions fail, the MITVLA LAS histogram gives limits on projected distance only (~0.22 Mpc for >95% of MITVLA sources). Given this upper limit, relativistic jets could "squirt" for up to ~1 Myr, over kpc scales, and many such "squirtings" could occur in a few Eddington BH lifetimes. Interrupted flows are indeed seen in the 'double-double' radio galaxies, but these are relatively rare cases.

If the kpc-jets are as ultra-relativistic as the VLBI jets, with $\Gamma >\sim 5$, then Doppler boosting would be important for i<35 degrees (~18% of an isotropic population). They would dominate any flux-limited sample, and highly asymmetric sources would then be expected to be the norm. The arm length ratio asymmetries would be as high as ~10, which is *not* observed. This still leaves room for mildly relativistic jets, though.

What does all this mean? From other work, we know that both pc-scale and kpc-scale jets (e.g. in 3C273 and M87) show superluminal, or at least

mildly relativistic, moving features. However, these relativistic flows probably arise from the bulk motion of jet fluid v_{jet}, and not from a jet advance or "pattern" speed v_{head}. Previous work on the arm length asymmetries of collinear double radio sources, and using 'spectral ageing' analyses of the electron populations near the terminal hot spot 'backflow'[30], all place firm statistical limits of ~0.2–0.3c on the jet head advance speed v_{head}. More recent work has considered the observed asymmetries and Doppler beaming effects arising from initially relativistic jet bulk flow motion. These show that there is significant deceleration of initially relativistic flows, as they emerge out onto the kpc-scales. Furthermore, we theoretically expect v_{head} to be less than v_{jet}. Due to momentum transfer, the relativistic pc-scale jets may become non-relativistic on kpc scales. In the case of 'heavy' jets ($\rho_{jet}/\rho_{amb} \gg 1$), v_{head} may become comparable to v_{jet}, and so the flow speed is then definitely not relativistic $v_{jet}/c \ll 1$. This may be true for the 'edge-darkened', subsonic FR I jets. However, if the jets are 'light', then it is quite possible, though by no means certain, that the flow speed is relativistic. This may be the case for the 'edge-brightened', supersonic, FR II jets. There may then be a connection between jet speed and morphology; a theory has been proposed that links the kpc-scale morphology to the inner workings of the central engine and its outflow, operating via a 'magnetic switch' mechanism[34]. However, no strong observational constraint on the flow speed v_{jet} is possible, without independent knowledge of the jet density contrast with the ambient medium. In the future, more attention needs to be paid to potentially revealing connections like these. In particular, the correlation between radio source age, morphology and jet dynamics needs to be more fully investigated and clarified. The main lesson to be learnt is that we can now start to relate statistical distributions of observable RLAGN parameters to the physics of their central engines.

5 Black Hole Physics

AGN range over many orders of magnitude in time, space, wavelength, power and energy. This attests to the fact that the observed phenomena come from a complicated, interwoven network of physical processes operating over both small and large regions, and over short and long times. The causal progression of physical phenomena in RLAGN is clear: black hole; magnetosphere and accretion disk; pc to kpc to Mpc jets; hotspots; lobes & bridges; then the IGM. The connection between theories and observations starts with the physics of the growth and evolution of the central MBHs. In modelling the growth and evolution of AGN, we must start off with the formation and physics of the central MBH; we focus specifically on the MBH here.

Classical general relativity is well understood for Schwarzschild and Kerr-Newman BHs. It is believed that rotation and magnetic fields are essential for extracting power from the BH [2]. What is required is to model the Kerr BH plasma magnetosphere, as well as the magnetized, turbulent accretion disk, and the launching of the VLBI pc-scale jet at its base. This is far from easy, and many simplifying assumptions still continue to be made even in recent theoretical and computational research: 1) single-component plasma; 2) quasi-neutral plasma; 3) axisymmetry; 4) magnetic forces dominate inertial ones (inertia free); 5) infinite conductivity (ideal-MHD conditions: 'frozen-in flux'); 6) degenerate electromagnetic fields (which cannot exist in charge accelerating 'gaps'); 7) (Lorentz) force-free; 8) (time) stationary; and 9) Newtonian regime.

In the past, almost all studies of the plasma magnetospheres of compact objects have retained all assumptions 1–9 above (in particular: axisymmetric, stationary, ideal-MHD plasma physics in the Newtonian regime). This is quite natural as a first step towards a broad understanding of the problems involved. However, it is a foregone conclusion that many of the results of these preliminary attempts will be largely inapplicable to real accreting MBH systems.

The 'Blandford-Znajek' (BZ) black hole magnetospheric power extraction process [2] was one of the first (1977) analytic attempts to define a general relativistic theory for axisymmetric, stationary, ideal-MHD, BH magnetospheres. Since then, there have been many analytic refinements made [31 46] based on this initial physical idea. For a recent review of axisymmetric, stationary, ideal-MHD, analytic models of BH magnetospheres and outflows, see [1].

It is a well-known fact that the luminosity output of AGN can be variable in time, over many different timescales [47]. If these obvious variabilities are driven by central MBH physics, then it is very likely that the power extraction from the central MBH cannot be time stationary. The first time-dependent, analytic models for axisymmetric, ideal-MHD BH magnetospheres [37] were proposed in 1988. The long-term, secular (i.e. cosmological) evolution of accreting MBHs in AGN was analyzed in [38]. The general theme of time-dependent AGN MBH magnetospheres, disks and outflows, and how they may cause the observed time variability in AGN luminosity, has been addressed in [39]. There have been further attempts to drop more assumptions, such as axisymmetry, together with stationarity, but the resulting theory seems no more approachable than before. It seems that the future pursuit of a more realistic understanding of these inherently complex physical scenarios must be continued by way of numerical simulations.

To better understand RLAGN variability in the core region, we will have to first understand time-dependent plasma physics phenomena, such as: magnetic

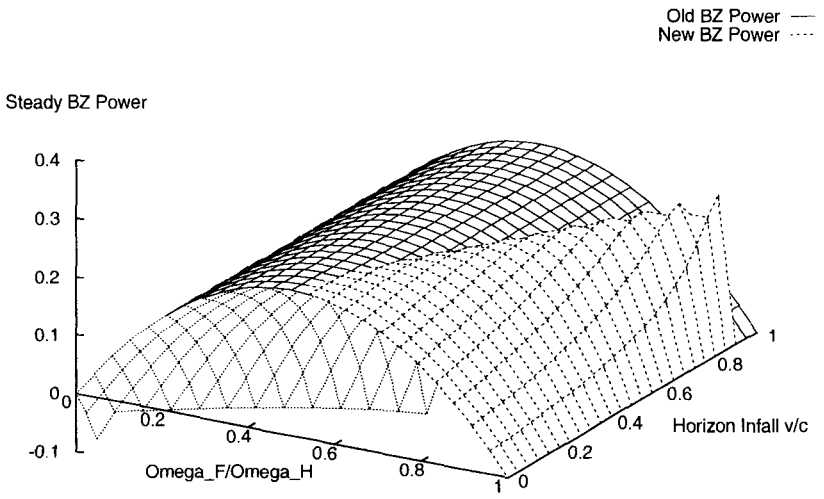


Figure 6: Comparison plot of the steady component of the Blandford-Znajek (BZ) power output from a *non-stationary*, axisymmetric, force-free, ideal-MHD, Kerr black hole magnetosphere (the surface drawn with dashed lines), with the corresponding BZ power for the time-stationary case (solid lines). The independent variables are the ratio of the magnetic field to black hole angular velocity, and the plasma radial infall velocity; power is plotted in arbitrary units.

reconnection, coronal magnetic ejections, dynamo theory and winds. Axisymmetric simulations (2.5.D) have been performed for Kerr BH magnetospheres [23 35], assumed force-free. These have indeed developed jets along the polar axis, after some ~100 dynamical rotation periods of the central BH. There also appear 'Torsional Alfven Wave Trains' (TAWTs), which seem to be important in the transport of plasma, magnetic field and angular momentum away from the hole, magnetosphere and disk. Such a steady, magnetic 'helix' seems to be a persistent and generic feature in many of the recent simulations. A 'magnetic switch' mechanism [34] may determine the jet speed in RLAGN; if real, this may be an important physical mechanism which affects the radio source morphology on much larger (kpc) scales. For the latest developments in the rapidly growing area of BH GRMHD physics, theories, simulations, and their relation to observations, see [7 42].

The last 4 assumptions (6-9) above have each been dropped in various papers, but not all together at the same time. An important one to shed is *stationarity* for the BH magnetosphere; this is especially critical, in view of the

fact that AGN are known to be variable in their observed properties, over all timescales from years down to minutes. The analytic formalism developed first in 1989[37 39], for axisymmetric, *non-stationary*, ideal-MHD, force-free Kerr BH magnetospheres, has since been reclarified and recast into an improved set of equations, and the associated 'Transfield Equation' has been derived[28 40].

As an example of a preliminary result from this non-stationary analysis, the steady-state BZ power output has been calculated as the time stationary limit of the general non-stationary case. The original (1977) BZ power output formula was found to be $P_{BZ,stat} \propto x(1-x)$, whose value is determined by the sole parameter $x = \Omega_F/\Omega_H$, the ratio of the magnetic field and BH angular velocities. However, the steady part of the *non-stationary* power output has a dependence on an extra parameter, the radial infall velocity at the BH horizon: $y = -v_{rad}/c$. The dependence on a 2nd parameter y is clearly shown in Fig. 6. The derivation of these results is based on the same equations that yield the new 'Transfield Equation' for axisymmetric, non-stationary BH magnetospheres[28 40]. There is a divergence to unphysical (negative) power output for large values of the infall velocity $y > x$; this can only be avoided if the angular velocity ratio $x \to 1$, in which case the radial infall velocity will approach the expected free-fall limit c. What this means is that the magnetic field lines fall onto the BH horizon at the speed of light; and, at this point, the angular velocity of the hole and the field should be equal $(\Omega_F/\Omega_H)_{horizon} \to 1$. If this condition is not met, it seems that the power output would assume unphysically divergent negative values. Hence, it is suggested that the appropriate boundary conditions near the horizon are that: $v_{rad}/c \to -1$, and that $(\Omega_F/\Omega_H)_{horizon} \to 1$.

From this axisymmetric, *non-stationary* analytic model, it is suggested that there is 'no slip' between the field and the horizon ($\Omega_F = \Omega_H$). However, this may only be possible when the horizon resistivity $\eta = 0$, i.e. with ideal-MHD conditions. On the other hand, the original (stationary) BZ result that there be *zero* power for $\Omega_F = \Omega_H$ seems unnatural, and is in any case expected to change within a more realistic, *non-stationary* analysis. This suggested 'no slip' condition for the BH magnetic field says nothing about its behaviour away from the horizon; indeed, we generally expect Ω_F to decline towards the Keplerian accretion disk velocities, as one progresses further away from the event horizon.

A critical set of questions still outstanding in the entire investigation into AGN in general, is concerned with: (1) how MBH "seeds" form, (2) how they grow to super-stellar mass, and (3) how their physical evolution is related to that of their host galaxies. Somehow, the answers to these questions are related to the recently discovered 'BH mass vs. bulge luminosity' and 'BH

mass vs. central velocity dispersion' relations; these strong correlations have been observed to occur in both active and inactive galaxies [33], which may imply a simple, unified formation scheme for MBHs in all types of galaxies. If these current speculations about the observations prove to be correct, then we may be seeing a first hint that there is actually a continuum connecting normal bright galaxies to RQAGN to RLAGN; and that the main governing parameter determining the physical properties and evolution of their central regions is the mass M of the central MBH. However, at this point, this last suggestion is still a pure speculation.

In summary, much more work is required to understand these complex MBH physics. In general, the astrophysics of MBH growth and evolution needs to be much better understood (convenient assumptions need to be removed from the analytical theory, and more realistic simulations need to be done). Many of the BH high energy plasma physics processes also occur in other accreting compact systems, even though the details may be different. It may be very rewarding to try to apply lessons learnt from one type of highly energetic system, to another of a similar nature. For example, 'magnetic reconnection' in the BH magnetosphere and accretion disk will need to be investigated; perhaps, for this piece of the puzzle, inspiration can be drawn from the abundant observations of plasmas in the Sun, and elsewhere in our solar system.

6 Future Work

In the case of observations, high-resolution images of RLAGN from the MIT VLA and other (CLASS, PMNATCA, VSOP) surveys should be published, and made publicly available, over the next few years. Attention needs to be paid to the statistical completeness of these large RLAGN samples. While not all of them may be complete, suitably chosen subsets of them may be reasonably complete, so these sub-samples could be used for statistical analyses (as long as they are large enough). Searching for statistically complete radio imaging samples will become possible once more of these images are made available. Optical identifications need to be attempted for radio component positions drawn from high-resolution interferometer imaging. This is possible with online Internet databases, such as DPOSS, DSS, APS, APM and (SUPER)COSMOS. Cross-identifications with non-optical surveys (e.g. IRAS, ISO, ROSAT and EGRET) will also be necessary to constrain RLAGN physics, via the compilation of spectral energy distributions (SEDs). The NASA Astrophysics Data Archives will be especially useful in this regard. Over the past decade, there has been an explosion of Internet web browser usage, and one of

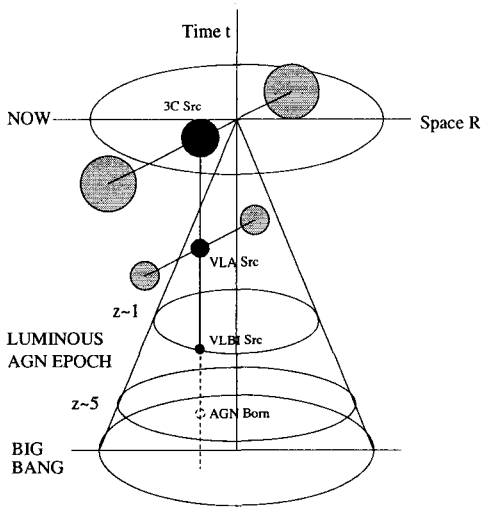


Figure 7: Spacetime evolution of MBHs and their jets in radio-loud AGN.

the results of this new phenomenon has been the creation of prototype 'Virtual Observatories'. The US Decadal Review recommendations for 2000–2010 proposed a National Virtual Observatory (NVO) along these lines, and a future Global Virtual Observatory (GVO) has also been discussed. Even today, the available Internet astronomical resources are vast [21], though mostly still not organized for truly automated processing. The redshift sparsity problem for RLAGN may be solved by wide-field fiber spectroscopy (e.g. using 2dF, 6dF) or, better and more easily, by reaping photometric redshifts, if possible from a large-area optical survey such as SDSS. Where data are gathered on the optical counterparts of RLAGN, the selection effects would need to be accounted for in the statistical modelling of the joint radio-optical observation selection process. Finally, for RLAGN, the long-term evolution of their radio luminosities and jet lengths may be derived from their estimated luminosity functions (LFs), and linear size functions (SFs), once the redshifts become available for sufficiently large, and reasonably complete, samples [3].

In the case of the theories, many clues and insights should be forthcoming from a detailed, global consideration of all the large, statistical samples and high-resolution images of RLAGN which will become publicly available over the next few years. To understand the physical behaviour of the growth and evolution of the MBHs, and their derivative radio structures, detailed,

global Monte-Carlo physical modelling of the population of individual AGN sources is required. Many various individual aspects of RLAGN population modelling have already been proposed in the literature, for: MBHs, accretion disks, pc-scale jets, and especially the kpc-scale jets & lobes[3 29 48]. Bearing in mind the central role played by growing and accreting MBHs, there are clearly identifiable functions for which physical models need to be proposed. Some of the more important ones are the following (here, a is the normalized BH spin, and we are careful to distinguish between the 'cosmic' time τ and the source 'lifetime' t):

- Creation Rate of Galaxies, $C_{gal}(\tau)$;
- Creation Rate of MBHs, $C_{mbh}(M, a, \tau)$;
- Time-dependent MBH Mass Function, $\Phi_{mbh}(M, a, \tau)$;
- Rate of Change of MBH Mass, $\dot{M}(M, a, t)$;
- Power Output, $P(M, a, t)$;
- Radio Luminosity, $L_{radio}(M, a, t)$;
- Jet Propagation Speed averaged over kpc scales, $< V_{head}(M, a, t) >$.

All of these need to be put into Monte-Carlo models of the evolving population of MBHs and their derivative structures. Fig. 7 is a schematic diagram showing the basic idea behind this proposal. Time permitting, a basic model will be attempted along these lines, after the reduction and analysis of the MITVLA snapshot survey data is completed. To help set up suitable simulations, an important ingredient will be the modelling of the long-term evolution of the MBH population and their associated physical parameters (M, a, t); as well as of the radio luminosity output, and the dimensions and morphologies of the jets and lobes. The basic theories set out in certain papers addressing these long-term evolutionary processes[20 38] will serve as indispensable starting points from which to explore possible theoretical interpretations of the large body of available data on radio-loud AGN.

Acknowledgments

I acknowledge the 2nd KIAS Astrophysics Workshop organizers (C.-H. Lee and H.-Y. Chang) for their invitation to give a talk; and to Dr. S. Hyung (KAO) and Dr. S.-J. Park (KAO) for their support of my work. This research was funded in part by the Korea Astronomy Observatory (KAO), and by the Korean government (KOSEF/KOFST) 'Brainpool' program.

References

1. Beskin, V.S., 1997, Russ. Acad. of Sci. Physics-Uspekhi **40(7)**: 659–688.
2. Blandford, R.D. & Znajek, R.L., 1977, *MNRAS* **179**: 433–456.
3. Blundell, K.M., Rawlings, S. & Willott, C.J., 1999, *AJ* **117(2)**: 677–706.
4. Bock, D.C.-J., Large, M.I. & Sadler, E.M., 1999, *AJ* **117(3)**: 1578–1593; *http://www.astrop.physics.usyd.edu.au/SUMSS.*
5. Bremer, M.N. *et al.*, eds., 1998, *Observational Cosmology with the New Radio Surveys*, Canary Islands Workshop, Tenerife, Spain, 13–15 Jan 1997 (Kluwer: Holland).
6. Chakrabarti, S.K., ed., 1999, *Observational Evidence for Black Holes in the Universe*, Calcutta, India, 11–17 Jan 1998 (Kluwer: Holland).
7. Chang, H.-Y. *et al.*, eds., 2001, *Explosive Phenomena in Astrophysical Compact Objects*, 1st KIAS Astrophysics Workshop, KIAS, KIST, Seoul, Korea, 24–27 May 2000, AIP Conf. Proc. v. 556 (AIP: USA).
8. Condon, J.J., 1989, *ApJ* **338**: 13–23.
9. Condon, J.J. *et al.*, 1998, *AJ* **115(5)**: 1693–1716; *ftp://cdsarc.u-strasbg.fr/cats/VIII/65; http://www.cv.nrao.edu/nvss.*
10. Dunlop, J.S. & Peacock, J.A., 1990, *MNRAS* **247**: 19.
11. Ferrarese, L. & Merritt, D., 2000, *ApJL* **539**: L9–L12.
12. Fletcher, A. *et al.*, 1998, in *Observational Cosmology with the New Radio Surveys*, eds. M.N. Bremer *et al.* (Kluwer: Holland), pp. 249–254.
13. Fletcher, A.B., 1998, MIT Ph.D. thesis, *The Nature and Evolution of I. Extragalactic Radio Sources, and II. Young Stellar Clusters.*
14. Gebhardt, K. *et al.*, 2000, *ApJL* **539**: L13–L16.
15. Gregory, P.C. *et al.*, 1996, *ApJS* **103**: 427; *ftp://cdsarc.u-strasbg.fr/cats/VIII/40; http://www.cv.nrao.edu/jcondon/gb6ftp.html*
16. Griffith, M.R. *et al.*, 1991, *ApJS* **75**: 801.
17. Henstock, D.R. *et al.*, 1995, *ApJS* **100**: 1–36; *http://www.astro.caltech.edu/tjp/cj.*
18. Hirabayashi, H. *et al.*, 2000, *PASJ* **52**: 997–1014; *http://www.vsop.isas.ac.jp/survey; http://www.vsop.isas.ac.jp/vsopsymp/papers/index.html.*
19. Jackson, C.A. & Wall, J.V., 1999, *MNRAS* **304**: 160–174.
20. Kaiser, C.R. *et al.*, 1997, *MNRAS* **292**: 723.
21. Kidger, M.R., Pérez-Fournon, I. & Sánchez, F., eds., 1999, *Internet Resources for Professional Astronomy*, 9th Astrop. Winter School, 17–28 Nov 1997, Canary Islands, Spain (Cambridge: UK).
22. King, L.J. *et al.*, 1999, *MNRAS* **307(2)**: 225–235;

ftp://ftp.aoc.nrao.edu/pub/sources.jvas;
http://www.jb.man.ac.uk/~ceres1/ceres1.html.
23. Koide, S. *et al.*, 2000, *ApJ* **536**: 668.
24. Kormendy, J., 2001, *RevMexA&A Ser. Conf.* **10**: 69–78.
25. Krolik, J.H., 1999, *Active Galactic Nuclei: From the Central Black Hole to the Galactic Environment* (Princeton: USA).
26. Lawrence, C.R. *et al.*, 1984, *ApJL* **278**: L95–L98.
27. Lawrence, C.R. *et al.*, 1986, *ApJS* **61**: 105–157.
28. Lee, C.-H. *et al.*, eds., 2002, *Current High-Energy Emission around Black Holes*, 2nd KIAS Astrophysics Workshop, KIAS, KIST, Seoul, Korea, 3–8 Sep 2001 (World Scientific), these proceedings.
29. Lister, M.L. *et al.*, 1994, *ApJ* **427(1)**: 125-133.
30. Liu, R., Pooley, G. & J.M. Riley, 1992, *MNRAS* **257**: 545–571.
31. Macdonald, D. & Thorne, K.S., 1982, *MNRAS* **198**: 345–382.
32. Magorrian, J. *et al.*, 1998, *AJ* **115(6)**: 2285–2305.
33. McLure, R.J. & Dunlop, J.S., 2001, *astro-ph/0108417.*
34. Meier, D.L. *et al.*, 1997, *Nature* **388**: 350–352.
35. Meier, D.L., Koide, S. & Uchida, Y., 2001, *Science* **291**: 84–92.
36. Myers, S.T. *et al.*, 2001, in *Gravitational Lensing: Recent Progress and Future Goals*, ASP Conf. Ser. v. 237, eds. T.G. Brainerd & C.S. Kochanek (ASP: USA), p. 51; *http://www.aoc.nrao.edu/~smyers/class.html; http://www.jb.man.ac.uk/~ceres1/ceres1.html.*
37. Park, S.J. & Vishniac, E.T., 1989, *ApJ* **337**: 78–83.
38. Park, S.J. & Vishniac, E.T., 1990, *ApJ* **353**: 103–107.
39. Park, S.J., 2000, *JKAS* **33(1)**: 19–28.
40. Park, S.-J., 2002, in *Current High-Energy Emission around Black Holes*, 2nd KIAS Astrophysics Workshop, KIAS, KIST, Seoul, Korea, 3–8 Sep 2001, eds. Lee, C.-H. *et al.*, (World Scientific), these proceedings.
41. Pearson, T.J. & Readhead, A.C.S., 1988, *ApJ* **328**: 114–142; *http://www.astro.caltech.edu/~tjp/cj/.*
42. Punsly, B., 2001, *Black Hole Gravitohydromagnetics* (Springer: USA).
43. Rengelink, R.B. *et al.*, 1997, *A&AS* **124**: 259–280; *ftp://cdsarc.u-strasbg.fr/cats/VIII/62; http://www.strw.leidenuniv.nl/~dpf/wenss.*
44. Shaver, P.A. *et al.*, 1996, *Nature* **384**: 439–441.
45. Spinrad, H. *et al.*, 1985, *PASP* **97**: 932–961.
46. Thorne, K.S., Price, R.H. & Macdonald, D.A., eds., 1986, *Black Holes: The Membrane Paradigm* (Yale: USA).
47. Ulrich, M.-H., Maraschi, L. & Urry, C.M., 1997, *ARA&A* **35**: 445–502.
48. Wan, L., Daly, R.A. & Guerra, E.J., 2000, *ApJ* **544**: 671–685.

49. White, R.L. *et al.*, 1997, *ApJ* **475**: 479-493;
ftp://cdsarc.u-strasbg.fr/cats/VIII/59;
http://sundog.stsci.edu/top.html.
50. Wilson, A.S. & Colbert, E.J.M., 1995, *ApJ* **438**: 62–71.
51. Wright, A.E. *et al.*, 1996, *ApJS* **103**: 145.
52. Wright, A.E., 1996, in IAU Symp. 175, *Extragalactic Radio Sources*, 10–14 Oct 1995, Bologna, Italy, eds. R.D. Ekers *et al.* (Kluwer: Holland), pp. 507–512; *ftp://ftp.atnf.csiro.au/pub/data/pmn;*
http://www.parkes.atnf.csiro.au/databases/surveys/pmn/casouth.pdf.
53. Xu, W. *et al.*, 1995, *ApJS* **99**: 297–348;
http://www.astro.caltech.edu/tjp/cj.

ASCA AND RXTE OBSERVATIONS OF THE ACCRETION DISK IN THE X-RAY BINARIES

T. DOTANI
3-1-1 Yoshinodai, Sagamihara, Kanagawa 229-8510
E-mail: dotani@astro.isas.ac.jp

We have analyzed the ASCA and RXTE data of the X-ray binaries to study the structure of the accretion flow in the vicinity of the compact object. We focused on the black hole binaries, especially the superluminal jet sources, and the unique neutron star binary, the Rapid Burster. We found two distinct states in the disk emission from GRS 1915+105. One state corresponds to the standard α-disk, while the other state has a high disk temperature and an apparent super-Eddington luminosity. We identify this as the so-called slim disk. Quick and repeated transitions between the slim disk and the α-disk may be the origin of the violent time variations and the ejection of jets in GRS 1915+105. Presence of the slim disk was also indicated in GRO J1655–40 and XTE J1550–564 when the source luminosity was high. Based on the similarities of the light curves between GRS 1915+105 and the Rapid Burster, we have studied the ASCA data of the Rapid Burster during the type II burst activities to look for the evidence of the slim disk. We found that the disk parameters were more or less constant through the burst cycle, while the blackbody emission, which originated from the neutron star surface, showed large changes. This result may be explained if the slim disk is developed during the type II bursts. We discuss on various observational indications of the slim disk in bright low-mass X-ray binaries including a black hole or a weakly magnetized neutron star.

1 Introduction

X-ray binaries are the most powerful X-ray sources in Galaxy. They are close binary systems consisting of a compact object (a neutron star or a black hole) and a normal star. Mass accretion on to the compact object from the normal star, mostly through the Roche-lobe overflow, powers the X-ray emission. It is widely accepted that the accreting matter overflowed from the Roche-lobe takes a geometrically thin, optically thick structure,[22] the so-called α-disk, far from the compact object. However, nature of the accretion flow near the compact object, where most of the high energy emission is produced, is not well understood yet.

It is suggested from theoretical studies that the accretion flow may take several different forms near the compact object. One possibility is that the α-disk directly extends down to the vicinity of the compact object. Because the viscous heating balances the local radiative cooling in the α-disk, its X-ray emission is approximated by the superposition of a blackbody emission of var-

ious temperature; this is the so-called multi-color disk model.[17] X-ray emission from the bright low-mass X-ray binaries were found to include emission well approximated by the multi-color disk model. When we take the advection cooling into account, another solution of the accretion flow is possible.[10,1,19] Such an accretion flow is called an advection dominated accretion flow (ADAF), and two types of solutions, optically thin ADAF and optically thick ADAF, are known. The latter is sometimes called a slim disk. Because only a part of the gravitational potential energy is released as radiation in the slim disk, temperature distribution of the disk is different from that of the α-disk. To approximate the X-ray emission from the slim disk, the so-called p-free disk model is sometimes used.[8] Here, p is the parameter which represent the radial dependence of the disk temperature ($T(r) \propto r^{-p}$). Note that $p = 3/4$ for the α-disk. The slim disk can be hotter and more luminous than the α-disk, because X-ray emission is possible from inside the last stable orbit (3 times the Schwarzschild radius for a non-rotating black hole). Optically thin ADAF produces only a small amount of X-ray emission, and may be developed when the mass accretion rate is rather small.

Because high-quality X-ray data from the X-ray binaries have been accumulated by ASCA, RXTE, and other satellites, it becomes gradually possible to identify different types of accretion flow observationally. In this paper, I summarize the current status of observational studies of the accretion flow in the bright X-ray binaries containing a black hole or a weakly magnetized neutron star. This paper is based on the works by Yamaoka (2001)[28] (GRS 1915+105), Kubota (2001)[11] (GRO J1655–40, XTE J1550–564), and Mahasena (2001)[13] (the Rapid Burster).

2 Past Observations

2.1 Black Hole Transients

A black hole in a binary system with a low-mass companion is always observed as a transient source. It shows an outburst every several decades, which is caused by the episodic increase of the mass accretion on to the black hole probably due to the disk instabilities. Black hole transients are especially suited to investigate the structure of the accretion flow, because the mass accretion rate on to the black hole decreases gradually with the decay of the outburst.

Several black hole transients were observed with the Ginga satellite, and the changes of the energy spectra were traced through the decay of the outbursts.[26] Energy spectra of the black hole binaries basically consist of two components, a multi-color disk component and a power-law component. Some other struc-

tures, e.g. an iron emission line and a reflection structure, are sometimes overlayed on the continuum spectrum. When the energy spectra of the black hole transients were analyzed with the two-component model, it was found that the inner radius of the multi-color disk emission stayed almost constant through the decay of the outbursts.[26] Furthermore, the inner disk radii were consistent to $3R_s$ (R_s is the Schwarzschild radius of the black hole), in which a stable Kepler orbit does not exist for a non-rotating black hole. These results indicate that the α-disk extends down to the last stable orbit ($= 3R_s$) when the mass accretion rate is relatively high.

2.2 Micro-Quasars and Ultra-Luminous X-ray Sources

A new class of black hole transients has been discovered in 1990s. They accompanies superluminal jets, and are sometimes called micro-quasars. Examples are GRS 1915+105[15] and GRO J1655–40.[7] X-ray emission from these sources also consist of a multi-color disk component and a power-law component. However, the disk parameters are often found to take unusual values; a high disk temperature and a small inner disk radius. The disk (color) temperature sometimes exceeds 2 keV. Because the disk temperatures observed with Ginga from the black hole transients were at most ~1 keV, the disk temperatures of the micro-quasars are unusually high. Several possibilities have been discussed to explain the high disk temperature, including a rapid spin of the black hole.

A high disk temperature has been observed from another kind of black hole binaries, i.e. ultra-luminous X-ray sources (ULXs). Presence of bright X-ray sources (as luminous as 10^{40} erg/sec) has been known in nearby galaxies. These sources are considered to be black hole binaries accreting near the critical rate. Because of the high luminosities, they are called ULXs. Spectral properties of the ULXs were studied with the ASCA satellite. The black hole in the ULX should have a large mass (~100 $M_\odot$) as suggested by its large luminosity. This means that the inner disk temperature is relatively low. However, the observed disk temperatures were rather high. The results are summarized in figure 1. As seen in the figure, super-critical mass accretion rate is necessary to explain the high disk temperature of some sources.

The high disk temperatures we obtained from the micro-quasars and the ULXs are not easily understood in the framework of the α-disk. This has motivated us to study the nature of the accretion flow in X-ray binaries in detail. In the next section, new results obtained from the micro-quasars and the neutron star binary are presented.

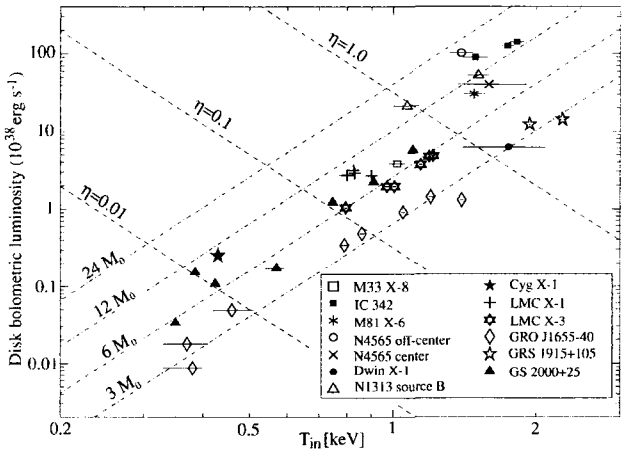


Figure 1: Disk parameters (the highest color temperature and the bolometric luminosity) deduced by fitting the multi-color disk model to the energy spectra of various X-ray sources are shown. The sources analyzed are the ultra-luminous X-ray sources in nearby galaxies and the black hole X-ray binaries in Galaxy. Four dot-dashed lines labeled by the black hole mass represent constant-mass grids when the α-disk extends down to $3R_s$. η represent a normalized mass accretion rate by the Eddington limit. Adopted from Makishima et al. (2000).[14]

3 Micro-quasars

3.1 GRS 1915+105

GRS 1915+105 is an X-ray transient discovered by GRANAT in 1992,[4] and was later identified as the first superluminal jet source in Galaxy. The source is notorious for the violent time variations, which may be related to the ejection of the relativistic jets. We used the RXTE archive data obtained during January 1999 and May 2000, and the ASCA data during 1994 and 2000 for the present analysis. There were in total 164 sets of RXTE data and 7 sets of ASCA data. Among the ASCA data, we used only those which had the simultaneous RXTE data, because the higher energy data were important to constrain the spectral parameters.

Before we performed the spectral fitting, we classified the data appropriately based on the spectral and timing properties. When the time variations were small, the data sets were classified into the soft, hard, and very high states using the count rates and their ratios in several energy bands. When the time variations were large, two types of variations were observed in the RXTE data. One showed relatively regular variations with a large amplitude

continuously, while the other had intermittent intervals of a low flux with a slow rise during the large amplitude, regular variations. Because these time variations accompany large changes in the energy spectra, we divided the data into subsets for these variable state data. The higher and lower flux intervals of the former time variations are separated and are referred to as the peak and the valley, respectively. In the case of the latter time variations, we call the higher and the lower flux intervals in the regular variations as oscillation-high and oscillation-low, respectively, and the low flux intervals as the rise.

When we calculated the energy spectra, some of them were found to have a dominant hard component. This means that we need to know the shape of the hard component accurately to estimate the parameters of the disk emission correctly. For this purpose, we utilized the time variations of the source. Because the rapid time variabilities are in general dominated by the hard component, energy dependence of the variability amplitude is expected to reflect its spectral shape. We calculated the energy dependence of the variability amplitude, i.e. frequency-resolved energy spectrum, for several different frequency ranges. We found that the frequency-resolved energy spectra had a break around 7 keV, and were well approximated by a broken power law, when the frequency range was higher than a few Hz. The photon index was 1.5–2.0 below $\sim$7 keV and was 2–3 above $\sim$7 keV.

Thus we used a model consisting of a multi-color disk model and a broken power law to fit the energy spectra of each data set. Most of the energy spectra were well reproduced by this two component model, but some showed significant residuals when the soft component was strong. This means that the multi-color disk model does not fit the data well. We found that the p-free disk model was required when the source was in the soft state with $p = 0.49 \pm 0.02$, while the multi-color disk model was adequate for the other states.

We carried out spectral fitting once again using the p-free disk model, instead of the multi-color disk model, for the soft state data. The parameters of the disk emission obtained are shown in figure 2. Presence of two branches is clear from the figure. In one branch, the disk temperature is relatively low and the bolometric luminosity is below the Eddington limit, which is calculated for the α-disk extending down to $= 3R_s$. On the other hand, the other branch corresponds to the high disk temperature and the bolometric disk luminosity exceeds the Eddington limit. GRS 1915+105 can switch the two branches quickly in the variable state. Because the two branches are well separated in the figure, it is considered that the accretion disk takes completely different structure in these two branches.

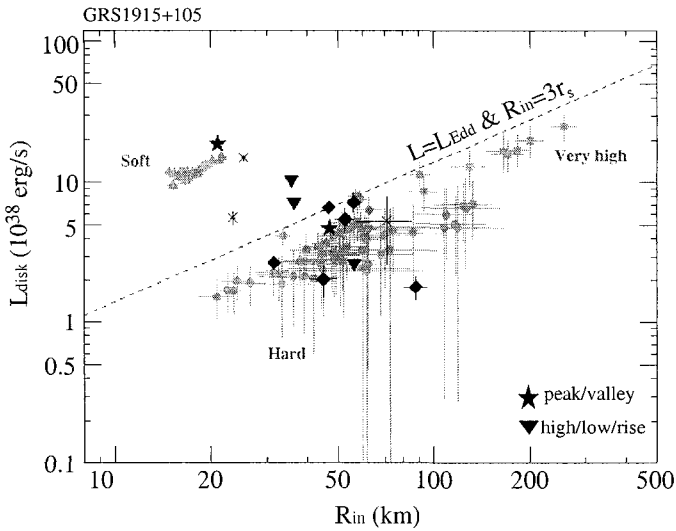


Figure 2: Best-fit disk parameters obtained for GRS 1915+105. A two-component model consisting of the multi-color disk model and a broken power law was used to fit the data. However, we needed to use the p-free disk model, instead of the multi-color disk model, for the soft state data to get an acceptable fit. Grey points are those obtained from the RXTE data, and the black points from the simultaneous fitting of the RXTE and the ASCA data. We assumed the source distance of 12.5 kpc and the inclination angle of 70°. The broken line indicates the relation between the inner disk radius and the bolometric disk luminosity for a non-rotating black hole, when the α-disk extends down to 3 times the Schwarzschild radius. Adapted from Yamaoka (2001).[28]

3.2 GRO J1655–40

GRO J1655–40 (Nova Sco 1994) is an X-ray transient discovered by BATSE in 1994,[29] and is also an superluminal jet source. The system parameters of GRO J1655–40 are well determined from the kinematic analysis of the superluminal jets. Mass of the central black hole is estimated to be 7.02 ± 0.22 $M_\odot$, inclination angle 69.5 ± 0.08 deg, distance to the source 3.2 ± 0.5 kpc.[9,20]

We used the ASCA data obtained on August 15, 1995, and February 26, 1997, and the RXTE data obtained between April 1996 and September 1997 for the present analysis. We extracted energy spectra from each set of data, and fitted a model consisting of a multi-color disk model and a power law. The fit was acceptable for most of the data except for those in the very high state. Although the total flux was almost constant, clear anti-correlation was observed between the flux of the multi-color disk component and that of the

power law component in the very high state. The anti-correlation was not due to the fitting procedure, because it was noticed even in the raw count rate data. As discussed in Kubota *et al.* (2001),[12] the anti-correlation may be due to the Compton up-scattering of the soft photons to the hard component. We followed their analysis and included another model, a Comptonized blackbody, to represent the Compton up-scattered component from the multi-color disk emission. Thus the fit model included three components, a multi-color disk, a power law, and a Comptonized blackbody. The disk parameters obtained from the spectral fits are shown in figure 3. According to the figure, when the disk temperature is high, e.g. higher than $\sim$1.2 keV, the disk luminosity starts to deviate from the relation $L_{disk} \propto T_{in}^4$ showing saturation. This change of the slope suggests that the structure of the disk changes when the disk temperature is high. In fact, saturation of the disk luminosity is the characteristics just expected for the slim disk.[27] We discuss this point later in more detail.

3.3 XTE J1550–564

XTE J1550–564 is a superluminal jet source discovered on September 7, 1998, by RXTE ASM.[23] The outburst of the source, which led its discovery, was followed immediately by the second outburst. These two outbursts were covered by the RXTE PCA observations, which were conducted 190 times during October 1998 and April 1999. These sets of data were used for the present analysis. The analysis presented here was performed assuming the source distance of 4 kpc, although the recent optical observations indicate 2.5 kpc.[21]

Analysis method we adopted was the same as that used for GRO J1655–40. We show the results in figure 4. As seen in the figure, some data points show large excursion from the relation $L_{disk} \propto T_{in}^4$; the data are referred to as an abnormal branch. A hint of abnormal branch is recognized also in GRO J1655–40 (data points with label (1) in figure 3). Deviation of the disk luminosity from the relation $L_{disk} \propto T_{in}^4$ is again noticed when the disk temperature is higher than $\sim$1 keV. It is noteworthy that the abnormal branch appears just at the position where the disk luminosity starts to saturate.

4 The Rapid Burster

The violent time variations observed from GRS 1915+105 have some similarities to those of the Rapid Burster. The Rapid Burster is a unique transient neutron-star binary which shows both type I and II bursts during the outburst. It is known that the burst pattern of the Rapid Burster evolves with the development of the outburst. Among the various patterns, the rapidly repetitive bursts followed by the flat-top burst resembles the time variations

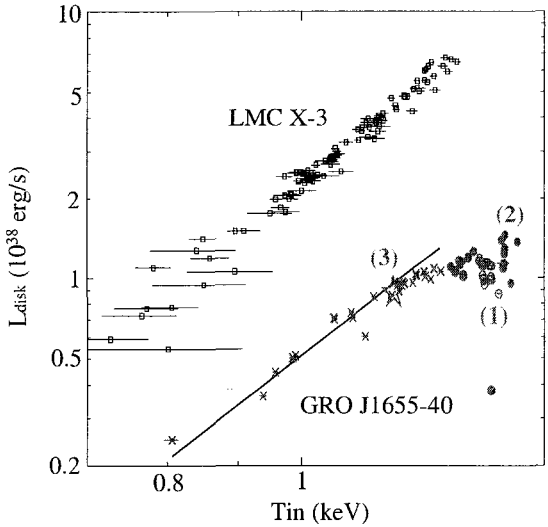


Figure 3: Disk parameters obtained from GRO J1655–40 are plotted together with those for LMC X-3 for comparison. Abscissa represents a color temperature. The labels in the figure indicate the source states: (3) soft state, (2) very high state, and (1) also soft state but the p-free disk model is required. When a Comptonized blackbody is included in the model, its flux is added to that of the disk emission. A thin solid line in the figure indicates the relation between the disk temperature and the disk luminosity for a non-rotating black hole when the α-disk extends down to 3 times the Schwarzschild radius. It is noted that the disk luminosity of GRO J1655–40 saturates when the disk temperature exceeds $\sim$1.2 keV. Adopted from Kubota (2001).[11]

of GRS 1915+105. We compare the two light curves in figure 5. In spite of the similarities between the rapidly repetitive bursts of the Rapid Burster and the rapid, large-amplitude variations of GRS 1915+105, there are some noticeable differences between them. In the case of GRS 1915+105, the rapid, large-amplitude variations are followed by the low-flux phase, which may correspond to the collapse of the inner part of the α-disk. Such collapse of the α-disk, if happened in the Rapid Burster, should produce a large flux at the surface of the neutron star. This may correspond to the flat-top bursts. After the low-flux phase, the inner α-disk of GRS 1915+105 may be refilled during the subsequent high flux phase with little time variations. In the case of the Rapid Burster, this corresponds to the low-flux phase after the flat-top burst. This difference may reflect the difference in the persistent flux level; the Rapid Burster usually show very little persistent emission between the bursts. It is expected from these similarities that the structure of the accretion flow may

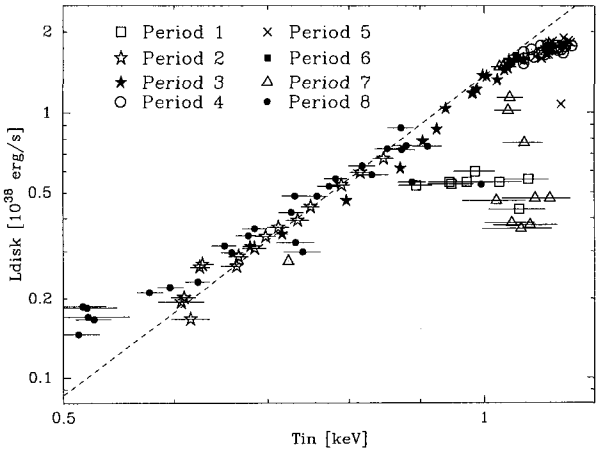


Figure 4: Same as figure 3 but the source is XTE J1550–564. Different symbols correspond to the different epochs of the observations. Adopted from Kubota (2001).[11]

also change largely in the Rapid Burster. To study the structure of the accretion flow, we have analyzed the ASCA GIS data of the Rapid Burster during the outburst in 1999 March.

We divided both the burst and the persistent emission into several phases, and the data are added for each phase to improve statistics. An energy spectrum is calculated for each phase from these composite data. A single component model, such as a blackbody or a power law, did not fit the energy spectra of either the persistent or the burst emission. Thus, we adopted a two component model, a multi-color disk and a blackbody, to fit the energy spectra consistently. The fit was acceptable for most of the spectra, and the best-fit parameters are shown in figure 6. The burst profile is mostly produced by the change of the blackbody emission, while the disk luminosity is more or less constant. The disk parameters do not show large changes with the burst profile and the persistent emission.

The enhanced blackbody emission during the bursts should be produced by the increase of the mass accretion rate from the accretion disk. However, the increased accretion flow should not deplete the inner part of the accretion disk, because the disk parameters do not change very much. This may be possible if the inner part of the α-disk switches to the slim disk during the bursts. This transition causes an enhanced mass accretion rate on to the neutron star,

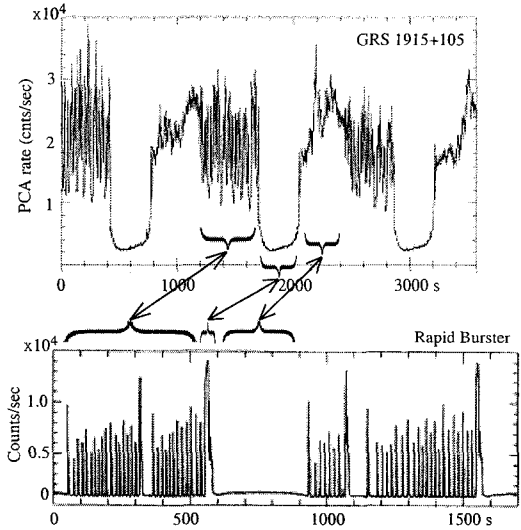


Figure 5: Comparison of the time variations between GRS 1915+105 and the rapid Burster. Correspondence of the light curves, in which the structures of the accretion flow are considered to be similar, is indicated by arrows. The light curves are adopted from Belloni *et al.* (1997)[2] and from Guerriero *et al.* (1999),[6] respectively, for GRS 1915+105 and the Rapid Burster.

but the change of the disk parameters may be rather small because of the significant X-ray emission from the slim disk.

5 Discussion

5.1 Black Hole Binaries

Analysis of GRS 1915+105 data has shown that the accretion disk can take at least two different branches; one has relatively high temperature and the other has low temperature. Belloni *et al.* (2000)[3] analyzed the variabilities of GRS 1915+105 in a model independent way and found that they could be basically explained by the transitions between three states. Among the three states, two correspond to the soft state; one state has a high disk temperature and the other a low disk temperature. Thus, we have confirmed the results by Belloni *et al.*, and have further identified the presence of the slim disk in the soft state with a high disk temperature. This identification is based on the following results: (1) disk luminosity exceeds the Eddington limit for the α-disk when the disk extends down to $3R_s$, (2) temperature profile of the disk

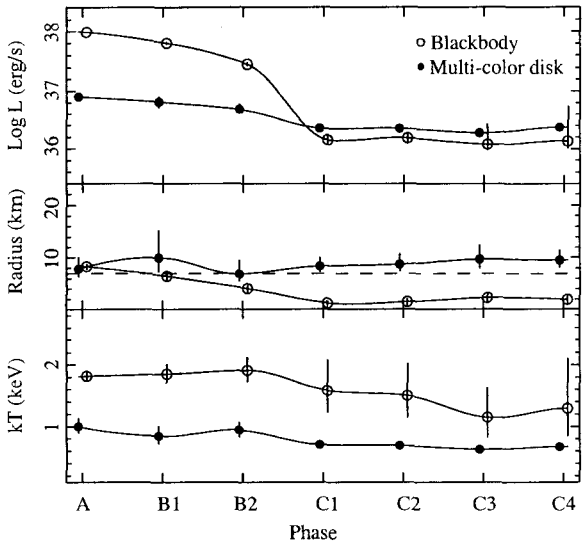


Figure 6: Example of the best-fit spectral parameters obtained from the model fits to the phase-resolved composite energy spectra of the Rapid Burster. The energy spectra were calculated using the ASCA GIS data. Phase A corresponds to the burst peak, phase B to the burst decay, and phase C to the persistent emission. The fit model is a two component model consisting of a multi-color disk model and a blackbody. Adapted from Mahasena (2001).[13]

coincide with that of the slim disk ($p \sim 0.5$), and (3) large jump in the disk parameters, which indicates a complete change in the disk structure.

We also found that ejection of the jets was related to the transition between the slim disk and the α-disk. Although it was already known that the inner disk radius changed largely when the jets were ejected;[2,16] our analysis suggests that the large change of the apparent disk radius is due to the transition between the α-disk and the slim disk. The violent time variations are also produced by the rapid and repeated transitions.

Time variations in other micro-quasars, such as GRO J1655–40 and XTE J155–564, are not very large, and the changes of the disk parameters are more or less continuous. However, clear deviations of the disk parameters from the α-disk are observed from GRO J1655–40 and XTE J1550–564. The deviations are (1) saturation of the disk luminosity when the inner disk temperature becomes high, and (2) temperature profile of the disk coincide with that of the slim disk ($p \sim 0.5$). Saturation of the disk luminosity was also observed in one of the ULXs, IC342 source 1.[27] This is interpreted as an evidence of the

slim disk. In this sense, IC342 source 1 may be the same kind of sources as GRO J1655–40 and XTE J1550–564. It is not clear why the change of the disk parameters associated with the transition between the α-disk and the slim disk is very large in GRS 1915+105, while it is rather small in GRO J1655–40 and XTE J1550–564. Difference in the black hole parameters, such as the spin, may be related.

Although we have studied the disk structures through the spectral analysis, their change may affect the time variations of the X-ray emission. In fact, change of the power spectrum between the upper and lower levels of the violent time variations was observed from GRS 1915+105.[24] If we compare the results with the current analysis, low-frequency noise seems to be prominent in the slim disk. However, the tendency is not very clear in GRO J1655–40 and XTE J1550–564. Larger vertical scale of the slim disk may enhance the interaction between the accretion flow and the outgoing photons, which is expected to cause larger time variations of the outgoing flux.

In this respect, the flip-flop variations observed from GX339–4[18] and GS1124-68[25] may have the same origin as the violent time variations in GRS 1915+105.[5] Although the disk parameters of these sources during the flip-flop variations were not known, clear change was observed in the amplitude of the low-frequency noise associated to the variations. This suggests that the flip-flop variations could result from the transitions between the α-disk and the slim disk. Although we need further analysis to confirm this analogy, it may become a good clue to clarify what causes the transition between the α-disk and the slim disk.

5.2 Neutron Star Binaries

Based on the phenomenological similarities in the light curves of GRS 1915+105 and the Rapid Burster, we studied the change of the disk parameters during the type II bursts. We found that the luminosity of the disk emission showed only a moderate change during the bursts in spite of the large change of the blackbody emission. Because it is generally believed that some kinds of disk instability causes the type II bursts, little change in the disk parameters is somewhat unexpected. Naive interpretation of type II bursts is the results of the collapse of the inner α-disk, which produces the enhanced mass accretion on to the neutron star. According to this interpretation, the inner disk radius is expected to increase, and the disk temperature to decrease, during the type II bursts, which were not observed clearly.

Only the small change of the disk parameters during the type II bursts may be explained if we assume that an inner part of the α-disk makes transition

to the slim disk. Because the slim disk radiates significant X-ray emission, transition of the inner part of the α-disk to the slim disk may not change the total luminosity from the accretion disk largely. Of course, because nature of the X-ray emission is different between the α-disk and the slim disk, disk parameters may change slightly, which was not clearly detected in the present analysis. Development of the slim disk increases the emission from the surface of the neutron star from two reasons: (1) mass accretion rate on to the neutron star increases, because the surface density of the slim disk is lower than the α-disk, (2) slim disk bring some of the gravitational potential energy released in the α-disk on to the neutron star surface.

If the slim disk really develops in the Rapid Burster during the type II bursts, it may also exist in other low-mass X-ray binaries when the luminosity is high. However, it is not so easy to identify the presence of the slim disk by the spectral analysis, because the X-ray emission from the neutron star surface is always comparable to the disk emission. Instead, power spectra may be used to infer the presence of the slim disk. As discussed in the previous subsection, low-frequency noise tends to become strong when the slim disk develops in GRS 1915+105. It is known that the bright low-mass X-ray binaries, so-called z sources, can take three spectral states or branches in the correlated spectral and timing behaviors. Among the branches, horizontal branch has dominant low-frequency noise. In the horizontal branch, the energy spectrum is on average the hardest among the branches. This means that the relative strength of the X-ray emission from the neutron star surface is the largest. If the slim disk is really developed in the horizontal branch, relatively strong emission from the neutron star surface is naturally explained, because some of the gravitational potential energy is advected on to the neutron star surface. Hirano *et al.* (1995)[8] has studied the radial profile of the disk temperature applying the p-free disk model to the Ginga data of Cyg X-2, which is one of the bright low-mass X-ray binaries. They found that p was around 0.67 in the horizontal branch and increased to 0.85 in the normal branch. If we consider the systematic effect in estimating p with the lack of the soft band coverage (say < 2 keV),[11] an absolute value of p may not be very useful. The smallest value of p observed in the horizontal branch supports the idea that the slim disk is developed in the horizontal branch. Although more detailed observations may be necessary to confirm the presence of the slim disk, transitions between different disk structures may play an important role to produce the correlated spectral and timing variations in bright low-mass X-ray binaries.

Acknowledgments

I thank Drs. Yamaoka, Kubota, and Mahasena for useful discussions and for providing the figures from their theses.

References

1. M.A. Abramowicz *et al.*, *ApJ* **332**, 646 (1988).
2. T. Belloni *et al.*, *ApJ* **479**, L145 (1997).
3. T. Belloni *et al.*, *A&A* **355**, 271 (2000).
4. A.J. Castro-Tirado *et al.*, *IAU Circ.*, 5590 (1992).
5. T. Dotani in *Observational Evidence for Black Holes in the Universe*, ed. S.K. Chakrabarti (Kluwer, the Netherlands, 1999), p.341.
6. R. Guerriero *et al.*, *MNRAS* **307**, 179 (1999).
7. B.A. Harmon *et al.*, *Nature* **374**, 703 (1995).
8. A. Hirano *et al.*, *ApJ* **446**, 350 (1995).
9. R.M. Hjellming, M.P. Rupen, *Nature* **375**, 464 (1995).
10. S. Ichimaru, *ApJ* **214**, 840 (1977).
11. A. Kubota, PhD thesis, University of Tokyo, 2001.
12. A. Kubota *et al.*, *ApJ* **560**, L147 (2001).
13. P. Mahasena, PhD thesis, University of Tokyo, 2001.
14. K. Makishima *et al.*, *ApJ* **535**, 632 (2000).
15. I.F. Mirabel, L.F. Rodriguez, *Nature* **371**, 46 (1994).
16. I.F. Mirabel *et al.*, *A&A* **330**, L9 (1998).
17. K. Mitsuda *et al.*, *PASJ* **36**, 741 (1984).
18. S. Miyamoto *et al.*, *ApJ* **383**, 784 (1991).
19. R. Narayan, I. Yi, *ApJ* **452**, 710 (1995).
20. J.A. Orosz, C.D. Bailyn, *ApJ* **477**, 876 (1997).
21. C. Sánchez-Fernández *et al.*, *A&A* **348**, L9 (1999).
22. N.I. Shakura, R.A. Sunyaev, *A&A* **24**, 337 (1973).
23. D.A. Smith, *IAU Circ.*, 7008 (1998).
24. J. Swank *et al.* in *Accretion Processes in Astrophysical Systems: Some Like it Hot!*, ed. S.S. Holt, T.R. Kallman (AIP Conference Proceedings 431, 1998), p.327 (astro-ph/9801220).
25. M. Takizawa, *et al.*, *ApJ* **489**, 272 (1997).
26. Y. Tanaka, in *Ginga Memorial Symposium*, ed. F. Makino, & F. Nagase (Institute of Space and Astronautical Science, Kanagawa, 1992), 19.
27. K. Watarai *et al.*, *ApJ* **549**, L77 (2001).
28. K. Yamaoka, PhD thesis, University of Tokyo, 2001.
29. S.N. Zhang *et al.*, *IAU Circ.*, 6046 (1994).

SEARCHING FOR EVIDENCE OF TIDAL DISRUPTION EVENT IN LONG-TERM X-RAY LIGHT CURVE OF SEYFERT GALAXY MCG-2-58-22

CHUL-SUNG CHOI
Korea Astronomy Observatory
36-1 Hwaam-dong, Yusong-gu, Taejon 305-348, Korea
cschoi@kao.re.kr

TADAYASU DOTANI
Institute of Space and Astronautical Science
3-1-1 Yoshinodai, Sagamihara, Kanagawa 229-8510, Japan
dotani@astro.isas.ac.jp

HEON-YOUNG CHANG, AND INSU YI
Korea Institute for Advanced Study
207-43 Cheongryangri-doing Dongdaemun-gu, Seoul 130-012, Korea
hyc@ns.kias.re.kr, iyi@ns.kias.re.kr

Stars can be captured or tidally disrupted by the supermassive black hole in galactic nuclei. It has long been thought that the tidal disruption of stars by a supermassive black hole can result in flare-like activity in the central regions of active galactic nuclei (AGNs) and even in normal galaxies, which occurs on a timescale of years. We report that MCG-2-58-22 (Mrk926) shows such a variability in the X-ray band. Long term variability of X-ray emission of active galactic nuclei and normal galactic centeres is of an interest in study dramatic phenomena expected from exotic processes such as a tidal disruption of a star passing-by a supermassive black hole. In this paper we present a long-term X-ray light curve (2−10 keV) of the bright Seyfert 1 galaxy MCG-2-58-22 by compiling data from various X-ray satellites, which covers more than 20 years. We find that two distinct time variations are displayed in the light curve. One is a gradual and secular decrease of the X-ray flux has a timescale of several ten years, and the other is abrupt increases of X-ray flux by a factor of $2-4$ compared with the flux level of gradual change parts. We detected 3 flares in total; a possible duration is ~ 2 years and an interval $\sim 6-8$ years. We discuss possible origins of these variabilities. Though a standard disk instability theory may explain a displayed time variability in X-ray light curve, the subsequent accretions of stellar debris of the tidal disruption event due to the supermassive black hole in MCG-2-58-22 cannot be ruled out as an alternative explanation.

1 Introduction

X-ray observations of active galactic nuclei (AGNs) show that most of them are variable with a various range of amplitudes and on many timescales (Lawrence et al. 1985; Grandi et al. 1992; Mushotzky, Done, & Pounds 1993; Nandra et

al. 1997; Ulrich, Marachi, & Urry 1997; Ptak et al. 1998; Turner et al. 1999). The short-time variability amplitude is known to be anti-correlated with the source luminosity for Seyfert 1 galaxies (Barr & Mushotzky 1986; Nandra et al. 1997; Ptak et al. 1998). The variability is thought to be linked to the central engine so that its behavior provides information on the central engine. Variability of the X-ray emission is a powerful probe of physical processes occurring in the inner regions of AGNs. In particular, the rapid variability is widely believed to be related to the central part, and actually has been used to constrain physical properties of the central engine. As substantial variability cannot be observed from a source on timescales shorter than the light crossing time of the source, so that this gives an upper limit to the size of $R < c\delta t$. Variability can also be used to derive the efficiency, η, of the source in converting accreting matter to luminosity. A change in luminosity δL in time δt implies an efficiency of $\eta = 0.05\delta L_{43}/\delta t_{100}$, where $\delta L_{43} = \delta L/10^{43}$ ergs^{-1}, $\delta t_{100} = \delta t/100$sec. Several AGN require efficiencies larger than those produced in nuclear processes, providing the strongest evidence to date that AGN are indeed powered by accretion.

Whereas small X-ray variability is a common property of AGNs and may reveal the intrinsic property of the central engine, X-ray outbursts are relatively rare and may be irrelavant to the inherent state of the central engine. It is yet worthwhile to study a long term variability since it may reveal other interesting features, such as, a gradual change of a global structure of the accretion disk around the central supermassive blackhole, or a stochastic flare due to some exotic events, that is, stellar tidal disruptions. Markowitz & Edelson (2001) analyzed 300 day light curves of Seyfert 1 galaxies in 2−10 keV band. They showed that the X-ray variability of Seyfert 1 galaxies is described by a single, universal power-density spectrum, and that the cutoff moves to longer time scales for higher luminosity sources. Soft X-ray outbursts which have about 2 orders of magnitude variation were observed from NGC 5905 (Bade, Komossa, & Dahlem 1996) and Zwicky 159.034 (Brandt, Pounds, & Fink 1995), of which various possible scenarios have been reviewed by Komossa & Bade (1999).

Until now, on the other hand, long-term modulations in a timescale longer than several years in light curves of AGNs have mostly been studied in the optical range. For instance, Webb (1990) reported the results of 61 years of optical observations of 3C120 and claimed a gradual and possibly sinusoidal variability component of period 12.43 years as well as high amplitude flares on much shorter timescales. Peterson et al. (1998) reported the spectroscopic monitoring of nine Seyfert 1 galaxies in optical band to determine a size of the broad line emission regions. The reason why studies of long term variability of AGNs have been done in the optical range is because X-ray variability longer

than a few years requires gathering sparse, more seriously, heterogeneous sets of observational data. Complied observational data of these kinds may provide effectively results of long-term flux monitoring as long as the observed fluxes are properly converted to those in the appropriate energy range for comparisons.

In the present work, we study a long-term X-ray light curve of MCG-2-58-22 covering more than two decades and discuss possible origins of the long-term behaviors. In the previous work, Choi et al. (2001) analyzed the ROSAT and ASCA data of this source and noticed flare-like events. This motivates us to perform a thorough analysis of the long-term X-ray flux variations of this source. For this purpose, we have gathered X-ray flux measurements of MCG-2-58-22 from the literature as well as raw X-ray data from the HEASARC public archives at NASA/GSFC and from the SIRIUS data base at ISAS. The observational data we gathered include those of HEAO-1 (Rothschild et al. 1979), Einstein (Giacconi et al. 1979), EXOSAT (Taylor et al. 1981), Ginga (Turner et al. 1989), ROSAT (Pfeffermann 1986), ASCA (Tanaka, Inoue, & Holt 1994), and RXTE (http://heasarc.gsfc.nasa.gov/docs/xte). As a result of this, quasi-periodic variabilities have been found from a long-term X-ray light curve which covers more than 20 years. The light curve indicates at least two timescales, one of which seems likely to be due to tidal disruption events.

2 Data Analysis

2.1 MCG-2-58-22

MCG-2-58-22 is a bright Seyfert 1 galaxy at $z = 0.04732$, which corresponds to a distance of 284 Mpc for $H_0 = 50$ km s^{-1} Mpc^{-1} (Huchra et al. 1993). The X-ray luminosity of MCG-2-58-22, $L_X \sim 10^{44}$ erg s^{-1}, is known to be variable in a timescale of a few 10^3 seconds to years (Grandi et al. 1992; Nandra & Pounds 1994; Choi et al. 2001). Many observational characteristics are typical of Seyfert 1 galaxies, such as a time variability, a power-law type continuum spectrum, a soft excess phenomenon, and its variable nature (Ghosh & Soundararajaperumal 1992; Nandra & Pounds 1994; Weaver et al. 1995; Nandra et al. 1997; George et al. 1998; Turner et al. 1999). Mass of the central supermassive black hole (SMBH) of MCG-2-58-22 is estimated from the UV and optical observations as a few times of $10^8 M_\odot$ (Padovani & Rafanelli 1988; Wandel 1991). MCG-2-58-22 has also been studied in other energy bands. Mundell et al. (2000) observed MCG-2-58-22 using the VLBA at 8.4 GHz and detected the nuclear source without any extended structures from the parsec-scale image, suggesting that this source could be a "bare" Seyfert 1 nucleus (see also Weaver et al. 1995). In the optical region, this source has displayed a year-scale continuum variation (de Ruiter & Lub 1986; Whittle 1992), as well

as very broad Balmer line variability in their profiles and luminosities (e.g., Winkler 1992 and references therein).

2.2 Flux Calibration

Since the different instruments of various observatories cover different energy range, in order to compare we convert the fluxes to those in the 2−10 keV band, which is covered by most of the instruments. Having surveyed the reported flux data, we further collect and analyze the archive data of MCG-2-58-22 from the HEASARC public archives (e.g., EXOSAT data of 1983 November, ROSAT data of 1990 November through 1993 December, ASCA data of 1993 May through 1997 December, and RXTE data of 1997 December through 1999 August) and from the SIRIUS database at ISAS (e.g., Ginga data of 1991 June). In the analysis of these data, we select good event data by applying standard data-screening criteria to the raw data, e.g., background subtraction, rejection of possibly contaminated data from the bright Earth and regions of high particle background, etc. Then, we finally obtain the X-ray fluxes in 2−10 keV band through a spectral fit to the screened data, assuming the power-law spectrum with a photon index of $\Gamma = 1.75$ and the hydrogen equivalent absorption column density of $N_H = 3.5 \times 10^{20}$ cm^{-2} in the line of sight to the source (e.g., Weaver et al. 1995; Piro, Matt, & Ricci 1997).

2.3 Error Estimates

The X-ray fluxes converted to the 2−10 keV band may include various systematic errors. For example, there could exist contaminating sources in the vicinity of MCG-2-58-22. We have checked various X-ray catalogs and confirmed that X-ray sources within 3 degrees from MCG-2-58-22 are dimmer than MCG-2-58-22 by an order of magnitude or more. Thus the effect of contaminating sources can be neglected in the present analysis. Another possibility is a spectral variability. The power-law slope of MCG-2-58-22 may not be constant, and the energy spectra may include some structures, such as the broad iron line and/or the Compton reflection hump. Effects of the spectral variability on long-term light curves have been evaluated as follows. We note that the reported photon indices in literatures fall within $\Gamma = 1.4 - 1.9$, except for the ROSAT observations. We have checked how the flux estimation is affected both by changes of photon indices and by structures which may exist in energy spectra (George et al. 1998; Nandra et al. 1997). The systematic error associated with these uncertainties was found to be much smaller than 10 %, when observations cover the 2−10 keV band. On the other hand, since the ROSAT PSPC did not cover the 2−10 keV range calculated fluxes include

relatively larger systematic errors. In the case of the ROSAT data, we may not assume a simple power-law, because the photon indices are different between the 0.1−2 keV ROSAT band ($\Gamma \sim 2.1$) and the typical 2−10 keV band ($\Gamma \sim 1.4 - 1.9$). Piro, Matt, & Ricci (1997) attempted various models to explain both Ginga and ROSAT data simultaneously. We utilize their models to estimate the systematic error in the flux conversion from the PSPC count rate to the 2−10 keV flux. We find that the flux conversion factors change by ~ 20 % for the different models if we take into account the possible changes of the power-law slope in 2−10 keV. Thus the converted ROSAT flux may include the systematic error as large as 20 %. In addition to these systematic errors, we also need to consider the cross-calibration error among the instruments. Although the magnitude of cross-calibration errors is usually unknown, we assume it is no more than 10 %. Therefore, if the 2−10 keV band is covered by the instruments the systematic error in the flux conversion is dominated by the cross-calibration error among the instruments and may be evaluated as $\sim$ 10 %. On the other hand, if the 2−10 keV band is not covered by the instrument as the case of the ROSAT PSPC the systematic error in the flux conversion is dominated by the uncertainty of the spectral shape, which is approximately 20 %.

3 Combined Light Curve

The combined X-ray fluxes are shown in Figure 1 with error bars. One may see immediately that the light curve shows at least two distinct features, i.e., one is a gradual and secular decrease of the flux and the other abrupt increases, or flares denoted by vertical arrows. The flux decreases during the period of longer than 15 years at least, from 1977 September to 1993 December. It is interesting to note the 3 highest data points in the light curve in 1983 November, 1991 June through November, and 1997 June through December. The flux increased at least by a factor of $\sim 2 - 4$ during the flares compared to a long-term trend of the flux decrease. It is not a matter of the cross-calibration, since flare events are observed even in a given satellite (flare events in 1983, 1991, 1997 for EXOSAT, ROSAT, ASCA and RXTE, respectively). Furthermore, the flares in 1991 and 1997 were observed by the imaging instruments, ROSAT and ASCA. This suggests strongly that MCG-2-58-22 is indeed the source of these flares. We have also checked whether or not these events are related to an artificial effect, such as background fluctuation and solar X-ray contamination, and concluded that these are intrinsic phenomena to the source. Durations and intervals between flares may give important information on the characteristics of these events. It is difficult to measure the *exact* peak flux and thus the

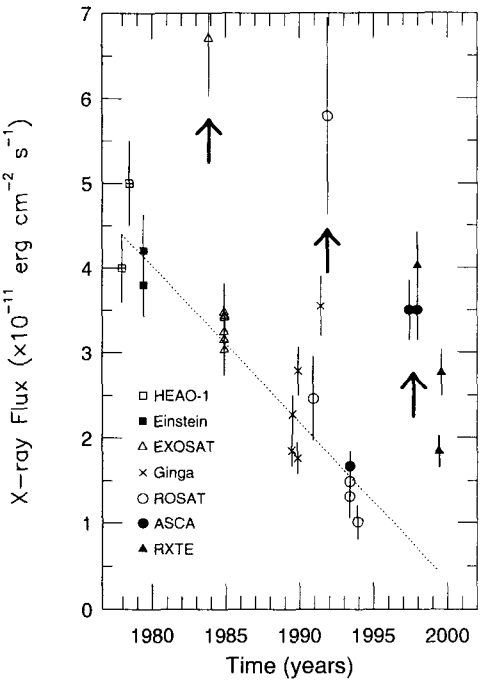


Figure 1: Long-term X-ray light curve of MCG-2-58-22 in the energy range of 2 – 10 keV. The flux data of HEAO-1 and Einstein are adopted from Piccinoti et al. (1982), and Turner et al. (1991), otherwise calculated in this study using a spectral fit to the data. Different symbols are used to distinguish the data of different satellites. The systematic uncertainty to the flux conversion particularly to the ROSAT data is estimated to be small, at most $\sim 20\%$.

duration of flares because of the sparse sampling of data. Based on the 1991 flare, we estimate its possible duration is $\sim$ 2 years. The flares have occurred almost quasi-periodically from 1983 to 1997. Namely, the time interval of the flares between the first and the second flares is $\sim$ 8 years and is $\sim$ 6 years between the second and the third flares. We have also investigated possible spectral changes associated to these flares using the ROSAT and ASCA data. Even though clear spectral changes are found associated to the flares, more formal analysis is further required.

4 Possible Explanations

As shown in Figure 1, MCG-2-58-22 clearly shows two characteristic variabilities, i.e., the gradual and secular decrease of the X-ray flux and flares occurring almost quasi-periodically. The detected two characteristic variabilities imply that the source is undergoing at least two distinctly different physical processes. One of the possible origins of the long-term variabilities is an instability which could arise from an accretion disk, and results in modulation of mass-accretion rate and consequently observed flux. If the disk experiences a viscous-thermal instability caused by the sudden change of hydrogen ionization in the disk, the instability timescale is roughly estimated as a few $\sim 10^2$ years. The viscous timescale near the innermost radius of the standard α-disk can be a few ten years for a black-hole mass of 10^8 $M_\odot$. On the other hand, the flares can not be explained by any disk instabilities, considering the dynamical timescale and the thermal timescale at the inner disk region are $\sim 10^{-2}$, $\sim 10^{-1}$ years, respectively. Observed timescales seem shorter than theoretically expected timescales. Therefore, disk instabilities cannot well explain observed flare events. Magnetic reconnection may be considered as an alternative machanism for the flares, but it seems to be unlikely to explain the duration time and the frequency of the flares.

In addition to the disk instabilities, the variabilities can be explained by a similar mechanism occurring in the galactic black hole candidates (GBHCs). Phenomenologically, the presence of the characteristic timescale is known in both GBHCs in the hard state and the SMBHs (Hayashida et al. 1998; Markowitz & Edelson 2001). In both types of sources, the characteristic timescale corresponds to the cutoff in the power density spectrum (PDS), and the timescale may be scaled almost linearly with the mass of the black hole. The canonical galactic black hole, Cyg X-1, has roughly 10 $M_\odot$ mass, which shows the cutoff of the PDS around $\sim$ 10 secs. The PDS of Cyg X-1 may be reproduced by the superposition of the shots, whose typical duration determines the characteristic timescale of the source (e.g. Negoro, Kitamoto & Mineshige 2001). If we assume that the SMBH in MCG-2-58-22 has a mass of 10^8 $M_\odot$ (e.g., Padovani & Rafanelli 1998 and Wandel 1991), the characteristic timescale becomes about a few years. This is just about the duration of the flares we detected. The flares in MCG-2-58-22 may be reminiscent of the shot noise of the galactic black holes in the hard state. Nonetheless, the slope expected from the shot noise model is inconsistent with that obtained by the Fourier analysis we have performed. It is, however, fair to say that physical processes involved in the variability actaully cause the difference since the data length and sampling intervals may result in undesired artifacts.

Temporal increases of the mass accretion rate could arise in principle from other mechanisms. An entertaining possibility of such mechanism is the tidal disruption of stars by the SMBH (Rees 1988). The frequency that a star would pass within a distance r from the SMBH can be approximated by $\approx 10^{-3}\ M_8^{4/3}(N_*/10^6\ \mathrm{pc}^{-3})(\sigma/300\ \mathrm{km\ s}^{-1})(r/r_t)\ \ \mathrm{yr}^{-1}$, where M_8 is the mass of the SMBH in units of $10^8\ \mathrm{M}_\odot$, σ is the virial velocity of the stars, r_t is the tidal radius of the SMBH. Although the number density of the stars and their velocity dispersion near the SMBH are not known, flare events we observe in MCG-2-58-22 cannot be regarded as independent tidal disruption events since its event rate is too high. Instead, the subsequent flares may be produced when the bound material from the tidally disrupted star returns to pericenter at a rate of a few years (Rees 1988; Ulmer 1999). The observed peak luminosity $\sim 6 \times 10^{44}$ erg s^{-1} corresponds to the luminosity that the debris of 0.1 $\mathrm{M}_\odot$ is being swallowed steadily with 10 % efficiency in a year. Observed flares is not inconsistent those from tidally disrupted stellar debris due to the SMBH, even though observation would depend on many unknown parameters, such as, the type of the disrupted star, the spin of the SMBH, the minimum radius of the trajectory. Estimated masses of the SMBH in MCG-2-58-22 seems a bit larger than a favored upper limit. However, in case of a Kerr black hole, the tidal disruption may occur for larger BH masses provided that the star approaches from a favorable direction (Beloborodov et al. 1992). Moreover, the atmosphere of giant stars could have been stripped instead of complete disruption, while the favored upper limit is a derived by theoretical works of the disruption processes of a solar type star (e.g., Nolthenius & Katz 1982; Carter & Luminet 1983; Rees 1988; Laguna et al. 1993; Ulmer 1999).

5 Conclusion

We report that long-term X-ray variability behaviors of MCG-2-58-22 by analyzing archival data of various X-ray satellites. We notice two characteristic variabilities in the light curve: one is the gradual and secular decrease of the X-ray flux which has a timescale of several ten years and the other is the flare. Although it is difficult to measure the duration and the interval of the flare accurately due to the sparse sampling of the observational data, a possible duration is $\sim$ 2 years and an interval $\sim 6 - 8$ years. Two distinct timescales can be accounted for with a supermassive black hole system with an unstable accretion disk, that is, an instability of the disk and a tidal disruption of stars by a supermassive black hole. Further spectral analysis and observational data in other wavelength should be explored to confirm this scenario. For instance, we note that long-term optical light curve of MCG-2-58-22 was obtained by

Winkler et al. (1992) demonstrate a clear variation of about 0.3 mag in the V-band during the period which overlapds Ginga observations in 1989 and 1990. It is difficult to ascertain that whether the optical variation is physically related with the X-ray variation, since the optical observation period is too short to compare it with that of the X-ray light curve. If it turns out that optical and X-ray light curves vary almost simultaneously, flare events are more likely due to an external event, a tidal disruption of star.

References

1. Bade, N., Komossa, S., Dahlem, M. 1996, A&Ap, 309, L35
2. Barr, P., & Mushotzky, R. F. 1986, Nature, 320, 421
3. Beloborodov, A. M., Illarionov, A. F., Ivanov, P. B., Polnarev, A. G. 1992, MNRAS, 259, 209
4. Brandt, W. N., Pounds, K. A., & Fink, H. 1995, MNRAS, 273, L47
5. Carter, B., & Luminet, J. P. 1983, A&Ap, 121, 97
6. Choi, C. S., Dotani, T., Yi, I., & Kim, C. 2001, in Proc. of X-ray Astronomy 2000, eds. R. Giacconi, L. Stella, and S. Serio, in press
7. de Ruiter, H. R., & Lub, J. 1986, AAPS, 63, 59
8. George, I. M., Turner, T. J., Netzer, H., Nandra, K., Mushotzky, R. F., & Yaqoob, T. 1998, ApJS, 114, 73
9. Ghosh, K. K., & Soundararajaperumal, S. 1992, ApJ, 398, 501
10. Giacconi, R., et al. 1979, ApJ, 230, 540
11. Grandi, P., Tagliaferri, G., Giommi, P., Barr, P., & Palumbo, G. G. C. 1992, ApJS, 82, 93
12. Hayashida, K., Miyamoto, S., Kitamoto, S., Negoro, H., & Inoue, H. 1998, ApJ, 500, 642
13. Huchra, J., Latham, D. W., Costa, L. N., Pellegrini, P. S., & Willmer, C. N. A. 1993, AJ, 105, 1637
14. Komossa, S., & Bade, N. 1999, A&Ap, 343, 775
15. Laguna, P., Miller, W. A., Zurek, W. H., Davies, M. B. 1993, ApJL, 410, L83
16. Lawrence, A., Watson, M. G., Pounds, K. A., & Elvis, M. 1985, MNRAS, 217, 685
17. Markowitz, A. & Edelson, R. 2001, ApJ, 547, 684
18. Mushotzky, R. F., Done, C., & Pounds, K. A. 1993, ARAA, 31, 717
19. Mundell, C. G., Wilson, A. S., Ulvestad, J. S., & Roy, A. L. 2000, ApJ, 529, 816
20. Nandra, K., & Pounds, K. A. 1994, MNRAS, 268, 405
21. Nandra, K., George, I. M., Mushotzky, R. F., Turner, T. J., & Yaqoob,

T. 1997, ApJ, 476, 70
22. Negoro, H., Kitamoto, S., & Mineshige, S. 2001, ApJ, 554, 528
23. Nolthenius, R. A., & Katz, J. I. 1982, ApJ, 263, 377
24. Padovani, P., & Rafanelli, P. 1988, A&Ap, 205, 53
25. Peterson, B. M. et al. 1998, ApJ, 501, 82
26. Pfeffermann, E. et al. 1986, Proc. SPIE, 733, 519
27. Piccinotti, G., Mushotzky, R. F., Boldt, E. A., Holt, S. S., Marshall, F. E., Serlemitsos, P. J., & Shafer, R. A. 1982, ApJ, 253, 485
28. Piro, L., Matt, G., & Ricci, R. 1997, A&AS, 126, 525
29. Ptak, A., Yaqoob, T., Mushotzky, R. F., Serlemitsos, P., & Griffiths, R. 1998, ApJ, 501, L37
30. Rees, M. 1988, Nature, 333, 523
31. Rothschild, R. et al. 1979, Space Sci. Instr., 4, 265
32. Tanaka, Y., Inoue, H., & Holt, S. S. 1994, PASJ, 46, L37
33. Taylor, B. G., Andresen, R. D., Peacock, A., & Zobl, R. 1981, Space Sci. Rev, 30, 479
34. Turner, M. J. L. et al. 1989, PASJ, 41, 345
35. Turner, T. J., George, I. M., Nandra, K., & Turcan, D. 1999, ApJ, 524, 667
36. Turner, T. J., Weaver, K. A., Mushotzky, R. F., Holt, S. S., & Madejsk, G. M. 1991, ApJ, 381, 85
37. Ulmer, A. 1999, ApJ, 514, 180
38. Ulrich, M. -H., Marachi, L., & Urry, C. M. 1997, ARAA, 35, 445
39. Wandel, A. 1991, A&Ap, 241, 5
40. Weaver, K. A., Nousek, J., Yaqoob, T., Hayashida, K. A., & Murakami, S. 1995, ApJ, 451, 147
41. Webb, J. R. 1990, AJ, 99, 49
42. Whittle, M. 1992, ApJS, 79, 49
43. Winkler, H., Glass, I. S., van Wyk, F., Marang, F., Spencer Jones, J. H., Buckley, D. A. H., & Sekiguchi, K. 1992, MNRAS, 257, 659

EXISTENCE OF X-RAY JETS ON KILOPARSEC SCALES IN RADIO-LOUD AGNS

J.M. BAI
Astronomy Program, SEES, Seoul National University, Seoul 151-742, Korea
Yunnan Astronomical Observatory, NAO, Chinese Academy of Sciences, Kunming 650011, China
E-mail: jmmbai@astro.snu.ac.kr

Myung Gyoon LEE
Astronomy Program, SEES, Seoul National University, Seoul 151-742, Korea

Based on the knowledge of the nature and physics of blazars, we argue that a detectable synchrotron X-ray jet exists on kiloparsec scales in most radio-loud AGNs. In blue blazars and blue-blazar-like radio galaxies, the parsec scale synchrotron X-ray jet continuously extends to kiloparsec scales, which is consistent with the *Chandra* recent observations for the blue-blazar-like radio galaxy, Cen A that shows the X-ray jet extending continuously from 10 parsec scales to kpc scales. Although the parsec scale X-ray jet in red blazars and red-blazar-like radio galaxies are inverse-Compton emission, a detectable synchrotron X-ray jet exists on kiloparsec scales in these sources. In typical red blazars and red-blazar-like radio galaxies, the X-ray jet is bright on 10 kpc scales whether the jet is highly relativistic on large scales or not. In extreme and thus the most powerful red blazars in which the synchrotron peak of the inner jet lies in the infrared bands and the jet may be still highly relativistic on large scales, the synchrotron peak may shift back to the optical bands on 100 kpc scales. Consequently, the X-ray jet in these sources gets fainter along the jet from 10 kpc to 100 kpc scales, while the optical and IR jet may get brighter, and on larger scales ($>\sim$ 100kpc, as in the case of PKS 0637−752) the X-rays from the jet may be dominated by inverse-Compton emission. These predictions can be tested with the ongoing observations of the Chandra X-ray Observatory.

1 Introduction

Extragalactic jets are a spectacular feature associated with the activity of radio-loud AGNs. They are pipelines through which the mass, momentum and energy are transported from the central nucleus to the extended radio lobes (e.g. Rees 1971; Begelman et al. 1984; Laing 1993), and naturally play an important role for understanding the nature and physics of the invisible central engine in AGNs which is widely believed to be an accretion disk surrounding a super massive black hole (e.g. Rees 1984). Jets are usually observed in the radio bands with VLA on kiloparsec (kpc) scales and with VLBI on parsec scales. A handful of radio jets have been detected in the IR/optical bands by the Hubble Space Telescope on kpc scales. At X-ray energies, about four nearby

radio jets have been detected by EINSTEIN and ROSAT X-ray observatories in Cen A (Turner 1997; Feigelson et al. 1981), M87 (Neumann et al. 1997; Schreier et al. 1982), 3C 273 (Röser et al. 2000; Harris and Stern 1987) and NGC 6251 (Mack et al. 1997), respectively.

Since it was launched in July 1999, Chandra X-ray Observatory (CXO) has detected several more radio jets on kpc scales with high spatial and spectral resolution at X-ray energies. While the radio (and probably the optical) emission of the jets is certain to be synchrotron, the radiation mechanisms responsible for the kpc scale X-ray jets have not been well understood yet. The X-ray jet in the distant quasar PKS 0637-752 has been explained in synchrotron self-Compton (SSC) models (Schwartz et al. 2000) and as inverse-Compton scattering of the cosmic microwave background (CMB, Tavecchio et al. 2000; Celotti et al. 2001). The X-ray jet in the famous quasar 3C 273 can be interpreted as synchrotron emission, or synchrotron self-Compton emission or inverse-Compton scattering of the CMB (Marshall et al. 2001; Sambruna et al. 2001; Röser et al. 2000). The X-ray jet in M87, Cen A and Pictor A has also been interpreted in several models (Neumann et al. 1997; Perlman et al. 2001; Turner et al. 1997; Wilson et al. 2001). All these explanations are based on spectral fitting in the X-ray band or broad band spectra from radio to X-rays, uncertainty of which is usually high due to the weak signal of the jet.

Blazars, including BL Lac objects and flat-spectrum radio quasars (FSRQs), are compact, flat spectrum radio sources with highly variable and polarized nonthermal continuum emission extending up to X-ray and often gamma-ray energies, which are generally understood as consequences of a relativistic jet oriented close to the line of sight (e.g. Blandford & Rees 1978; Blandford & Königl 1979; Marscher 1980; Ghisellini et al. 1993; Kollgaard 1994). The compact jet of a blazar which appears as a bright core in VLA maps but usually exhibits an elongated jet in VLBI maps, connects the central region with the outer, more extended, large scale jets and radio lobes, and dominates the radio emission of the source, as in the cases of 3C 273 and PKS 0637−752. Blazars have been observed intensively from radio to gamma-rays, and the radiation mechanisms of the inner jets (the compact, parsec and subparsec scale jets) have been almost certain. The broad band spectral energy distributions (SEDs) of blazars have two components, exhibiting a self-similar double-hump structure. Correlated variations across the SEDs are consistent with the picture that a single electron population in the relativistic jet gives rise to both components, via synchrotron at low energies and inverse-Compton scattering at high energies (e.g. Ulrich et al. 1997; Ghisellini et al. 1998). In red blazars (such as 3C 273 and PKS 0637-752) which have synchrotron peak at IR/optical

wavelengths, the X-rays are inverse-Compton emission, while in blue blazars which have synchrotron peak at UV/X-rays, the X-rays are an extension of synchrotron emission (Padovani & Giommi 1996; Kubo et al. 1998). Besides the emission mechanisms, other aspects of the inner jets in blazars have been intensively studied as well.

The knowledge about the inner jet in blazars may serve as a starting point and a useful tool to study the emission mechanisms in the kpc scale jet. The kpc scale jet is an extension of the inner jet, connecting the inner jet to the radio lobe, and thus must be somewhat similar to the inner jet. Conclusions about the inner jet may be applied or generalized to the kpc scale jet, and the conservative quantities can be used to constrain the models for the kpc scale jet. In this paper, we argue that a detectable synchrotron X-ray jet exists on kpc scales in radio-loud AGNs, based on the knowledge of the inner jet in blazars.

2 X-ray jets in blazars

In blue blazars, the inner jet produces X-rays via synchrotron process. It is possible that the parsec scale synchrotron X-ray jet extends to the kpc scale, i.e., there may be a synchrotron X-ray jet on kpc scales. In red blazars, the inner X-ray jet is inverse-Compton emission, so it seems that red blazars cannot produce an X-ray jet on kpc scales through synchrotron process. Furthermore, for some sources, the trough between the synchrotron and Compton peak happens to lie in the range of X-rays, and consequently there are not any X-ray jets in these sources on kpc scales. However, it is probable that there exists a synchrotron X-ray jet in red blazars on kpc scales. Statistical studies for complete samples of blazars revealed spectral trends that as the bolometric luminosity increases, the luminosities of the emission-lines and Compton component increase but the frequency of the peak of the synchrotron component decreases (Sambruna et al. 1996; Fossati et al. 1998). These trends support a simple paradigm in which the electron acceleration is similar in both red and blue blazars but the relativistic electrons in the inner jet of red blazars suffer more Compton cooling because of larger external photon densities as indicated by the higher emission-line luminosity, leading naturally to lower characteristic electron energies and thus lower frequencies of the synchrotron peak (Ghisellini et al. 1998; Urry 1999), i.e. it is strong Compton cooling that prevents synchrotron peak from moving to higher frequencies. On kpc scales, the external photon densities of the jet in red blazars are much smaller, so the electrons may reach much higher characteristic energies. These high energy electrons may produce synchrotron X-ray jet on kpc scales, which can be seen below in

detail.

2.1 Physics of the inner jet in blazars

Statistical analysis for all γ-ray loud blazars shows a strong correlation between the characteristic electron energy γ_{peak} (i.e. the Lorentz factor of the electrons emitting at the peaks of the synchrotron and inverse-Compton components) and energy density $(U_r + U_B)$ in the jet comoving frame for all γ-ray loud FSRQs (Ghisellini et al. 1998),

$$\gamma_{peak} \propto (U_r + U_B)^{-0.5\pm0.06} = C(U_r + U_B)^{-0.5\pm0.06}, \quad (1)$$

where U_r and $U_B = B^2/(8\pi)$ are the energy density of radiation (produced in or outside the jet) and magnetic field, respectively. If including γ-ray loud BL Lac objects, the correlation is $\gamma_{peak} \propto (U_r + U_B)^{-0.6\pm0.04}$, slightly different from equation (1). The difference may be caused by the larger uncertainty of U_r, since BL Lac objects have weak or no emission lines, leading to larger uncertainties to the radiation energy density of the external radiation, U_{ext}. Taking into account both synchrotron and inverse-Compton emission, the electron cooling rate in the jet is

$$m_e c^2 \dot{\gamma} = (4/3)\sigma_T c \gamma^2 (U_r + U_B), \quad (2)$$

where m_e is the mass of electron and σ_T is the Thomson cross-section. Using equation (1), at $\gamma = \gamma_{peak}$, $\dot{\gamma} \propto \gamma^2(U_r + U_B) \sim constant$, i.e., at γ_{peak} the radiative cooling rate is nearly the same for all source.

The electrons must be accelerated in the regions where they produce radiation in jets since the timescales of radiative energy losses of electrons are very short to both synchrotron and Compton processes. Taking into account the accelerations, the continuity equation governing the temporal evolution of the electron density, $N(\gamma, t)$ in energy space thus is

$$\frac{\partial N(\gamma,t)}{\partial t} + \frac{\partial}{\partial \gamma}[(\dot{\gamma}_{acc} - \dot{\gamma}_{cool})N(\gamma,t)] - \frac{N(\gamma,t)}{t_{esc}} = Q(\gamma,t) \quad (3)$$

where $\dot{\gamma}_{acc}$ is electron acceleration rate, t_{esc} is the time scale for an electron to escape from the emitting region, and $Q(\gamma, t)$ is the electron injection rate. For $\dot{\gamma}_{acc} < \dot{\gamma}_{cool}$, because of strong radiative cooling, $N(\gamma, t)$ is much steeper than for $\dot{\gamma}_{acc} > \dot{\gamma}_{cool}$. The spectrum of $N(\gamma, t)$ breaks at γ_{break} when the electron heating rate balances to cooling rate,

$$\dot{\gamma}_{acc} = \dot{\gamma}_{cool}. \quad (4)$$

This break is responsible for both the synchrotron and inverse-Compton peaks, and thus γ_{break} is actually γ_{peak}. Assuming that the jet is composed of electrons and protons and that electrons are accelerated by shocks in the jet, $\dot{\gamma}_{acc}$ is independent of γ (Ghisellini 1999; Maraschi & Tavecchio 2001). According to equation (4), we have

$$\gamma_{peak} = \gamma_{break} = [(4/3m_e c)\sigma_T(U_r + U_B)\dot{\gamma}_{acc}]^{-1/2}, \quad (5)$$

which is consistent with equation (1), provided that the electron acceleration rate, $\dot{\gamma}_{acc}$ is on average similar in all sources. Equation (1) thus suggests that electron acceleration is the similar in both red and blue blazars, independent of luminosity, γ and U, and that only the cooling differs.

It is reasonable to assume that equation (1) can be applied to the kpc scale jet, and that the proportionality constant C is the same on small and large scales. The kpc scale jet is an extension of the inner jet. It can be treated as an "inner jet" originating from a weaker "nucleus", the end of the inner jet. Furthermore, on kpc scales the jet is still relativistic as indicated by the observed jet−counterjet intensity asymmetry. The differences between the kpc scale jet and the inner jet are not more than the differences between red and blue blazars. Probably, the electron acceleration mechanism in the kpc scale jet is the same as that in the inner jet, and the electron heating rate balances to the cooling rate.

2.2 The kpc scale synchrotron X-ray jet in blazars

Given the relativistic electrons with γ_{peak} in the magnetic field B of a relativistic jet with speed β in units of the speed of light, Lorentz factor Γ and viewing angle θ, the peak frequency ν_s of synchrotron radiation in the observer's frame is

$$\nu_s = 3.7 \times 10^6 \frac{\delta}{1+z} B\gamma_{peak}^2, \quad (6)$$

where z is the redshift of the source and δ is the Doppler factor $\delta = [\Gamma(1 - \beta\cos\theta)]^{-1}$. For blue blazars, $U_r(pc) \sim U_B(pc)$ in the inner jet. According to equation (1), $\gamma_{peak}^2 = C^2[2U_B(pc)]^{-1} = 4C^2\pi/B_{pc}^2$. Substituting γ_{peak} into equation (6) gives

$$\nu_s(pc) = 3.7 \times 10^6 \frac{\delta_{pc}}{1+z} \frac{4\pi C^2}{B_{pc}}. \quad (7)$$

On kpc scales, none of the jets shows any direct evidence of superluminal motion with one exception, M87 (Biretta et al. 1999), so in most cases the kpc scale jet probably has only mildly relativistic speeds, i.e., $\Gamma(kpc)$ is slightly

greater than 1. Since the inner jet is highly relativistic, the contribution of the emission from the inner jet to $U_r(kpc)$ cannot be neglected (Celotti et al. 2001). The emission of the kpc scale jet itself also contributes a main part of $U_r(kpc)$, but $U_r(kpc) \sim U_B(kpc)$ still holds. According to equation (1), $\gamma_{peak}^2 = C^2[2U_B(kpc)]^{-1} = 4C^2\pi/B_{kpc}^2$, yielding

$$\nu_s(kpc) = 3.7 \times 10^6 \frac{\delta_{kpc}}{1+z} \frac{4\pi C^2}{B_{kpc}}. \tag{8}$$

At point r of the jet, the magnetic flux L_B is

$$L_B = \pi\psi^2 r^2 c\Gamma^2 U_B = \frac{c}{8}\psi^2 r^2 \Gamma^2 B^2, \tag{9}$$

assuming that the emitting region located at r has a transverse dimension $R = \psi r$. Conservation of magnetic flux L_B gives

$$\frac{B(kpc)}{B(pc)} = \frac{\Gamma_{pc}}{1000\Gamma_{kpc}} \sim \frac{1}{100}, \tag{10}$$

i.e., $B(kpc) \sim 10^{-2}B(pc)$, assuming that ψ is constant along the jet and $\Gamma_{pc} \sim 10$ (Blandford 1993; Marcher 1993; Celotti et al. 2001). Combining equations (7), (8) and (10) yields

$$\nu_s(kpc) = 10^2 \nu_s(pc) \frac{\delta_{kpc}}{\delta_{pc}} \sim 10\nu_s(pc). \tag{11}$$

The synchrotron emission of blue blazars peaks in the range of ~0.01 keV to ~1 keV, so on kpc scales, the synchrotron peak of the jet emission lies above 0.1 keV, suggesting that a synchrotron X-ray jet exists in blue blazars on kpc scales, i.e. the parsec scale synchrotron X-ray jet continuously extends to kpc scales.

For red blazars, U_r/U_B is typically in the range of 10 – 100, so $\gamma_{peak}^2 \sim 4C^2\pi/(10B_{pc}^2)$, yielding

$$\nu_s(pc) = 3.7 \times 10^6 \frac{\delta_{pc}}{1+z} \frac{4\pi C^2}{10B_{pc}}. \tag{12}$$

On kpc scales, as in the case of blue blazars, $U_r(kpc) \sim U_B(kpc)$, and

$$\nu_s(kpc) = 3.7 \times 10^6 \frac{\delta_{kpc}}{1+z} \frac{4\pi C^2}{B_{kpc}}. \tag{13}$$

Thus

$$\nu_s(kpc) = 10^3 \nu_s(pc) \frac{\delta_{kpc}}{\delta_{pc}} \sim 10^2 \nu_s(pc). \quad (14)$$

The synchrotron peak of red blazars lies in the range of $\sim 10^{13}$ Hz to $\sim 10^{15}$ Hz, so on kpc scales, the synchrotron peak of the jet emission lies above ~ 0.01 keV, which is like those blue blazars whose synchrotron peak lies in the UV and soft X-ray energies, suggesting that a synchrotron X-ray jet exists in red blazars on kpc scales as well.

3 Discussion

In Section 2, it has been argued that a synchrotron X-ray jet exists in both blue and red blazars. Blue and red blazars are just the extrema of the blazar class. Similarly, a synchrotron X-ray jet exists in intermediate blazars on kpc scales. Furthermore, according to the unified schemes, Fanaroff-Riley class I radio galaxies (Fanaroff & Riley 1974) are intrinsically the same as BL Lac objects, and Fanaroff-Riley class II radio galaxies and steep-spectrum radio quasars (SSRQs) are intrinsically the same as FSRQs, with the relativistic jet oriented at a larger angle to the line of sight than blazars (e.g. Urry & Padovani 1995). Therefore, radio galaxies and SSRQs have a synchrotron X-ray jet on kpc scales as well. The X-ray jets in Cen A, M87, NGC 6251, Pictor A are thus synchrotron emission. Furthermore, Cen A and M87 are blue-blazar-like radio galaxies (Bai & Lee 2001), so their large scale X-ray jets should be a continuous extension of their parsec scale synchrotron X-ray jets, which is consistent with the *Chandra* recent observations for Cen A that shows the X-ray jet extending continuously from 10 parsec scales to kpc scales (Kraft et al. 2000). M87 is also very near, and the X-ray jet connecting parsec scale jet to kpc scale jet can be observed by *Chandra*, providing a test to the above result.

According to equations (8), (9) and (13), the synchrotron peak of the emission of the large scale jet shifts gradually to higher frequencies along the jet for both red and blue blazars. Consequently, in blue blazars and blue-blazar-like radio galaxies, the X-ray jet may get faint along the jet. In typical red blazars and red-blazar-like radio galaxies, from ~ 10 kpc to ~ 100 kpc along the jet, the synchrotron peak may lie rightly in the energy range of the CXO, resulting in a relatively bright X-ray jet. This may account for the X-rays observed from the "inner jet" of 3C 273 between 5 arcsecond and 10 arcsecond from the core (Marshall et al. 2001). The viewing angle θ of the jet of 3C 273 to the line of sight is $\cos\theta = 0.95$ (Davis et al. 1991), so at the distance of 3C 273 ($z = 0.158$), 5 arcsecond and 10 arcsecond from the core of 3C 273 correspond to ~ 38.43 kpc and ~ 76.86 kpc from the core along the jet,

respectively, assuming $H_0 = 75\mathrm{km}^{-1}\mathrm{Mpc}^{-1}$ and $q_0 = 0.5$. The X-rays from the "inner jet" of 3C 273 are thus probably synchrotron emission.

If in some cases (possibly in some powerful FSRQs) the jet is still highly relativistic on large scales, i.e., $\Gamma \sim \Gamma_{pc} \sim 10$, the energy density of the CMB, U_{CMB}, gets significant in the jet comoving frame. Assuming $L_B = 10^{45}$erg/s and redshift zero, at $\sim$10 kpc $U_{CMB} \sim U_B$, and at $\sim$100 kpc $U_{CMB} \sim 100U_B$ (Celotti et al. 2001). From $\sim$1 kpc to $\sim$10 kpc along the jet, the above analysis is still valid, except that equations (11) and (14) turn to be $\nu_s(kpc) = 10^3\nu_s(pc)$ and $\nu_s(kpc) = 10^4\nu_s(pc)$, respectively. On 100 kpc scales,

$$\nu_s(100kpc) \sim 10^{-1}\nu_s(10kpc), \tag{15}$$

indicating that the synchrotron peaks gradually shift to lower frequencies from 10 kpc to 100 kpc scales. For extreme (and thus the most powerful) red blazars, in which the synchrotron peak of the inner jet lies in the infrared bands, the synchrotron peak on 100 kpc scales may shift back to the optical bands, and consequently the X-ray jet gets fainter along the jet from 10 kpc to 100 kpc scales, while the optical and IR jet may get brighter. On larger scales ($>$ 100kpc) the X-rays from the jet may be dominated by inverse-Compton emission. The X-ray jet discovered by Chandra in PKS 0637-752 is about 0.5 Mpc from the nucleus (Schwartz, 2000; Celotti et al. 2001), and thus is probably dominated by inverse-Compton emission. If in some powerful Fanaroff-Riley class II radio galaxies (counterparts of powerful red blazars) the jet is still highly relativistic on large scales, and is viewed at a very large angle, Doppler beaming effects may cause the X-ray jet too weak to be detected.

4 Conclusions

In conclusion, we predict a detectable synchrotron X-ray jet on kpc scales in most radio-loud AGNs. In blue blazars and blue-blazar-like radio galaxies the kpc scale X-ray jet is a continuous extension of the parsec scale synchrotron X-ray jets and gets faint along the jet, which is consistent with the *Chandra* recent observations for the blue-blazar-like radio galaxy, Cen A. In typical red blazars and red-blazar-like radio galaxies the X-ray jet is bright on 10 kpc scales whether the jet is highly relativistic on large scales or not. In extreme red blazars in which the synchrotron peak of the inner jet lies in the infrared bands and the jet is still highly relativistic on large scales, the X-ray jet gets fainter along the jet from 10 kpc to 100 kpc scales while the optical and IR jet may get brighter, and on larger scales ($>\sim$ 100kpc, as in the case of PKS 0637$-$752) the X-rays from the jet may be dominated by inverse-Compton emission.

These predictions can be tested with the ongoing observations of the Chandra X-ray Observatory. Although blazars are characterized by core-dominant morphology on kpc scales, there are still a handful of blazars showing VLA jet. These are blue blazars 0414+009, 548−322, 2201+044, 2155−304 (Laurent-Muehleisen et al. 1993), 0829+089 and red blazars 3C 371 (Wrobel & Lind 1990), PKS 0521−365 (Keel 1986), 0954+658, 2007+777 (Kollgaard et al. 1992), 0752+258 (Antonucci & Ulvestad 1985). In all these sources we predict a detectable synchrotron X-ray jet on kpc scales. Apart from some powerful Fanaroff-Riley class II radio galaxies in which the jet may be highly relativistic on kpc scales ($\Gamma_{kpc} \sim 10$) and oriented at a very large angle to the line of sight, we also predict a detectable synchrotron X-ray jet on kpc scales in radio galaxies and steep-spectrum radio quasars. As time goes on, more and more radio-loud AGNs will be observed by the CXO. If a kpc scale X-ray jet is detected in each of these sources, the synchrotron origin of kpc scale X-ray jets will be verified.

Acknowledgments

This work was financially supported by the BK21 Project of the Korean government.

References

1. R.R.J. Antonucci, & J.S. Ulvestad, ApJ **294**, 158 (1985).
2. M.C. Begelman, R.D. Blandford, M.J. Rees, Rev. Mod. Phys. **56**, 255 (1984).
3. J. M. Bai, & M.G. Lee, ApJ **549**, L173 (2001).
4. J.A. Biretta, W.B. Sparks, & F. Macchetto, ApJ **520**, 621 (1999).
5. R.D. Blandford, & M.J. Rees, 1978, in *Pittsburgh Conference on BL Lac Objects* ed. A.N. Wolfe (Pittsburgh, University of Pittsburgh Press, 1978).
6. R.D. Blandford, in *Astrophysical Jets* ed. D. Burgarella, M. Livio, & C.P. O'Dea (Cambridge: Cambridge University Press, 1993).
7. R. D. Blandford, & A. Königl, ApJ **232**, 34 (1979).
8. A. Celotti, G. Ghisellini, & M. Chiaberge, MNRAS **321**, L1 (2001).
9. R.J. Davis, S.C. Unwin, & T.W.B. Muxlow, Nat **354**, 374 (1991).
10. E.D. Feigelson *et al*, ApJ **251**, 31 (1981).
11. G. Fossati *et al*, MNRAS **299**, 433 (1998).
12. G. Ghisellini *et al*, ApJ **407**, 65 (1993).
13. G. Ghisellini *et al*, MNRAS **301**, 451 (1998).
14. G. Ghisellini Astron. Nachr. **320**, 232 (1999).

15. D.E. Harris, & C.P. Stern, ApJ **313**, 136 (1987).
16. W.C. Keel, ApJ **302**, 296 (1986).
17. R.I. Kollgaard *et al*, AJ **104**, 1687 (1992).
18. R.I. Kollgaard, Vistas in Astronomy **38**, 29 (1994).
19. R.P. Kraft *et al*, ApJ **531**, L9 (2000).
20. R.A. Laing, in *Astrophysical Jets* ed. D. Burgarella, M. Livio, & C.P. O'Dea (Cambridge: Cambridge University Press, 1993).
21. H. Kubo *et al*, ApJ **504**, 693 (1998).
22. S.A. Laurent-Muehleisen *et al*, AJ **106**, 875 (1993).
23. K.-H. Mack, J. Kerp, and U. Klein, A&A **324**, 870 (1997).
24. H.L. Marshall *et al*, ApJ **549**, L167 (2001).
25. L. Maraschi, & F. Tavecchio, astro-ph/0107566 (2001).
26. A.P. Marscher, in *Astrophysical Jets* ed. D. Burgarella, M. Livio, & C.P. O'Dea (Cambridge: Cambridge University Press, 1993).
27. A.P. Marscher, ApJ **235**, 386 (1980).
28. M. Neumann *et al*, A&A **318**, 383 (1997).
29. P. Padovani, & P. Giommi, MNRAS **279**, 526 (1996).
30. E.S. Perlman *et al*, ApJ **551**, 206 (2001).
31. M.J. Rees, Nat **229**, 312 (1971).
32. M.J. Rees, ARA&A **22**, 471 (1984).
33. H.-J. Röser *et al*, A&A **360**, 99 (2000).
34. R.M. Sambruna *et al*, ApJ **463**, 444 (1996).
35. R.M. Sambruna *et al*, ApJ **549**, L161 (2001).
36. E.J. Schreier, P. Gorenstein, E.D. Feigelson, ApJ **261**, 42 (1982).
37. D.A. Schwartz *et al*, ApJ **540**, L69 (2000).
38. D.A. Schwartz, Chandra News **7**, 8 (2000).
39. F. Tavecchio *et al*, ApJ **544**, L23 (2000).
40. T.J. Turner *et al*, ApJ **475**, 118 (1997).
41. M.-H. Ulrich, L. Maraschi, & C.M. Urry, ARA&A **35**, 455 (1997).
42. C.M. Urry, P. Padovani, PASP **107**, 804 (1995).
43. C.M. Urry, Astropart. Phys. **11**, 159 (1999).
44. A.S. Wilson, A.J. Young, & P.L. Shopbell, ApJ **547**, 740 (2001).
45. J.M. Wrobel, & K.R. Lind, ApJ **348**, 135 (1990).

PART 2

Accretion Disk/Formation of Jets

This page is intentionally left blank

MAGNETIC STRESSES IN THE INNER REGIONS OF ACCRETION DISKS AROUND BLACK HOLES

J.H. KROLIK
Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
E-mail: jhk@pha.jhu.edu

In earlier work on angular momentum flow through accretion disks, the net angular momentum flux was treated as a freely-choosable boundary condition, and was determined by a heuristic argument that there should be no stress at the marginally stable orbit. Now that MHD turbulence is seen to be the primary agent of torque in disks, the stress at the marginally stable orbit is calculable, and it is found to be non-zero. Consequences include an increase in the energy available for radiation, reduced black hole spin-up, and generation of disk fluctuations.

1 Angular Momentum Flux in Accretion Disks

The most important job an accretion disk must perform is to transport angular momentum. In order for matter to move inward, it must lose its angular momentum; this angular momentum, once removed from its original owner, must then be carried away. The result is a flow of angular momentum from place to place within the disk, some attached directly to moving matter, some conveyed by non-material stresses such as those exerted by magnetic fields. Viewed in this way, it is clear that one of the fundamental tasks of a theory of accretion disks is to predict the character of this angular momentum flow.

In a time-steady disk, the total quantity of angular momentum stored in the disk is constant, but there can still be a net flux of angular momentum travelling radially through the disk; all that is required is that this flux be conserved. Thus, the condition of angular momentum conservation in a time-steady disk determines the structure only up to this arbitrary constant.

One way to identify the net angular momentum flux is to look at the angular momentum carried by accreting matter so close to the black hole that all dynamical communication between the accretion flow and the disk has been cut off. The angular momentum flux at that point is simply the accretion rate $\dot{M}$ times the mean specific angular momentum of accreted matter j_{in}. Thus, the net flux can be specified in terms of the mean angular momentum per unit mass carried into the black hole.

But where is the point where the accretion flow ceases to interact with the disk? In the first attempt to answer this question, Novikov and Thorne [1] suggested that the marginally stable orbit is the place where the accretion

flow decouples. From that point inward, the proper time required for matter to plunge all the way to the event horizon is at most a few orbital periods, so the inventory of mass stored between the marginally stable orbit and the event horizon can never be large compared to the mass is stored in the body of the disk, where inflow is very slow. Because they could not imagine how material with so little inertia could substantially influence a much heavier object, Novikov and Thorne argued that stress effectively ceases at r_{ms}, the radius of the marginally stable orbit. The specific angular momentum carried by matter through the event horizon would then be the angular momentum of a circular orbit at r_{ms}.

However, in a prescient footnote in a subsequent paper[2], Page and Thorne acknowledged a loophole to this argument:

"It is conceivable that the disk material might contain extremely strong magnetic fields, and that these fields might transport a torque from the infalling material at $r < r_{ms}$ to the disk at $r \geq r_{ms}$. In this case the boundary condition at r_{ms} would be modified, and the solution ... would be changed."

Page and Thorne understood that electromagnetic forces are not necessarily limited by the inertia of the matter in which the fields are embedded. Nonetheless, they, and nearly every worker in the field for the next twenty-five years, assumed that the stress associated with angular momentum transport effectively ceased at and inside r_{ms}. In fact, some sought to support this assumption by making use of the argument from dimensional analysis that this stress should scale with the local pressure[3]; if so, because the pressure falls rapidly as matter squirts away from the inner edge of the disk and into the plunging region, whatever agency caused the stress should disappear very close to this radius[4].

2 Magnetic Fields in Disks

However, thanks to the work of Balbus and Hawley (as summarized, for example, in their excellent review[5]), we now have very good reason to think that magnetic fields are not just important, but *essential* to the process of accretion. The argument may be boiled down to two simple statements: whenever magnetic fields are present in orbiting plasma whose angular frequency diminishes outward, they tend to exert a torque that transfers angular momentum outward; and there is an extremely robust magneto-rotational instability that leads to the creation of strong magnetic fields in exactly those circumstances, provided only that any externally-imposed magnetic field is weak.

It's easy to see why magnetic fields generically push angular momentum in the right direction. Except in the coldest disks, the accreting matter is sufficiently ionized as to be an excellent electrical conductor, and magnetic flux is therefore frozen into the plasma. At the same time, in just about every likely astronomical potential, and in particular in point-mass potentials, the angular frequency of a circular orbit is greater at small radii than large. When that's the case, any field line linking material at different radii will be stretched as the material at smaller radius orbits faster than the material farther out. The effective magnetic "tension" then creates a force that tries to slow down the motion of the faster-moving fluid element and accelerate the slower; that is, it tries to move angular momentum from smaller radius to larger. Equivalently, the fact that $d\Omega/dr < 0$ means that magnetic field components are correlated in the sense that $\langle B_r B_\phi \rangle < 0$ (here r and ϕ refer to cylindrical radius and azimuthal angle, respectively).

The argument of the previous paragraph demonstrates that magnetic field torques have the right sign; the second step is to show that they're strong enough to be interesting. Independently rediscovering an instability found decades earlier by Velikhov [6] and by Chandrasekhar [7], Balbus and Hawley [8] observed that poloidal MHD perturbations in rotating plasma grow exponentially on the rotational timescale whenever $d\Omega/dr < 0$ and $v_A < h\Omega$, where v_A is the Alfvén speed and h is the thickness of the plasma in the direction along the rotational axis. The second of these two criteria, when applied to an accretion disk in vertical hydrostatic balance, amounts to the condition that $B^2/8\pi$ is smaller than the pressure p. Since both these criteria apply in the vast majority of cases, we should then expect that the magneto-rotational instability is a powerful source of turbulence in just about every disk. A host of nonlinear numerical simulations (e.g., Stone et al. [9] and Brandenburg et al. [10]) has since shown that the instability saturates at a level such that $\langle B_r B_\phi \rangle / p \sim 10^{-2} - 10^{-1}$. Moreover, as expected, the correlation between B_r and B_ϕ is a significant fraction of the total field intensity: $\langle B_r B_\phi / B^2 \rangle$ is typically a few tenths.

These simulations have been conducted in a somewhat limited setting, however. For the most part, they have assumed a "shearing box" geometry. That is, in order to simulate an annular section of a disk, they approximate the system of rotating coordinates by a box with an imposed shear and "radial" boundary conditions that are quasi-periodic. Implicitly, then, they assume that the time required for the turbulence to equilibrate, i.e. for stirring on large scales by the magneto-rotational instability to be balanced by dissipation on small scales, is short compared to an inflow time—put another way, in these simulations, the inflow time is effectively infinite.

In a real disk, this assumption is appropriate in most locations. The

characteristic dissipation time in the dynamically stable part of the disk is $\sim P_{orb}/(\alpha\beta)$, where P_{orb} is the orbital period, α is the ratio of stress to pressure, and β is the ratio of total pressure to magnetic energy density. On the other hand, the inflow time is $\sim (r/h)^2 P_{orb}/\alpha$, so that the ratio of dissipation time to inflow time is $\sim \beta^{-1}(h/r)^2$. For typical values of β ($\sim 10 - 100$) and h/r (< 1), inflow is indeed sufficiently slow that there is plenty of time for the turbulence to equilibrate.

However, near r_{ms} the sense of this comparison changes. Although the dissipation time scales in much the same way, the inflow time becomes drastically shorter—after all, just inside r_{ms}, the inflow time is $\sim P_{orb}$. In this region, then, there is no longer time for the amplitude of the turbulence to equilibrate. Rather, the magnetic field is effectively flux-frozen and simply adopts whatever strength results from accepting the magnetic flux brought in from outer regions.

Flux-freezing near r_{ms} predicts, then, that the field strength should be not terribly different from its value at slightly larger radii. Because the sense of orbital shear doesn't change, there is still a strong correlation between B_r and B_ϕ, and therefore the stress is likewise about the same as at larger radii. *But if this is so, there is no reason for the stress to disappear at* r_{ms}.[11]

3 Consequences of Non-zero Stress at the Marginally Stable Orbit

Given the fundamental importance of angular momentum flow to accretion disks, it should not be surprising that altering the inner stress boundary condition, and therefore the net angular momentum flux, should have a large number of ramifications. The most basic of these is that continued force within r_{ms} means not only continuing torque, but also continuing work done by falling material on its neighbors. In the traditional Novikov-Thorne picture, matter falls on free-fall trajectories from r_{ms} to the event horizon; free-fall implies constant energy and angular momentum, so the energy and angular momentum it has upon entering the black hole are identical to what it had at r_{ms}. The total energy released in the course of accretion, and therefore potentially available for radiation, is then the binding energy of a circular orbit at r_{ms}.

However, if matter is subject to continuing forces within r_{ms}, neither its energy nor its angular momentum need stay constant. Given the sign of angular momentum transfer, the net effect should on average be for matter to lose both energy and angular momentum. It would then enter the black hole with less of both, releasing more energy for possible radiation and retarding spin-up of the black hole. Gammie[12] has solved a fully general-relativistic toy-model in which he imagines steady flow solely in the equatorial plane; he found that

reasonable magnetic field strengths at r_{ms} could lead to substantial increases in the effective accretion efficiency, especially for rapidly-rotating black holes. Order-unity increases in efficiency could then lead[13] to spin equilibrium for the black hole at values of a/M considerably less than the limiting value predicted by Thorne[14].

A second corollary of non-zero stress at r_{ms} is that there should also be non-zero dissipation there. Wherever there is stress in the disk, there is also work done; however, the work that matter does in the course of losing enough angular momentum to fall in does not in general match how much potential energy it loses in the same fall. Any excess lost energy must be dissipated. Thus, in contrast to the Novikov-Thorne model, which predicts that radiative disks should become cooler as one approaches r_{ms} from the outside, non-zero stress near r_{ms} enforces continued heating in this region. If the radiation time is shorter than the infall time, there should then be a supplementary high-frequency contribution to the disk's emitted spectrum[13].

When extra radiation is produced in the innermost rings of the disk, relativistic effects on this radiation are especially strong. In the Novikov-Thorne model, very little of the radiation emitted from disks around slowly-spinning black holes is bent back toward the disk, although the fraction of "returning radiation" increases as a/M rises toward unity. Photons made not far outside r_{ms} are, however, subject to strong gravitational trajectory-bending, and a sizable fraction of them find their paths bent back down so that they strike the disk (some also, of course, follow geodesics that lead to the event horizon and oblivion). It is possible, then, that much of the energy dissipated at larger radius comes not from decay of local turbulence, but rather from turbulence generated at very small radii whose energy has been transferred farther out in the disk by photons[13]. Processes like this may help explain the extraordinarily close coordination of fluctuations in different wavelength bands seen in many active galactic nuclei.

A third consequence is entailed by the fact that all these stresses are intrinsically turbulent. The energy and angular momentum per unit mass delivered to the black hole can therefore be expected to fluctuate from place to place and from time to time. Similarly, any additional stress and dissipation in the disk will likewise fluctuate. These fluctuations may provide a noise source capable of explaining the continual variations in output that are the hallmark of accreting black hole lightcurves.

Finally, there has long been interest in the possibility of extracting energy from rotating black holes *in the absence of accretion* via magnetic effects[15]. In the context of these studies it is often assumed that the magnetic field structure is "force-free"; that is, the plasma inertia is so small that the magnetic

field adopts whatever form minimizes the field energy without regard for what the matter is doing. This assumption is valid when $B^2/8\pi \gg \rho c^2$. In the picture suggested by magnetically-regulated accretion, although the force-free approximation may be appropriate at altitudes high above the disk midplane, field loops are controlled by dense accreting matter in the plane. The boundary conditions determining the structure of the force-free region are therefore set in a region where the assumption does not apply, or is at most marginal [11].

4 Simulations

Although the existence of non-zero stress at r_{ms} seems likely on the basis of the arguments raised so far, it is still necessary to quantitatively compute how large it might be. The only known technique that offers any promise for progress is large-scale numerical simulation. Moreover, unlike the simulations described earlier, these obviously cannot be confined in shearing boxes; in order to have any bearing on these issues, they must cover a span of radii ranging from immediately outside the event horizon to at least several times r_{ms}. Such efforts are, of necessity, extremely computationally demanding.

The problems are made harder by two additional considerations. Unlike the earlier shearing-box simulations, in which following the flow of angular momentum was the critical issue, for these following the flow of energy is equally critical. That means dissipation must be treated in some sensible way, radiation must be computed at a realistic rate, and photon diffusion must also be properly tracked. In the vicinity of r_{ms}, all these processes take place on timescales not so different from a dynamical time, so crude approximations could be seriously misleading.

In addition, likewise unlike the shearing-box simulations, the action of interest here takes place, by definition, deep in the relativistic gravitational potential. Newtonian gravity was a perfectly reasonable approximation for examining annular segments of disks far outside the marginally stable orbit, but assuming it in this context would eliminate the central element of dynamically unstable orbits at radii inside r_{ms}. Genuine general relativistic dynamics are therefore essential for a true grasp of this problem.

Unfortunately, none of the work published to date has surmounted either of these additional complications. So far, in no treatment is the energy equation solved. Instead, the internal energy of the gas is simply assumed to follow an adiabatic equation of state for which the specific entropy is constant unless a shock occurs. Similarly, no general relativistic 3-d MHD code yet written has sufficient efficiency and numerical stability to be used for this problem (although several groups are hard at work to develop one). Instead, every

attempt to date has adopted Newtonian physics, but in a potential that crudely approximates dynamics in a Schwarzschild metric. The "Paczynski-Wiita" potential [16] has the "quantum defect" form $U = -GM/(r - 2GM/c^2)$, which gives it a marginally stable orbit at $r = 6GM/c^2$, the same radial coordinate at which the marginally stable orbit may be found around a non-rotating black hole.

Despite these failings, simulations of this sort may still serve as guides to what may occur, particularly the distribution of stress. Perhaps the most comprehensive of the simulations done to date can be found in my work with John Hawley [17,18], although qualitatively similar results have been found by others[19].

The key accomplishment of this effort is displayed in Figure 1. As that figure shows, the azimuthally- and vertically-averaged magnetic stress behaves quite smoothly as a function of radius, showing absolutely no sign of dying at r_{ms}. If anything, the stress increases somewhat with decreasing radius through the plunging region. At the same time, although the stress and the pressure roughly vary together in the disk body, the ratio of magnetic stress to pressure increases sharply inside $\simeq 2r_{ms}$. In other words, the "α"-parameter is *not* constant in the inner disk. In fact, in the higher-resolution simulation shown in this figure, the effective α increases by two orders of magnitude from $R = 10$ to the inner problem boundary at $R = 1.25$!

Consistent with this continuing stress, the simulations show that the specific angular momentum and energy of plunging matter continue to change, sometimes dramatically, well inside the marginally stable orbit. Some matter nearly doubles in binding energy before reaching the inner boundary of the simulation; other matter becomes nearly unbound. Whether any of these changes in energy result in significant radiative or dynamical consequences is not well-determined by these simulations for the reasons discussed earlier—they do not track either dissipation or radiation. It is entirely possible that, while energy is exchanged violently between different fluid elements in the plunging region, nearly all of it may simply ride along with those fluid elements until they disappear inside the black hole.

On the other hand, angular momentum is accurately accounted in these simulations, and the mean specific angular momentum of matter does steadily fall as it plunges inward from r_{ms} (Figure 2). And, of course, any angular momentum lost by accreting matter must be delivered back to the disk. Although the fractional change in this simulation ($\simeq 5\%$) is small, it is clearly non-zero. It is also quite possible that the limitations of this simulation cause this effect to be underestimated. In addition to the physical limitations already discussed, we have found that improving grid resolution has led to steadily

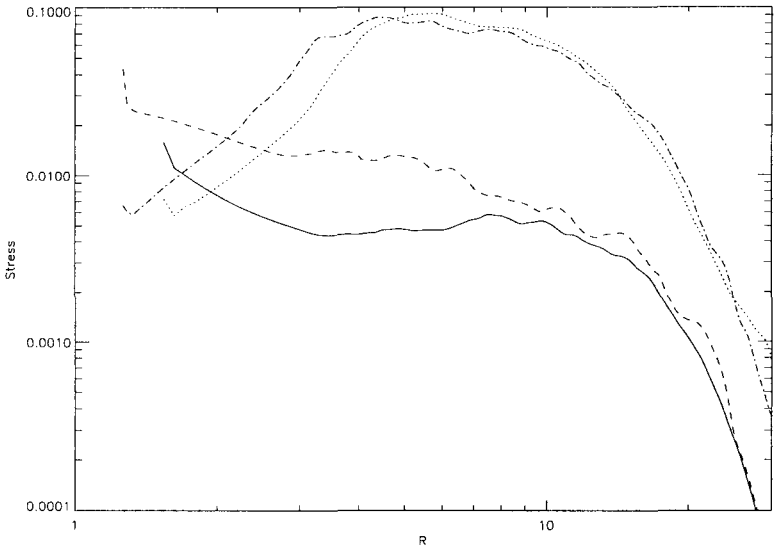


Figure 1: Vertically-integrated and azimuthally-averaged pressure and magnetic stress at late time in two simulations whose initial field configurations were purely poloidal. The solid curve shows the R-ϕ magnetic stress in the simulation of Hawley & Krolik (2001); the dotted curve shows the pressure in the same simulation. The dashed and dot-dashed curves show the magnetic stress and pressure, respectively, in the higher-resolution simulation of Hawley & Krolik (2002). In this figure the length unit is $2GM/c^2$, so $R_{ms} = 3$.

growing magnetic stresses (Figure 1).

To end this quick summary of simulation results, it is appropriate to display a measure of the fluctuations seen in these highly turbulent systems (Figure 3). The accretion rate through the simulation's inner boundary varies by as much as a factor of three; even its rms fluctuations are $\simeq 1/3$ of the mean accretion rate. Fluctuations in the luminosity are, however, only indirectly tied to fluctuations in the accretion rate. The energy lost in the course of accretion must first be dissipated into heat, individual particles must transfer their thermal energy into photons, and those photons must then diffuse out of the flow. Thus, the lightcurve may be better described as a convolution of the probability distribution for photon diffusive escape with the history of stress. For this reason, we also computed the Fourier power spectrum of the volume-integrated Maxwell stress, also shown in Figure 3. Both Fourier power spectra are broad and smooth, and exhibit substantial power at frequencies well below the orbital frequency at r_{ms}, which is $\simeq 0.05$ in these units.

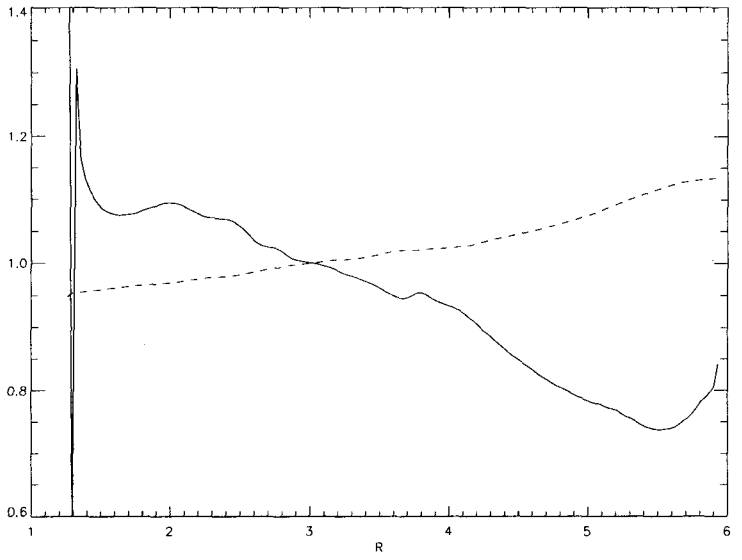


Figure 2: Mass-flux weighted azimuthally- and vertically-averaged specific binding energy (solid curve) and angular momentum (dashed curve) at the same time in the higher-resolution simulation shown in fig. 1. Both curves are normalized to their values at $R = 3$ to emphasize the continuing change at smaller radii. The spikes in the binding energy just outside the inner radius of the simulation are artifacts.

A closer look at the spatial distribution of these fluctuations reveals that the region around r_{ms} is a prolific source of "noise". The amplitude of fluctuations is greatest in the marginally stable region, and wave-trains can be seen propagating out through the disk from origins near r_{ms}. At a qualitative level, the special ability of this zone to create fluctuations can be readily understood, although we have not yet created a quantitative theory of these processes. Consider, for example, two adjacent fluid elements, threaded by a magnetic flux tube and orbiting just outside r_{ms}. If one of them interacts with some other nearby matter and loses a small amount of angular momentum, it will start to fall in. As it does so, its angular velocity increases simply due to conservation of angular momentum. The flux tube connecting it to its partner then mediates a torque that transfers angular momentum from the inner fluid element to the outer. That much would occur if the same events occurred in the stable part of the disk. However, because the inner fluid element now finds itself in the unstable orbit region, its orbital velocity becomes more and more

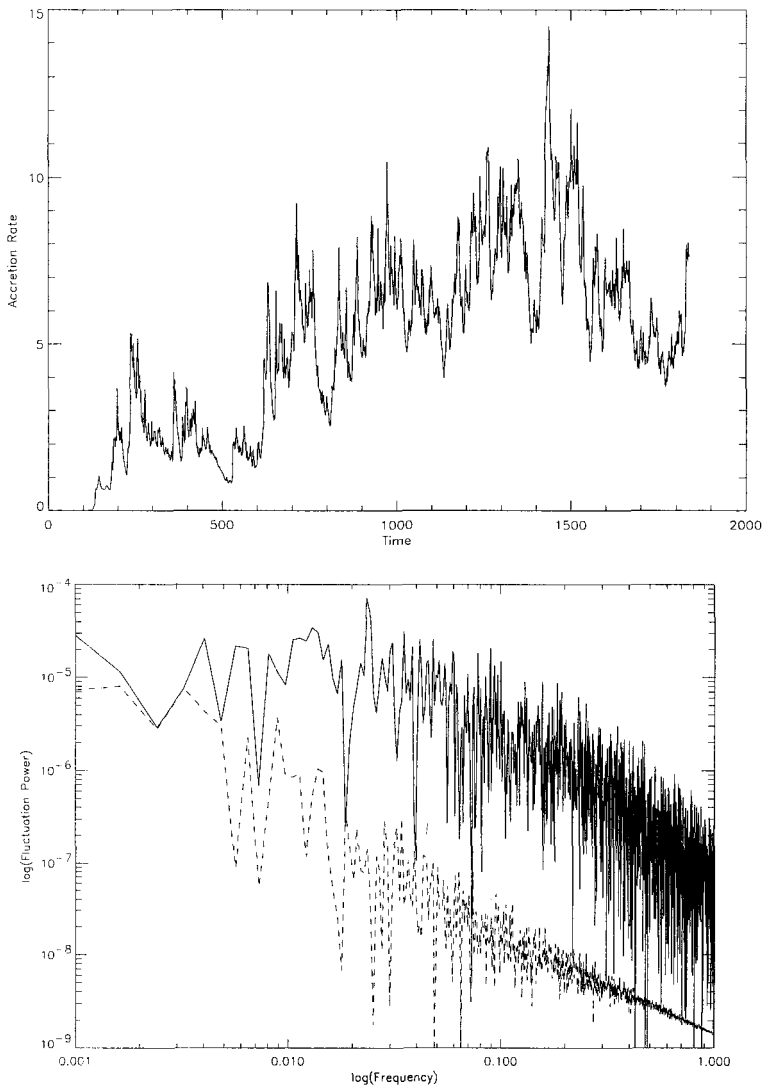


Figure 3: Upper panel: Mass accretion rate at the inner edge as a function of time in the higher-resolution simulation shown in fig. 1 and in fig. 2. Lower panel: Solid line is Fourier power density per logarithmic frequency interval of the accretion rate into the black hole (i.e. $\log(f|\widehat{\dot{M}}|^2)$. Dashed line is Fourier power in the same units for the volume-integrated Maxwell stress. To avoid transients associated with the initial start-up and linear growth phase of the magneto-rotational instability, the spectrum is computed for $t \geq 600$ time units.

radial, accentuating the torque. In essence, the ordinary magneto-rotational instability, which automatically results in an outward flow of angular momentum, acquires an "after-burner". Thus, the very nature of orbital dynamics near r_{ms} makes it a source of fluid "noise".

5 Summary

The recognition that MHD turbulence is the basic engine behind transport of angular momentum in accretion disks has enabled us to finally begin working toward a genuinely physical understanding of what happens in accretion flows. Now that we can build our picture of disk dynamics on fundamental physics, it has become possible to predict with some confidence what stresses occur where and with what strength.

In particular, the importance of magnetic stresses to accretion is now forcing us to move beyond the classical Novikov-Thorne model of energy and angular momentum release in relativistic accretion. Contrary to their (once plausible) assumption, stress does *not* cease at the marginally stable orbit.

Numerous consequences flow from this simple revision of a boundary condition. The radiative efficiency of accretion may be rather greater than initially envisioned, and may also be time-variable, as it is modulated by continuing MHD turbulence. Black hole spin-up may be rather slower, so that the equilibrium rotation rate achieved by black holes after long periods of accretion may be lower than previously thought. The spectra radiated by matter accreting onto black holes may have greater contributions from hotter regions. And the ubiquitous fluctuations seen in the lightcurves of accreting black holes, whether in AGN or Galactic binaries, may find their explanation in MHD turbulence in the vicinity of the marginally stable orbits around their central objects.

Acknowledgments

This work was partially supported by NASA grant NAG5-9187 to JHK.

References

1. Novikov, I.D. & Thorne, K.S. 1973, in *Black Holes*, eds. C. De Witt and B. De Witt (New York: Gordon & Breach), p. 343
2. Page, D. & Thorne, K.S., *Ap. J.* **191**, 499 (1974)
3. Shakura, N.I. & Sunyaev, R.A., *Astron. Astrop.* **24**, 337 (1973)
4. Abramowicz, M.A. & Kato, S., *Ap. J.* **336**, 304 (1989)
5. Balbus, S.A. & Hawley, J.F., *Revs. Mod. Phys.* **70**, 1 (1998)
6. Velikhov, E., *J.E.T.P.* **36**, 1398 (1959)

7. Chandrasekhar, S., *Hydrodynamic and Hydromagnetic Stability* (New York: Dover) (1981)
8. Balbus, S.A. & Hawley, J.H., *Ap. J.* **376**, 214 (1991)
9. Stone, J.M., Hawley, J.F., Gammie, C.F. & Balbus, S.A., *Ap. J.* **463**, 656 (1996)
10. Brandenburg, A., Nordlund, A., Stein, R.F. & Torkelsson, U., *Ap. J. Letts.* **458**, L45 (1996)
11. Krolik, J.H., *Ap. J. Letts.* **515**, L73 (1999)
12. Gammie, C.F., *Ap. J. Letts.* **522**, L57 (1999)
13. Agol, E. & Krolik, J.H., *Ap. J.* **528**, 161 (2000)
14. Thorne, K.S., *Ap. J.* **191**, 507 (1974)
15. Blandford, R.D. & Znajek, R.L., *M.N.R.A.S.* **179**, 433 (1977)
16. Paczyński, B. & Wiita, P.J., *Astron. Astrop.* **88**, , (23)1980
17. Hawley, J.F. & Krolik, J.H., *Ap. J.* **548**, 348 (2001)
18. Hawley, J.F. & Krolik, J.H., *Ap. J.* in press (2002)
19. Reynolds, C.S. & Armitage, P.J., *Ap. J. Letts.* **561**, L81 (2001)

RAYLEIGH SCATTERED Lyα IN ACTIVE GALACTIC NUCLEI

HEE-WON LEE
Department of Geoinformation Sciences
Sejong University, Seoul, Korea
hwlee@kunja.sejong.ac.kr

We compute the polarization of the broad Lyα emission line, which is Rayleigh-scattered from high column neutral components including the broad emission line region itself and the accretion disk in active galactic nuclei. Depending on the anisotropy of the radiation field in the broad emission line region and relative geometry with respect to the accretion disk, a large range of degree of polarization up to 10 percent is obtained. It is argued that the accretion disk reflection, which gives the polarization perpendicular to the disk normal axis, dominantly contributes to the polarized broad Lyα flux. We estimate the polarized flux Rayleigh-scattered in the broad emission line region that is modeled by a large number of spherical clouds distributed mainly in the equatorial region, and find that only minor contribution to the total polarized flux is obtained. Our results are applied to the polarized Lyα of the radio quiet quasar PG 1630+377.

1 Introduction

Active galactic nuclei (AGN) are powered by the accretion process around a supermassive black hole in the form of a geometrically thin but optically thick disk (Blandford 1990). Prominent emission lines with a typical width of 10^4 km s^{-1} are regarded as one of the distinguishing characteristics of AGN. One of the most widely used models for the broad emission line region (BELR) is the cloud model, where the BELR is composed of a large number of small clouds with column densities 10^{22-24} cm^{-2} and has a small covering factor ~ 0.1. Alternative models of the BELR have been presented by a number of researchers (e.g. Murray et al. 1995)

Spectropolarimetry is regarded as an important tool in testing the existence of an accretion disk in active galactic nuclei (AGN) and constraining the unification models (e.g. Miller & Goodrich 1990, Antonucci 1993, Lee & Blandford 1997). Spectropolarimetric observations of several quasars by the Hubble Space Telescope showed a steep increase of the degree of polarization blueward of the Lyman edge (Koratkar et al. 1998). In particular, in the quasar PG 1630+377, up to 20 percent of degree of polarization was observed. One interesting point is that the broad Lyα feature was also strongly polarized with the degree of polarization 7 per cent after correction for unpolarized continuum, whereas no evidence of strong polarization was found in other lines (Koratkar et al. 1995).

Korista & Ferland (1998) discussed the ultraviolet-optical albedo of the

BELR in the context of the cloud model. They emphasized that the scattering albedo around the hydrogen Lyα is particularly large, which is attributed to the Rayleigh scattering by a high column neutral hydrogen component with a minor additional contribution from Thomson scattering. The exact amount of neutral hydrogen in the broad line region is not constrained strongly by observation (e.g. Dickey 1988, Cole et al. 1998). The high column density component of H I has also been discussed in the context of spectropolarimetry of Hα as a theoretical possibility of broad polarized flux seen in several narrow line AGN (Lee & Yun 1998).

Because of a small but not negligible covering factor of the BELR, it is expected that the broad Lyα feature may contain the flux that is Rayleigh-scattered in the BELR. This may lead to a polarized Lyα BEL, if the BELR possesses an anisotropic kinematics or take a non-spherical geometric shape. It is also a theoretical possibility to assign a large column component in the accretion disk itself, where the BEL photons can be reflected at the surface. Because of the lack of the knowledge on the detailed physical structure of the accretion disk, a simple model is adopted in this work to estimate the polarization of Lyα from the Rayleigh scattering process.

2 Atomic Physics of Rayleigh Scattering

The Rayleigh scattering processes have been studied by many researchers (e.g. Isliker et al. 1989, Gavrila 1967, Sadeghpour & Dalgarno 1992). According to the famous Kramers-Heisenberg formula, the Rayleigh scattering cross section can be written as

$$\frac{d\sigma_{Ray}}{d\Omega} = r_0^2 \left| (\epsilon^{\alpha} \cdot \epsilon^{\alpha'}) \left(\sum_n M^b(n) + \int_0^{\infty} dn' M^c(n') \right) \right|^2 , \qquad (1)$$

where ϵ^{α} and $\epsilon^{\alpha'}$ are the polarization vectors associated with the incident and outgoing photons, respectively, and r_0 is the classical electron radius. Here, the matrix elements for the bound and continuum states are given by

$$M^b(n) = \frac{m\omega^2}{3\hbar} | < 1s \parallel r \parallel np > |^2 \frac{2\omega_{n1}}{\omega_{n1}^2 - \omega^2} \qquad (2)$$

$$M^c(n) = \frac{m\omega^2}{3\hbar} | < 1s \parallel r \parallel n'p > |^2 \frac{2\omega_{n'1}}{\omega_{n1}^2 - \omega^2}. \qquad (3)$$

The explicit reduced matrix elements can be found in many text books (e.g. Berestetskii, Lifshitz, & Pitaevskii 1972, Saslow & Mills 1969, Lee & Lee 1997). The scattering cross section for a H I component of column density 10^{23} cm^{-2}

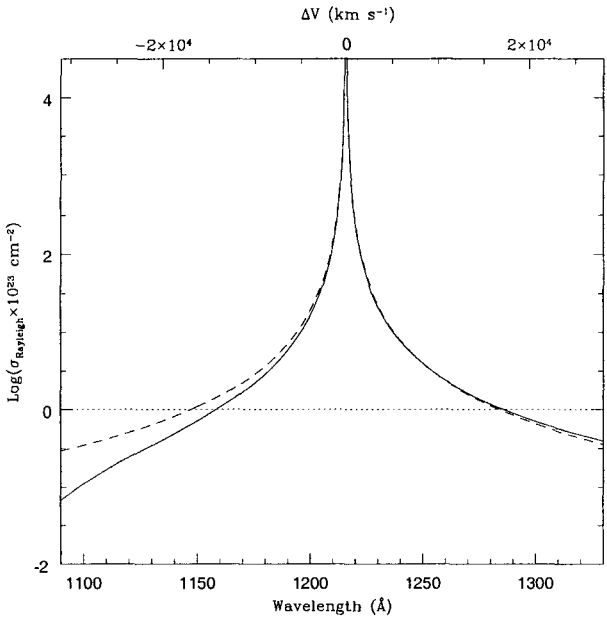


Figure 1: The Rayleigh scattering optical depth of a hydrogen slab of column density 10^{23} cm^{-2} as a function of the incident wavelength. The solid line represents the true value and the dotted one is obtained using the approximate relation given in Eq. 4 in the text. An excellent agreement is seen around the line center and a slight asymmetry is seen when $\Delta V \geq 10^4$ km s^{-1}. Note that the vertical scale is logarithmic.

is plotted in Fig. 1. A useful approximation to the Rayleigh scattering cross section can be also found in Ferland (1996). By taking the first term in the summation that represents the $2p$ state contribution, we see that an approximate functional relation of the column density with respect to the scattering band width ΔV is given by

$$\frac{\Delta V}{10^4 \text{ km s}^{-1}} \simeq \left[\frac{N_{HI}}{3 \times 10^{22} \text{ cm}^{-2}}\right]^{1/2} \tag{4}$$

in the range $\Delta V \leq 10^4$ km s^{-1}. The quality of this approximation is also illustrated in Fig. 1.

The cross section is asymmetrical with respect to the Lyα line center and it

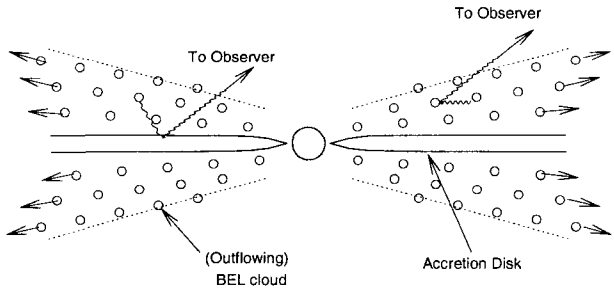


Figure 2: Schematic diagram of the broad emission line region adopted in this work. Broad emission line clouds represented by small circles are assumed to be located mainly equatorially and outflowing according to Eqs. 5, 6.

is larger redward of Lyα than in the blue part. Because the dynamical velocity scale of the BELR is about 10^4 km s^{-1}, most of the broad Lyα line photons can be Rayleigh-scattered in the BELR under the assumption that the H I column density is no less than 10^{22} cm^{-2}. Furthermore, since N V 1240 lies 5900 km s^{-1} redward of Lyα, it is also possible that the N V line photons can be Rayleigh-scattered by H I atoms inside the BELR.

3 Anisotropic Model of the BELR

In this work, we calculate the Lyα polarization adopting a simple kinematic BELR models with an anisotropic distribution of line emitting clouds, which also produces typical single-peak logarithmic profiles observed in many AGN (e.g. Netzer 1990).

We assume that the clouds are distributed mainly in the equatorial region

and outflowing radially. Validity of this assumption is in question, but it is not expected that the overall polarization computation will be sensitively dependent on the detailed dynamic models. The velocity law is assumed to be given by

$$\mathbf{v}(\mathbf{r}) = v_0(r/R_0)\hat{\mathbf{r}}. \tag{5}$$

The number density distribution of line emitting clouds is assumed to be given by

$$n(r) = \begin{cases} n_0(r/R_0)^{-3}, & \text{if } \mu \leq 0.1 \text{ and } R_i \leq r \leq R_o \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

to satisfy the mass conservation. Here, $\mu \equiv \cos\theta$ is the cosine of the polar angle. Here, R_i and R_o are the inner and outer radius of the BELR, and n_0 is the number density of line emitting clouds at the inner radius of the BELR. In this work we set $R_o = 10^{17}$ cm and $R_i = 10^{-2}R_o$. Netzer (1990) discusses the conditions for producing logarithmic profiles and this is one of many cases that satisfies the conditions.

In order to obtain a covering factor $C_{BELR} = 0.1$ of the BELR from this distribution, another constraint on n_0 should be imposed, that is, every line of sight with $|\mu| < 0.1$ must hit at least one cloud. Assuming that each cloud is a uniform sphere of radius $R_c = 10^{13}$ cm, we obtain

$$n_0 \geq (\pi R_c^2 R_i[1 - R_i/R_0])^{-1}. \tag{7}$$

From this, the total number of clouds N_{tot} is given by

$$N_{tot} \geq 0.4\ \pi n_0 R_0 R_i^2[1 - R_i/R_o] = 0.4 \times 10^6 \tag{8}$$

3.1 Rayleigh reflection from the accretion disk

A geometrically thin and optically thick accretion disk is often invoked to explain the X-ray cold component. A corona component above the surface of the accretion disk is also believed to exist. The detailed structure of the accretion disk is not known observationally, though. It seems natural that the accretion disk surface may reflect the radiation incident from the BELR.

In this paper, we assume that the line photons from BEL clouds are reflected from the accretion disk surface. Here, we only consider the Rayleigh-scattering by H I, neglecting other scattering processes. It is assumed for simplicity that the dimension of the accretion disk is comparable to that of the BELR, so that the disk surface effectively cover the half of the sky viewed from each single cloud.

Any anisotropy in the radiation field in the BELR may affect significantly the polarization especially in the case of a reflected radiation field. An

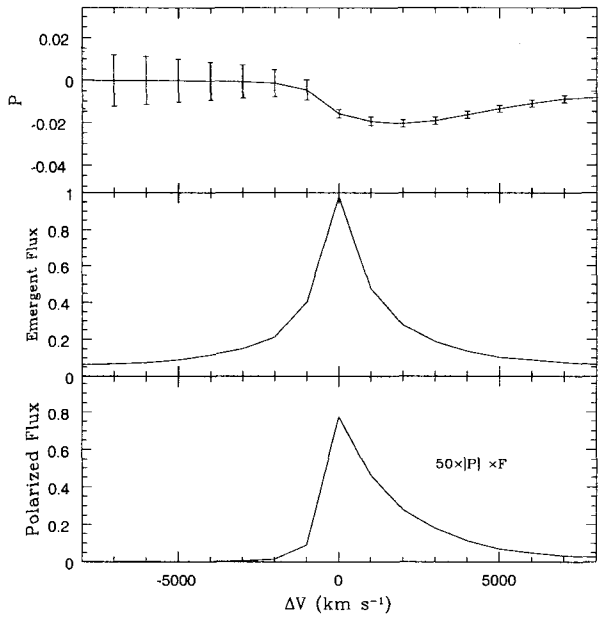


Figure 3: Polarization of the Rayleigh-scattered broad Lyα in the broad emission line region. The degree of polarization is typically 2 percent in the red part and 0.5 percent in the blue part. This asymmetry in the degree of polarization is illustrated again in the polarized flux which shows a significant excess in the red part.

anisotropic radiation field is expected in the cloud picture of the BELR, where line emitting clouds have a large column density in order to remain optically thick in the vicinity of the central engine of AGN (see Ferland et al. 1992). In this situation, the broad emission lines are interpreted to originate from the illuminated side of line emitting clouds. We will relax the assumption of isotropically radiating clouds and consider the effect of anisotropy of the BELR radiation field.

4 Result

4.1 Rayleigh-scattering from the BELR

In Fig. 3 the degree of polarization, scattered flux and polarized flux are plotted from the BELR model described in the previous section. It is seen that the

degree of polarization up to 2 percent is obtained in the red part. However, in the blue part the degree of polarization is ~ 0.5 percent. which is significantly smaller than that in the red part. This is explained by the assumption of the outflowing motion of the BEL clouds. The conspicuous red asymmetry is also illustrated in the polarized flux.

In this work, the polarization direction is represented by the sign of the degree of polarization, where the positive sign is for the polarization in the direction perpendicular to the normal axis of the accretion disk and the negative sign for the polarization in the direction parallel to the normal axis. As is expected in the equatorial distribution of BEL clouds adopted in this work, the polarization is in the direction parallel to the accretion disk normal axis.

Note also that the total flux has a single-peak profile and that the scattered flux also exhibits a single-peak profile. Even though the BELR is assumed to have an expanding (outflowing) motion, the position of the maximum in the scattered flux almost coincides with the line center. This result is a consequence of a low filling factor of the BEL emitting clouds in the BELR and the assumption of isotropic line emission of the BEL clouds. That is, all the BEL emitting clouds can be seen by the observer and therefore the red-shifting clouds contribute to the scattered and polarized fluxes almost as equally as the blue-shifting clouds do.

A more realistic model can be constructed by relaxing the assumption of isotropically emitting BEL clouds. Instead by constraining so that Lyα photons are emitted from the illuminated side we may enhance the polarization of the red part of the scattered flux and therefore strengthen the red asymmetry. However, a more dominant contribution is expected to be made by the N V 1240 line flux, which is located at 5900 km s^{-1} redward of Lyα.

4.2 Polarization of the radiation Rayleigh-reflected from a disk surface

Since we assume that the H I column density in the accretion disk is high enough, the incident radiation is treated to be monochromatic and the scattering medium is of scattering optical depth $\tau = 10$. The frequency redistribution of the scattered radiation is also neglected because it is order of magnitude smaller than the dynamical velocity scale of the BELR.

In Fig. 4, we show the composite polarized spectra of the flux component scattered in the anisotropical BELR and the accretion disk reflected component. We assume that the Rayleigh reflection component is polarized with the degree of polarization 10 percent in the direction perpendicular to the disk normal axis. Due to the dominant contribution from the Rayleigh reflection, the overall polarization direction is perpendicular to the disk normal

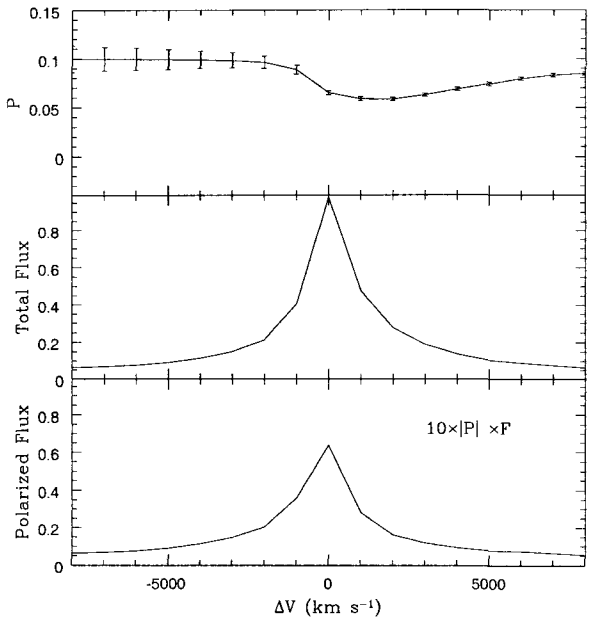


Figure 4: Synthetic polarization of the component Rayleigh-reflected from the accretion disk surface and the component Rayleigh-scattered in the broad emission line region. For the Rayleigh-reflection component it is assumed to be polarized by 10 percent in the direction perpendicular to the disk normal axis.i See the text for more detail.

axis. However, because of the opposite polarization directions between the two components, the polarization is weaker in the red part than in the blue part.

It is also noted that this result is highly sensitive to the anisotropy of the radiation field around the BEL clouds. In obtaining the synthetic polarization, we neglect the dilution effect from the unpolarized continuum. If we normalize the line flux so that the equivalent width of the broad Lyα is 100 Å, then the composite degree of polarization of the broad Lyα flux would reduce to about 3 per cent.

5 Application to PG 1630+377

Spectropolarimetry played an important role in constructing the unified model of Seyfert galaxies and it continues to be important for other active galaxies (e.g. Miller & Goodrich 1990, Tran et al. 1999, Young et al. 1996). In type 2

Seyfert galaxies, the hidden broad emission lines often appear in the polarized flux, which is naturally explained by assuming that there is a scattering region consisting of free electrons or dust along the polar directions whereas the broad line region is hidden by an opaque component in the equatorial region. In this scattering geometry, the polarization direction is perpendicular to the symmetry axis of the accretion disk. No apparent polarization flip shown in the observed polarized spectra implies that the relative geometry of scatterers and incident continuum radiation source is polar rather than equatorial.

The degree of polarization of the Lyα scattered in the BELR and reflected from the accretion disk is similar to the observed value of PG 1630+377 obtained by Koratkar et al. (1995). We note that the continuum of PG 1630+377 is too weakly polarized to estimate the accurate position angle. It is noted that the polarization blueward of the Lyman edge is in the same direction as that of the Lyα. Considering the result that the scattered flux in a disk-like BELR is polarized only weakly in the direction parallel to the symmetry axis, the main contribution of the broad Lyα is plausibly attributed to the Rayleigh reflection from the accretion disk.

However, as is seen in the previous section, the polarization strength and direction are sensitively dependent on the anisotropy of the radiation field around the line emitting clouds. Therefore an independent determination of the accretion disk normal axis will provide important information on the anisotropy of the BELR radiation field and on the scattering geometry, which may strongly constrain the unification scheme of AGN.

As shown in the previous section, the outflowing BELR model gives polarization parallel to the disk normal in the red part of the Lyα feature, which results in the asymmetric polarized flux. Therefore a close inspection of the profile difference in the total flux and the polarized flux may reveal the kinematic information of the BELR. In the case of PG 1630+377 the peak of the Lyα polarized flux appears to be locatedredward compared with that of the total flux. However, as Koratkar et al. (1995) noted, there may be contribution from N V ions and higher resolution spectropolarimetry will shed more light.

Many theoretical suggestions were given to the surprising phenomenon of the steep increase of polarization blueward of the Lyman edge in PG 1630+377 invoking electron scattering in the accretion disk (e.g. Beloborodov & Poutanen 1999, Blaes & Agol 1996). These suggestions usually give strong polarization in the direction dependent on the scattering optical depth. However, as Koratkar et al. (1995) noted, it is also a possibility that the steep rise may represent the strongly polarized flux of a broad line feature in the far UV region. Singlet transition lines such as He I 584, O V 630 and N IV 765 are good candidates for this feature.

Acknowledgments

The author thanks Sang-Hyeon Ahn, Chang-Hwan Lee and R. D. Blandford for useful discussions.

References

1. Antonucci, R. 1993, Annual Review of Astronomy and Astrophysics, 31, 473
2. Beloborodov, A. M., and Poutanen, J., 1999, ApJ, 517, L77
3. Berestetskii, V. B., Lifshitz, E. M., and Pitaevskii, L. P. 1972, Relativistic Quantum Mechanics, Pergamon Press, New York
4. Blaes, O., and Agol, E., 1996, ApJ, 469, L41
5. Blandford, R. D., 1990, in Saas-Fee Advanced Course 20, Active Galactic Nuclei, ed. R. D. Blandford, H. Netzer, & L. Woltjer (Berlin: Springer)
6. Cole, G. H. J., Pedlar, A., Mundell, C. G., Gallimore, J. F., and Holloway, A. J., 1998, MNRAS, 301, 782
7. Dickey, J. M., 1988, ApJ, 300, 190
8. Ferland, G., 1996, in Hazy, Univ. of Kentucky Dept. of Phys. and Astron. Internal Rep.
9. Ferland, G., Peterson, B. M., Horne, K., Welsh, W. F., and Nahar, S. N., 1992, ApJ, 387, 95
10. Isliker, H., Nussbaumer, H., and Vogel, M., 1989, A&A, 219, 271
11. Korista, K., and Ferland, G., 1998, ApJ, 495, 672
12. Koratkar, A., Antonucci, R. R. J., Goodrich, R. W., Bushouse, H., and Kinney, A. L., 1995, ApJ, 450, 501
13. Koratkar, A., Antonucci, R. R. J., Goodrich, R. W., and Storrs, A. , 1998, ApJ, 503, 599
14. Laor, A., Netzer, H., Piran, T., 1990, MNRAS, 242, 560
15. Lee, H. -W., and Blandford, R. D., 1997, MNRAS, 288, 19
16. Lee, H.-W., and Lee, K. W., 1997, MNRAS, 287, 211
17. Lee, H. -W., and Yun, J. H., 1998, MNRAS, 301, 193
18. Miller, J. S., and Goodrich, R. W., 1990, ApJ, 355, 456
19. Murray, N., Chiang, J., Grossman, S. A., and Voit, G. M., 1995, ApJ, 451, 498
20. Netzer, H., 1990, in Saas-Fee Advanced Course 20, Active Galactic Nuclei, ed. R. D. Blandford, H. Netzer, & L. Woltjer (Berlin: Springer)
21. Rees, M. J., 1987, MNRAS, 228, P47
22. Sadeghpour, H. R., and Dalgarno, A., 1992, J. Phys. B: At. Mol. Opt. Phys., 25, 4801
23. Saslow, W. M., and Mills, D. L., 1969, Phys. Rev., 187, 1025

24. Tran, H. D., Brotherton, M. S., Stanford, S. A., Van Breugel, W., Dey, A., Stern, D., and Antonucci, R., 1999, ApJ, 516, 85
25. Young, S., Hough, J. H., Efstathiou, A., Wills, B. J., Bailey, J. A., Ward, M. J., Axon, D. J., 1996, MNRAS, 281, 1206

BLACK HOLE ACCRETION IN TRANSIENT X-RAY BINARIES

KRISTEN MENOU[a]
Princeton University, Dept. of Astrophysical Sciences, Princeton, NJ 08544, USA
kristen@astro.princeton.edu

Recent work on the modes of accretion onto black holes (BHs) in Soft X-Ray Transients (SXTs) is reviewed, with an emphasis on uncertainties affecting models of accretion during quiescence (inner hot flow, outer thin disk). Various interpretations of the quiescent X-ray luminosity difference between systems containing neutron stars (NSs) and systems containing BH candidates are also summarized. A new scenario, which does not require BH candidates to possess an event horizon, is presented here. This scenario may be ruled out in the future, from detailed X-ray spectroscopic diagnostics or from the absence of type I X-ray bursts in systems containing BH candidates.

1 Introduction

One of the main interests for studying accretion in close binary star systems is that we know better in these systems than anywhere else the general conditions under which accretion proceeds (e.g. binary size and geometry, donor and accretor masses, etc...). Many such close binary systems are transient: they semi-regularly experience large amplitude outbursts followed by long periods of quiescence. Depending on the nature of the compact object, these systems are called Dwarf Novae (DN; for a white dwarf accretor; see Warner 1995 for a review) or Soft X-ray Transients (SXTs; for a neutron star or black hole accretor; see Lewin, van Paradijs & van den Heuvel 1995 for a review).

Although this contribution is focused on the nature of accretion in BH SXTs (mostly during quiescence), reference will also be made to NS SXTs and DN. The similarities between these three classes of systems are strong, so that one can hope to learn about the compact object from their subtle differences. In §2, I review the motivation for a two-component accretion flow structure (inner hot flow, outer thin disk) in quiescent BH SXTs. In §3, I discuss various uncertainties affecting models of these two components of the accretion flow. In §4, I review the observational status and theoretical interpretations of the quiescent X-ray luminosity difference between systems containing BHs and those containing NSs. In §5, I present an alternative interpretation of this luminosity difference that does not require BH candidates to possess an event horizon.

[a] Chandra Fellow

2 Accretion Flow Structure

The gas accreted onto compact objects in close binary systems is transferred via Roche-lobe overflow by a main-sequence or sub-giant companion. The large specific angular momentum of the gas naturally leads to the conclusion that a thin accretion disk forms in the system (the gas being allowed to radiatively lose energy much more efficiently than momentum when cooling is efficient; see, e.g., Frank, King & Raine 1992 for a review). The presence of thin accretion disks during outburst in transient close binaries has been observationally confirmed by detailed eclipse maps in DN (Horne 1993) and X-ray spectroscopic diagnostics in SXTs (Tanaka & Shibazaki 1996). The situation in quiescence appears more complicated, however.

According to the Disk Instability Model (DIM), the outbursts of transient close binaries are triggered by a thermal-viscous instability occurring when hydrogen becomes partially ionized somewhere in the disk (Meyer & Meyer-Hofmeister 1981; Cannizzo 1993; Lasota 2001a). Lasota (1996) pointed out that if fully-extended, quiescent disks in SXTs were accurately described by the DIM, accretion onto the compact object should proceed at a rate $\ll 10^{10}$ g s^{-1}, i.e. many orders of magnitude smaller than inferred from X-ray observations. In addition, at such a low rate, a thin accretion disk is not expected to emit a substantial amount of X-rays (Narayan, McClintock & Yi 1996). This has led to the general belief that there cannot be a disk extending all the way down to the compact object in quiescent SXTs, as would be expected in the DIM. Wheeler (1996) also pointed out that the optical light from quiescent BH SXTs (excluding the contribution from the companion star; see Narayan et al. 1996) corresponds to black body temperatures in excess of 10^4 K, which is too hot for the neutral disk expected in the DIM. Nonetheless, the presence of a disk in quiescence appears required to explain the double-peaked emission lines observed (e.g. Orosz et al. 1994) and the outbursts experienced by SXTs.

An attractive solution to this problem is to postulate that the thin disk is present only in the outer regions of the accretion flow, while it is replaced by a hot, X-ray emitting flow in the vicinity of the BH. This possibility was first proposed by Narayan et al. (1996) and further developed by Esin, McClintock & Narayan 1997 (see also Narayan 1996; Lasota, Narayan & Yi 1996; Narayan, Barret & McClintock 1997; Hameury et al. 1997). These specific models, in which the inner hot flow is assumed to be an advection-dominated accretion flow (ADAF), have been rather successful at explaining the spectral properties of BH SXTs in various spectral states (Esin et al. 1998).

Strong observational support for the presence of a hot flow in the inner regions of BH SXTs during low-luminosity spectral states exists for the un-

usual BH SXT labeled XTE J1118+480 (see McClintock et al. 2001a for a mass function determination). This system is located high above the Galactic plane, so that for the first time EUV and soft X-ray spectral measurements could be obtained for a BH SXT (given the small extinction to the source; Garcia et al. 2000). These spectroscopic data were obtained as part of an extensive multiwavelength campaign during the low/hard-state outburst that led to the discovery of this system (Hynes et al. 2000; McClintock et al. 2001b and references therein). The low-state spectral energy distribution of XTE J1118+480 strongly suggests the presence of a thermal component peaking at $\sim$ 50 eV, in addition to the hard power law seen at higher energies (McClintock et al. 2001b). As explicitly shown by Esin et al. (2001), these data are best interpreted as indicating that the thin disk in this system (responsible for the thermal emission) is truncated at several tens of Schwarzschild radii from the BH, where it is replaced by a hot accretion flow (responsible for the power law emission observed). It is difficult to interpret the thermal emission observed as power-law emission from a corona that has been reprocessed by a fully extended, underlying disk (because the thermal component is apparently more energetic than the power-law component; Esin et al. 2001).

3 Uncertainties

Based on the successful spectral modeling of Esin et al. (1998) and the evidence in XTE J1118+480 (as well as other arguments discussed in §2), there are good reasons to believe that the inner regions of thin disks in quiescent BH SXTs are replaced by hot accretion flows. Many important components of this scenario are poorly constrained, however. A major uncertainty comes from our lack of understanding of the nature of the transition between the two components of the accretion flow. Although several mechanisms for the "evaporation" of the disk into a hot flow have been explored in the literature (Meyer & Meyer-Hofmeister 1994; Honma 1996; de Kool & Wickramasinghe 1999; Rozanska & Czerny 2000; Manmoto & Kato 2000; Spruit & Deufel 2001; Ball, Narayan & Quataert 2001), none of these theories provide reliable predictions that can be put to the observational test. In what follows, the discussion will be focused on additional uncertainties affecting models of the inner hot flow and the outer thin disk separately.

3.1 Inner Hot Flow

The work on hot accretion flows in recent years has focused on self-similar, analytical solutions to the problem of accretion onto a BH (but see also the work by Ichimaru 1977; Rees et al. 1982; Abramowicz et al. 1995).

Narayan & Yi (1994; 1995a,b) derived a self-similar, advection-dominated accretion flow (ADAF) solution in which nearly all the gravitational potential energy released during accretion is advected by the flow rather than efficiently radiated locally by the gas (as is the case for thin disk accretion). This advection property requires preferential viscous heating of the ions over the electrons and a two-temperature structure. The latter is allowed if energy transfer between the ion and electron populations is rather inefficient, as is the case if only Coulomb collisions operate. A detailed analysis of the energetics of turbulence and the various dissipation mechanisms involved in this type of flow suggests that preferential ion heating can be achieved only for significantly sub-equipartition magnetic fields (Gruzinov 1998; Quataert 1998; Quataert & Gruzinov 1999). As initially pointed out by Narayan & Yi, two complications with the ADAF solution are that the flow has a positive Bernoulli energy constant (indicating that the gas may be subject to hydrodynamical outflows because it is technically unbound) and is unstable to convection in the radial direction (while the effects of convection on the ADAF structure are neglected).

Blandford & Begelman (1999) explored further the possibility that the hot flow is subject to outflows. They constructed a family of advection-dominated inflow-outflow solutions (ADIOS) in which mass, energy and angular momentum are gradually lost to an outflow as part of the gas is being accreted. They suggested that solutions with negative Bernoulli energy constant, which can be obtained under strong outflow assumptions, are more viable than the initial ADAF solution (see Abramowicz, Lasota & Igumenshchev 2000 for a different view).

Alternatively, Narayan, Igumenshchev & Abramowicz (2000) and Quataert & Gruzinov (2000) explored the role of convection in determining the structure and properties of the hot flow. These authors proposed a convection-dominated accretion flow (CDAF) solution in which radial convection is postulated to dominate over viscous processes to determine the sign of angular momentum transport in the flow. The properties of CDAFs are quite different from those of ADAFs, in particular the presence of an additional convection-driven, outward energy flux that may be at the origin of an outflow launched from the outer regions of the hot flow. An interesting feature of the CDAF solution is that it does not require preferential ion heating like the ADAF solution does (Quataert & Narayan 1999; Ball et al. 2001).

While the temperature structure of ADAFs, ADIOS and CDAFs is close to virial in all three cases, their density structure vary widely: $\rho \propto R^{-3/2}$ in an ADAF, $\rho \propto R^{-1/2}$ in a CDAF and a ρ profile that depends on the wind parameters in an ADIOS. Given an accretion geometry and an accretion rate at which the truncated disk feeds the hot flow in a quiescent BH SXT, the

X-ray emission properties will therefore be largely different depending on the exact nature of the hot flow. The ADAF solution may not be favored because of the neglect of radial convection, the apparent susceptibility to hydrodynamical outflows and the possible difficulty to preferentially heat ions. It is unclear, however, whether the hot flow in quiescent BH SXTs should have an ADIOS-like structure, a CDAF-like structure or none of the above.

Given this uncertain situation, the interest for numerical investigations of the structure of radiatively inefficient hot flows has grown considerably during the last few years. A large number of 2D and 3D hydrodynamical simulations (Igumenshchev, Chen & Abramowicz 1996; Igumenshchev & Abramowicz 1999; 2000; Stone; Pringle & Begelman 1999; Igumenshchev, Abramowicz & Narayan 2000) have indicated that hot flows with a small value of the viscosity parameter α (< 0.1 or so) are subject to large-scale convective motions (in better agreement with CDAF theory), while hot flows with a large α value (> 0.1 or so) are subject to strong outflows (in better agreement with ADIOS theory). An obvious shortcoming of all these studies is that an ad-hoc Shakura-Sunyaev-type viscosity has to be included in the Navier-Stokes equations solved to mimic the "viscous" interactions that are thought to be in fact of magnetic origin in this hot plasma (Balbus & Hawley 1991; 1998)

Results of global 3D magneto-hydrodynamical (MHD) numerical simulations of hot flows have appeared only very recently because of the large numerical power required. Hawley, Balbus & Stone (2001) present such a calculation, with strong outflow properties. In addition, these authors argue that treating convection in the hot flow separately from angular momentum transport, as is done in CDAF analytical theories, is incorrect. On the other hand, Igumenshchev & Narayan (2001) present a 3D MHD numerical simulation of hot, spherical accretion (gas with zero net angular momentum) in which convective motions play an important role. The apparent discrepancy between these results may be caused by differences in the physical problems addressed or perhaps different numerical implementations (e.g. allowing or not for explicit magnetic energy dissipation in the equations solved), but it is clear at this point that no consensus has emerged on the structure and properties to expect for hot flows. Additional numerical simulations will likely resolve this issue in the future.

3.2 Outer Thin Disk

The structure and properties of outer thin disks in quiescent BH SXTs are also subject to considerable uncertainties, mostly because of our ignorance of the nature of viscosity in this case. These uncertainties are large enough that we

are currently unable to predict the rate at which the disk feeds the inner hot flow in these systems (given a mass transfer rate). Although the DIM makes definite predictions for the accretion rate as a function of radius to expect in a quiescent disk ($\dot{M} \propto R^{2.5}$ or so), the assumption that the bulk of the disk accretes mass with a uniform efficiency of angular momentum transport (expressed as a constant value α parameter in the DIM) may not be valid.

While the transport in ionized disks during outburst is probably due to MHD turbulence resulting from the non-linear development of the Magneto-Rotational Instability (MRI; Balbus & Hawley 1991; 1998), Gammie & Menou (1998) pointed out that quiescent disks are so neutral that resistive diffusion is important and MHD turbulence may not be sustained during this phase. Menou (2000) reaffirmed by using the MHD numerical simulations of resistive disks of Fleming, Stone & Hawley (2000) in more detailed calculations.

Recently, however, Wardle (1999) and Balbus & Terquem (2001) pointed out that the role of Hall terms had generally been ignored in studies of weakly-ionized disks. Although Hall terms are strongest in low-density environments (hence the discussion for T-Tauri disks), they are not negligibly small in the quiescent disks of close binary systems. Using the scaling derived by Balbus & Terquem (2001; their Eq. 25) and assuming a reasonable sub-equipartition magnetic field, one finds that the resistive diffusion term dominates over Hall effects for typical densities $\sim 10^{-6}$ g cm^{-3} found in quiescent disks. Nonetheless, Hall terms constitute an additional non-ideal MHD effect that will have to be included in future discussions of the level of MHD turbulence to expect in quiescent disks in transient close binaries.[b]

Menou (2000) argued that, in the absence of transport by MHD turbulence during quiescence, another "viscosity" mechanism could then drive accretion (perhaps spiral density waves induced in the disk via tidal interaction with the companion star; Spruit 1987). Even in the absence of an alternative viscosity mechanism operating during quiescence, it is still possible for accretion to proceed via MHD-turbulent, X-ray ionized layers at the disk surface. This layered-accretion scenario, first considered by Gammie (1996) for T-Tauri disks (ionized by cosmic rays), has been discussed in detail by Menou (2001) for quiescent disks in DN. The feasibility of layered accretion has been numerically demonstrated by Fleming & Stone (2002).

It is unclear, however, how relevant the layered accretion model is for quiescent disks in SXTs, because the origin and geometry of the ionizing X-rays are not well understood in these systems (for instance, the low-level of X-ray emission in quiescent BH SXTs may not allow the surface layers to be suffi-

[b]Balbus & Terquem note how difficult it is to analytically predict the influence of Hall terms because of their potentially stabilizing *and* destabilizing effects.

ciently ionized for MHD turbulence to operate). Nevertheless, the possibility that accretion proceeds via surface layers in quiescent DN rather than in the bulk of the disk as assumed in the DIM shows that our ignorance of the nature of viscosity in quiescent disks results in large uncertainties on the structure and properties of these disks. In particular, layered accretion predicts $\dot{M} \propto R$ in quiescent disks, as compared to $\dot{M} \propto R^{2.5}$ or so for the DIM (Menou 2001). Therefore, the rate at which the quiescent disk feeds the hot flow in BH SXTs is currently unpredictable. This situation is unlikely to change until we better understand the mechanism responsible for accretion in quiescent disks.

4 Black Hole – Neutron Star Luminosity Difference

One of the most interesting properties of BH SXTs is their X-ray faintness during quiescence, as compared to NS SXTs. The existence of this difference was first pointed out by Narayan, Garcia & McClintock (1997) and Garcia et al. (1997), when they compared systems experiencing type I X-ray bursts (hence containing NSs) to those with mass functions in excess of $2-3M_{\odot}$ (presumably containing BHs). Recent deep *Chandra* observations allowed the detection of several faint quiescent BH SXTs, thus confirming with better statistics that quiescent BH systems are about 2 orders of magnitude fainter than their NS equivalents (see Garcia et al. 2001 for the latest data). Since the discovery of this luminosity difference, various interpretations of the observational data have been proposed and debated, as described below.

4.1 Accretion Scenarios

Narayan et al. (1997) interpreted the data on quiescent SXTs as evidence for event horizons in systems containing BH candidates. Indeed, assuming that accretion proceeds via an ADAF in the inner regions of quiescent BH SXTs (§2), a large fraction of the gravitational potential energy released during the accretion process may be stored as thermal energy in the flow and lost through the BH event horizon (thus leading to a very small radiative efficiency). On the contrary, independent of the structure of the accretion flow, one would expect a large radiative efficiency ($\sim 10-20\%$) in quiescent NS SXTs, from the necessity to radiate all the energy released by accretion at the stellar surface. Assuming that accretion proceeds at roughly the same rate in the two classes of systems, one therefore expects quiescent BH SXTs to be less luminous than their NS equivalents.

Menou et al. (1999) attempted to quantify the luminosity difference expected in this accretion scenario with more detailed models. They pointed out the importance of comparing systems with similar orbital periods, to guarantee

comparable mass transfer rates and (presumably) mass accretion rates (allowing a meaningful test). They also showed that quiescent NS SXTs are much less luminous than naively expected, in that only a small fraction ($\sim 10^{-3}$) of the mass transferred must actually reach the NS surface. This could be achieved via the action of an efficient "propeller effect" in these systems (see also Asai et al. 1998). Menou & McClintock (2001) attempted to further test the ADAF+propeller scenario for quiescent NS SXTs, by using multiwavelength spectral data available for the specific system Cen X-4. This was not very successful in that the data do not show any signature of the presence of an ADAF in this system. Chandler & Rutledge (2000) further questioned the validity of the propeller picture by showing the near absence of X-ray pulsations when the NS SXT Aql X-1 fades into quiescence, while it is supposed to enter the propeller regime at that point.

Recently, Loeb, Narayan & Raymond (2001) and Abramowicz & Igumenshchev (2001) suggested that the observed X-ray luminosity difference of a factor ~ 100 would be better understood if one accounts for the modified hot flow structure expected from strong convection (CDAF theory). It is unclear how reliable these interpretations are, however, because of various oversimplifications (e.g. neglect of the stellar magnetic field, unproven relevance of CDAF theory for the case of accretion onto a compact object with a hard surface; see also discussion in Lasota 2001b). The current situation with accretion scenarios for quiescent BH and NS SXTs largely reflects the numerous uncertainties associated with hot flow models, as discussed in §3.

4.2 Stellar Scenarios

Bildsten & Rutledge (2000) proposed that the quiescent X-ray emission of BH SXTs is dominated by coronal emission from the rapidly-rotating companion star in these systems. This possibility has been challenged by Lasota (2000) and cannot explain the excessive X-ray emission in at least two know quiescent BH SXTs (GRO J0422+32 and V404 Cyg; see Bildsten & Rutledge 2000; Narayan, Garcia & McClintock 2001). X-ray spectral diagnostics also challenge this scenario in several additional systems (Kong et al. 2001).

Brown, Bildsten & Rutledge (1998) pointed out that the (soft) X-ray emission of quiescent NS SXTs could be powered by NS thermal cooling, following heating by deep crustal compression during outburst. The X-ray luminosities of $\sim 10^{32} - 10^{33}$ erg s^{-1} obtained by Brown et al. (1998), in agreement with the observed values, were later confirmed in more detailed calculations by Colpi et al. (2001; see also Ushomirsky & Rutledge 2001). The unusually low quiescent X-ray luminosity of the ms X-ray pulsar SAX J1808-36 (Wijnands

et al. 2001a) was expected in this scenario given the outburst properties of this system (Brown et al. 1998). The NS cooling interpretation is further supported by detailed NS hydrogen atmosphere fits to the X-ray data which require an emission region of size comparable to the entire NS surface, as expected (Rutledge et al. 1999; 2000; 2001a,b). These atmosphere models are successful in fitting the soft X-ray spectra of several sources recently discovered in globular clusters as well (Rutledge et al. 2001c; Grindlay et al. 2001). The crustal heating scenario may also be able to explain the case of the NS in the SXT KS 1731-260, despite the unusual outburst and cooling properties of this system (Wijnands et al. 2001b; Rutledge et al. 2001d).

Despite successes, the crustal heating scenario is not without difficulties. The power law component seen in the X-ray spectra of several quiescent NS SXTs (Asai et al. 1996; Campana et al. 1998, Rutledge et al. 2001a,b) is not explained by this scenario and therefore requires a different origin. It is puzzling that in both Aql X-1 and Cen X-4, the thermal and power law components have comparable luminosities, while they should be unrelated according to the crustal heating scenario. This coincidence may be more easily explained in an accretion scenario (which is not incompatible with hydrogen-atmosphere-type emission).

Variability properties of quiescent NS SXTs may also challenge the crustal heating model. Rutledge et al. (2001a) attributed the long-term variability of Cen X-4 during quiescence to a variation in the power-law component only. Campana et al. (1997) reported significant variability of the quiescent X-ray emission of Cen X-4 (over a period of several days), but it is unclear whether this variability can be simply attributed to the contribution of the power law component in the ROSAT-HRI soft X-ray band. Ushomirsky & Rutledge (2001) suggest that transient accretion events could lead to specific variability patterns of the NS thermal emission in the crustal heating scenario. Nonetheless, Rutledge et al. (2001b) acknowledge that the long-term variability of the thermal component that they observed in Aql X-1 during quiescence is not easily explained by the crustal heating scenario.

4.3 Other Scenarios

Other interpretations of the X-ray emission from quiescent SXTs have been proposed, including the radio-pulsar interpretation of Campana & Stella (2000) for NS SXTs. These alternatives are critically reviewed by Narayan et al. (2001).

5 Evidence for Event Horizons?

Whether the quiescent X-ray emission of SXTs is powered by accretion or NS cooling, the scenarios discussed in the previous section attribute, directly or indirectly, the faintness of quiescent BH SXTs to the lack of a hard surface in these systems.[c] The observed luminosity difference has therefore been interpreted as evidence for event horizons in systems containing BH candidates.

This argument is constructed by elimination: to establish the presence of event horizons in systems containing BH candidates, one should eliminate the possibility of an exotic compact object with a hard surface and a mass in excess of $2-3M_{\odot}$ (the usual general relativity limit for a NS, beyond which collapse to a BH is traditionally unavoidable; Rhoades & Ruffini 1974; Kalogera & Baym 1996). In this section, a specific scenario which makes this exotic assumption is considered in more detail (systems containing BH candidates are still referred to as BH SXTs, for simplicity).

Assuming that the quiescent X-ray emission of SXTs is powered by accretion,[d] the radiative efficiency should be comparable in BH and NS SXTs (say $\sim 10\%$), if hard surfaces are present in both cases, because the compact objects should have comparable compactness in first approximation (see below). Is it possible for accretion to proceed at a rate ~ 100 times smaller in BH SXTs during quiescence (only because of their ~ 5 times more massive compact objects) and thus explain the observed luminosity difference? As emphasized by Menou et al. (1999), the mass transfer rates in BH and NS SXTs should be comparable, as long as one compares systems with similar orbital periods. In addition, a close inspection of the scaling with central mass of the accretion rate in quiescent disks, in both the DIM and the layered accretion model of Menou (2001), reveals only a weak dependence on the mass of the central object, that is insufficient to explain the observed luminosity difference. One is therefore tempted to focus on differences in the properties of the flow in the vicinity of the compact object to explain the observed luminosity difference.

5.1 Accretion vs. Spindown Luminosity

It is natural to expect BH candidates, even with a hard surface, to be more compact than NSs, because of their larger masses. It is then possible for NSs radii to exceed the radius of the marginally stable orbit for an object

[c] As an exception, if the quiescent X-ray emission of NS SXTs is powered by NS cooling while coronal activity dominates in quiescent BH SXTs, the lack of hard surface in BH SXTs is not crucial (but see objections to the coronal emission scenario in §4.2).

[d] The specific scenario considered here does not exclude an additional contribution from NS cooling to the X-ray emission of quiescent NS SXTs.

of their mass, while BH candidates, being more compact, could lie within their marginally stable orbit (currently favored NS equations of state allow for NS radii both smaller and larger than the corresponding marginally stable orbit radius; e.g. Kalogera & Baym 1996). Based on this difference between BH candidates and NSs, an accretion scenario can be constructed in which the quiescent X-ray luminosity difference is explained without requiring BH candidates to have an event horizon.

For compact enough BH candidates, accretion will be transsonic and transalfvenic because the flow radial speed approaches the speed of light as the compact object radius approaches a Schwarzschild radius. In this case, the gas reaching the BH candidate surface is viscously disconnected from the rest of the flow and, independently of the flow structure at large radii, emission from the gas shocking the surface should have a radiative efficiency $\sim$ 10% or so. This implies very low accretion rates, $\sim 10^{10-11}$ g s^{-1}, in quiescent BH SXTs.

It is reasonable to expect similar mass accretion rates in BH and NS SXTs during quiescence (see point above). In the case of NS accretion, however, if the NS radius exceeds the marginally stable orbit, the entire flow is viscously connected. Medvedev & Narayan (2001) have shown that the flow can adopt a hot configuration in this case, through which it effectively spins down the NS. Medvedev & Narayan (2001) note that the spindown luminosity dominates over the accretion luminosity for low enough accretion rates. In particular, for accretion rates $\sim 10^{10-11}$ g s^{-1}, as discussed above, and NS spin frequencies $\sim$ 300 Hz (a reasonable value for NSs in SXTs; e.g., Wijnands & van der Klis 1998; White & Zhang 1997), a spindown luminosity of $\sim 10^{33}$ erg s^{-1} is expected, much in excess of the corresponding accretion luminosity. This spindown luminosity is right at the level required to explain the quiescent X-ray luminosity of NS SXTs (slight changes in model parameters, such the NS spin rate, can easily account for the range of luminosities observed in quiescent NS SXTs). Thus, in the accretion scenario considered here, quiescent BH SXTs are less luminous than their NS equivalents because, in their case, the flow is viscously disconnected from the compact object and only the accretion luminosity, at a level of $\sim 10^{30-31}$ erg s^{-1}, is liberated (independently of the BH candidate spin rate). Note that the hot settling flow solution of Medvedev & Narayan (2001) is strictly valid only for an unmagnetized NS. One can show that for surface field strengths roughly $\leq 10^8$ G (as is reasonable for NS in SXTs), an unmagnetized approximation is relevant because the thermal pressure in the hot settling flow dominates over the magnetic field pressure at the stellar surface.

The accretion scenario outlined above is dynamically-consistent with recent work on hot accretion flows and energetically-consistent with the observed

X-ray luminosity difference. It is arbitrary only in that it postulates the existence of massive, very compact objects and it assumes that NSs lie beyond their marginally stable orbits. For simplicity, the issue of thermal emission from the massive compact object in BH SXTs and the role of magnetic fields associated with this object were ignored (the fields could arguably be weak and the object internal structure could be such that crustal heating like in the NS case is not expected).

The value of this scenario may not be as a solid alternative to other interpretations of the observed luminosity difference. Rather, it shows that it could be dangerous to have evidence for BH event horizons rely only on the luminosity difference between BH and NS systems. By using additional observables, however, the evidence will likely be made stronger. Detailed X-ray spectroscopic diagnostics may, for instance, rule out the presence of a hot settling flow in quiescent NS SXTs. Perhaps more importantly, accumulation of mass at the hypothetical surface of the compact object in BH SXTs during outburst may be expected to trigger type I X-ray bursts, by analogy with the NS case (and independently of the detailed internal structure of the hypothetical compact object). The non-detection of these bursts may therefore constitute the strongest evidence for event horizons in BH SXTs. Detailed burst calculations by Heyl & Narayan (2002) for a compact object of arbitrary size support the validity of this test.

Acknowledgments

The author thanks Lars Bildsten, Jean-Pierre Lasota, Jeff McClintock, Ramesh Narayan and Eliot Quataert for comments on the manuscript, and the Center for Astrophysical Sciences at Johns Hopkins University for hospitality. Support for this work was provided by NASA through Chandra Fellowship grant PF9-10006 awarded by the Smithsonian Astrophysical Observatory for NASA under contract NAS8-39073.

References

1. Abramowicz, M.A. et al. 1995, ApJ, 438, L37
2. Abramowicz, M.A. & Igumenshchev, I.V. 2001, ApJ, 554, L53
3. Abramowicz, M.A., Lasota, J.-P. & Igumenshchev, I.V. 2000, MNRAS, 314, 775
4. Asai, K. et al. 1996, PASJ, 48, 257
5. Asai, K. et al. 1998, PASJ, 50, 611
6. Balbus, S.A., & Hawley, J.F. 1991, ApJ, 376, 214
7. Balbus, S.A., & Hawley, J.F. 1998, Rev. Mod. Phys., 70, 1

8. Balbus, S.A. & Terquem, C. 2001, ApJ, 552, 235
9. Ball, G., Narayan, R. & Quataert, E. 2001, 552, 221
10. Blandford, R.D. & Begelman, M.C. 1999, MNRAS, 303, L1
11. Brown, E.F., Bildsten, L. & Rutledge, R.E. 1998, ApJ, 504, L95
12. Bildsten, L. & Rutledge, R.E. 2000, ApJ, 541, 908
13. Campana, S. et al. 1998, ApJ, 499, L65
14. Campana, S., Mereghetti, S., Stella, L. & Colpi, M. 1997, A&A, 324, 941
15. Campana, S. & Stella, L. 2000, ApJ, 541, 849
16. Cannizzo, J.K. 1993, in Wheeler J.C., ed., Accretion discs in Compact Stellar Systems (World Scientific, Singapore), p. 6
17. Chandler, A.M. & Rutledge, R.E. 2000, ApJ, 545, 1000
18. Colpi, M., Geppert, U., Page, Dany & Possenti, A. 2001, ApJ, 548, L175
19. de Kool, M. & Wickramasinghe, D. 1999, MNRAS, 307, 449
20. Esin, A.A. et al. 1998, ApJ, 505, 854
21. Esin, A.A. et al. 2001, ApJ, 555, 483
22. Esin, A.A., McClintock, J.E. & Narayan, R. 1997, ApJ, 489, 865
23. Fleming, T.P. & Stone, J.M. 2002, in preparation
24. Fleming, T.P., Stone, J.M. & Hawley, J.F., 2000, ApJ, 530, 464
25. Frank, J., King, A.R. & Raine, D.J. 1992, *Accretion Power in Astrophysics* (Cambridge University Press, Cambridge)
26. Gammie, C.F. 1996, ApJ, 457, 355
27. Gammie, C.F. & Menou, K. 1998, ApJ, 492, L75
28. Garcia, M.R. et al. 2000, IAU Circ. 7392
29. Garcia, M.R. et al. 2001, ApJ, 553, L47
30. Garcia, M.R., McClintock, J.E., Narayan, R. & Callanan, P.J. 1997, Proceedings of the 13th North American Workshop on CVs, eds. S. Howell, E. Kuulers, C. Woodward (San Francisco: ASP), p. 506
31. Grindlay, J.E., Heinke, C., Edmonds, P.D. & Murray, S.S. 2001, Science, 292, 2290
32. Gruzinov, A. 1998, ApJ, 501, 787
33. Hameury J.-M., Lasota J.-P., McClintock, J.E. & Narayan, R. 1997, ApJ, 489, 234
34. Hameury J.-M., Menou K., Dubus G., Lasota J.-P., Huré J.-M., 1998, MNRAS, 298, 1048
35. Hawley, J.F., Balbus, S.A. & Stone, J.M. 2001, ApJ, 554, L49
36. Hawley, J.F., Balbus, S.A. & Winters, W.F. 1999, ApJ, 518, 394
37. Hawley, J.F. & Stone, J. 1998, ApJ, 501, 758
38. Heyl, J.S. & Narayan, R. 2002, in preparation
39. Honma, F. 1996, PASJ, 48, 77
40. Horne, K. 1993, in "Accretion Disks in Compact Stellar Systems", ed.

J.C. Wheeler (Singapore: World Scientific), 117
41. Hynes, R. I. et al. 2000, ApJ, 539, L37
42. Ichimaru, S. 1977, ApJ, 214, 840
43. Igumenshchev, I.V. & Abramowicz, M.A. 1999, MNRAS, 303, 309
44. Igumenshchev, I.V. & Abramowicz, M.A. 2000, ApJS, 130, 463
45. Igumenshchev, I.V., Abramowicz, M.A. & Narayan, R. 2000, ApJ, 537, L271
46. Igumenshchev, I.V., Chen, X. & Abramowicz, M.A. 1996, MNRAS, 278, 236
47. Igumenshchev, I.V. & Narayan, R. 2001, ApJ, submitted, astro-ph/0105365
48. Kalogera, V. & Baym, G. 1996, ApJ, 470, L61
49. Kong, A.K.H., McClintock, J.E. Garcia, M.R., Murray S.S. & Barret, D. 2001, ApJ, submitted, astroph/0111134
50. Lasota, J.-P. 1996, IAU Colloquium 163: "Accretion phenomena and Related Outflows", eds. D. Wickramasinghe, L. Ferrario and G. Bicknell, ASP Conf. Series, astroph/9610068
51. Lasota, J.-P. 2000, A&A, 360, 575
52. Lasota, J.-P. 2001a, New A.R., 45, 449
53. Lasota, J.-P. 2001b, EDPS Conference Series in Astronomy & Astrophysics, in press, astroph/0110212
54. Lasota, J.-P., Narayan, R. & Yi, I. 1996, A&A, 314, 813
55. Lewin, W.H.G., van Paradijs, J. & van den Heuvel, E.P.J. 1995, *X-ray Binaries* (Cambridge: Cambridge University Press)
56. Loeb, A., Narayan, R. & Raymond, J.C. 2001, ApJ, 547, L151
57. Manmoto, T. & Kato, S. 2000, ApJ, 538, 295
58. McClintock, J. E. et al. 2001a, ApJ, 551, L147
59. McClintock, J. E. et al. 2001b, ApJ, 555, 477
60. Medvedev, M.V. & Narayan, R. 2001, ApJ, 554, 1255
61. Menou, K. 2000, Science, 288, 2022
62. Menou, K. 2002, Proceedings of "The Physics of Cataclysmic Variables and Related Objects" (ASP Conference Ser.), eds. B. Gaensicke, K. Beuermann & K. Reinsch, astro-ph/0108287
63. Menou, K. et al. 1999, ApJ, 520, 276
64. Menou, K. & McClintock, J.E. 2001, ApJ, 557, 304
65. Meyer, F. & Meyer-Hofmeister, E. 1981, A&A, 104, L10
66. Meyer, F. & Meyer-Hofmeister, E. 1994, A&A, 288, 175
67. Narayan, R. 1996, ApJ, 462, 136
68. Narayan, R., Barret, D. & McClintock, J.E. 1997, ApJ, 482, 448
69. Narayan, R., Garcia, M.R. & McClintock, J.E. 1997, ApJ, 478, L79

70. Narayan, R., Garcia, M.R. & McClintock, J.E. 2001, in Proc. IX Marcel Grossmann Meeting, eds. V. Gurzadyan, R. Jantzen and R. Ruffini (Singapore: World Scientific), astroph/0107387
71. Narayan, R., Igumenshchev, I.V. & Abramowicz, M.A. 2000, ApJ, 539, 798
72. Narayan, R., McClintock, J.E. & Yi, I. 1996, ApJ, 457, 821
73. Narayan, R. & Yi, I. 1994, ApJ, 428, L13
74. Narayan, R. & Yi, I. 1995a, ApJ, 444, 231
75. Narayan, R. & Yi, I. 1995b, ApJ, 452, 710
76. Orosz, J.A., Bailyn, C.D., Remillard, R.A., McClintock, J.E. & Foltz, C.B. 1994, ApJ, 436, 848
77. Quataert, E. 1998, ApJ, 500, 978
78. Quataert, E. & Gruzinov, A. 1999, ApJ, 520, 248
79. Quataert, E. & Gruzinov, A. 2000, ApJ, 539, 809
80. Rees, M.J., Phinney, E.S., Begelman, M.C. & Blandford, R.D. 1982, Nature, 295, 17
81. Rhoades, C.E. & Ruffini, R. 1974, Phys. Rev. Lett., 32, 6
82. Rozanska, A. & Czerny, B. 2000, 360, 1170
83. Rutledge, R.E. et al. 1999, ApJ, 514, 945
84. Rutledge, R.E. et al. 2000, ApJ, 529, 985
85. Rutledge, R.E. et al. 2001a, ApJ, 551, 921
86. Rutledge, R.E. et al. 2001b, ApJ, 559, 1054
87. Rutledge, R.E. et al. 2001c, ApJ, submitted, astroph/0105405
88. Rutledge, R.E. et al. 2001d, ApJ, submitted, astroph/0108125
89. Spruit, H.C. 1987, A&A, 184, 173
90. Spruit, H.C. & Deufel, B. 2001, A&A, submitted, astroph/0108497
91. Stone, J.M., Pringle, J.E. & Begelman, M.C. 1999, MNRAS, 310, 1002
92. Tanaka, Y. & Shibazaki, N. 1996, ARA&A, 34, 607
93. Ushomirsky, G. & Rutledge, R.E. 2001, MNRAS, 325, 1157
94. Wardle, M. 1999, MNRAS, 307, 849
95. Warner, B. 1995, *Cataclysmic variable stars* (Cambridge University Press, Cambridge)
96. Wheeler, J.C. 1996, in "Relativistic Astrophysics: A Conference in Honor of Igor Novikov's 60th Birthday", eds. B. Jones & D. Markovic (Cambridge Univ. Press), astroph/9606119
97. White, N.E. & Zhang, W. 1997, ApJ, 490, L87
98. Wijnands, R. et al. 2001a, ApJ, submitted, astroph/0105421
99. Wijnands, R. et al. 2001b, ApJL, in press, astroph/0107380
100. Wijnands, R. & van der Klis, M. 1998, Nature, 394, 344

X-RAY VARIABILITY OF GALACTIC BLACK HOLES AND SIMULATED MAGNETOHYDRODYNAMICAL FLOW

S. MINESHIGE
Yukawa Institute, Kyoto University, Sakyo-ku, Kyoto 606–8502, Japan
E-mail: minesige@yukawa.kyoto-u.ac.jp

H. NEGORO
RIKEN, Hirosawa, Wako, Saitama 351–0198, Japan
E-mail: negoro@crab.riken.go.jp

R. MATSUMOTO, M. MACHIDA, T. MANMOTO
Dept. of Physics, Chiba University, Inage-ku, Chiba 263–8522, Japan
E-mail: matumoto@astro.c.chiba-u.ac.jp, machida@astro.c.chiba-u.ac.jp, manmoto@astro.c.chiba-u.ac.jp

X-ray time variations of Cyg X–1 in the hard state comprise two components with different timescales; one characterizing the light variation and the other rapid spectral hardening around the peak. These seemingly indicate two (or more) physical processes being involved; e.g., disturbance propagation and magnetic reconnection, leading to magnetic flares. Such features are consistent with the view of MHD accretion flow. The unique properties of the MHD accretion flow are also discussed. To explain the observed X-ray spectral shape and variability during the hard state, magnetic energy needs to be dissipated and be radiated away in short timescales. Finally, we will touch on the peculiar variations of GRS 1915+105, which could be the first evidence of relaxation oscillations between the standard and slim disks.

1 Observed X-ray spectral variability

As is well known, complex time variability is one of the most prominent observational features of black-hole (BH) objects. The power-spectral densities (PSDs) commonly show $1/f$-like declines at high frequencies. Such variability properties should contain the important physics involved with accretion flow into BHs, and, hence, many studies have been performed to quantify and interpret these variability patterns in terms of the PSD shapes. Nevertheless, its origin is not yet clear until now (although magnetic flares are most promising at the present; see Wheeler 1977; Galeev, Rosner, & Vaiana 1979; Takahara 1979; Kawaguchi & Mineshige 1999). We wish to stress, here, that the Fourier-domain analyses have limitations, since important pieces of information are lost in the procedure. In fact, many fluctuating light curves, which are not alike the observed ones, can reproduce the power-law declines in PSDs as are observed. We should return to the time-domain analyses to get a good insight into physics underlying spectral variability.

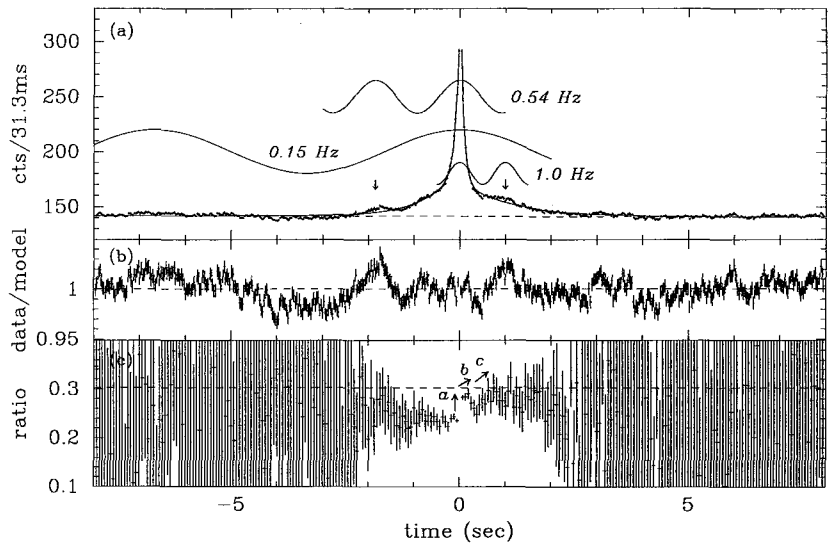


Figure 1: The variability properties of the superposed shot obtained based on the *Ginga* data of Cyg X–1. (a) Average time profile of the shots. (b) The ratio of the observed counts to that expected by the best-fit double-exponential model. (c) The hardness ratio variations. The dashed line is the average ratio. (From Negoro, Kitamoto, & Mineshige 2001)

There was a strong reason why the Fourier-domain analyses were so popular. The number of photons detected per certain periods is not enough to perform time-domain analyses with good statistics, while we can sum up many photons when performing the Fourier-domain analyses. To overcome this difficulty, the Osaka group of people developed a useful technique to be used in time-domain analyses; that is the technique of the superposed shot. If we have a close look at the X-ray intensity variations of Cyg X–1, we notice rather spiky features (referred to as shots) which are present on the timescale of several sec. By choosing only large events and superposed them by aligning their peaks, we can obtain a sort of the average variation patterns. The results are summarized in figure 1.

The upper panel shows an average time profile of the shots of Cyg X-1 in the 1.2–58.4 keV band obtained by superposing 872 large shots (with peak intensity of exceeding twice its average). Surprisingly, the shot profile is rather symmetric in time with respect to the peak epoch. The profile is well represented by double exponential models; i.e., the linear combination of the two exponential functions with different time constants; one about 0.1 s and

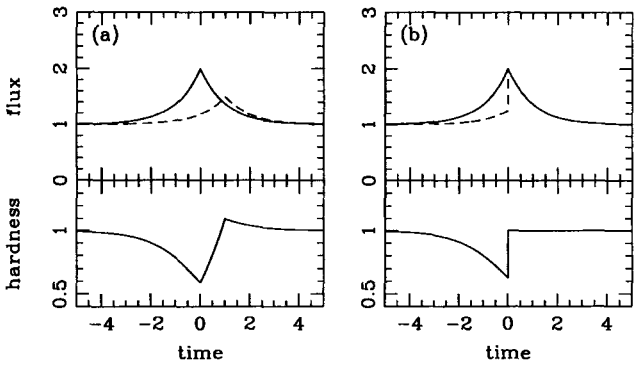


Figure 2: Schematic light curves (upper) and hardness ratios (lower) of the two models for hard time lag: (a) a delayed hard X-ray model, and (b) an abrupt hardening model. The solid and dashed lines in the upper panels represent soft and hard fluxes. It is easy to produce case (a) profile with physical models, but the observations point the case (b). (From Negoro, Kitamoto, & Mineshige 2001)

the other about 1 sec. The deviations from the best-fit double exponential model are illustrated in the middle panel of figure 1. Even more interesting is the hardness ratio (7.3–14.6 keV/1.2–7.3 keV) variation of the excess shot (shot component subtracted by the mean flux), which is depicted in the bottom panel of figure 1. It clearly exhibits a soft spectrum before the peak intensity, and rapid hardening in $\ll 0.1$ s around the peak, followed by complex changes.

An important discovery is that our current analysis points the presence of two (or more) physical timescales involved with X-ray variability: $\tau \simeq 0.1-1$ s, on which X-ray intensity rises and decays, and $\tau < 10$ ms, on which a spectral hardening occurs around the shot peak. We should caution that *the hard X-ray peak does not significantly lag behind the soft X-ray one* even within a shorter time resolution of ~ 2 ms of the RXTE (Fen, Li, & Chen 1999). It is very difficult to explain both timescales simultaneously in terms of a single physical mechanism, since usually one physical mechanism has one typical timescale. Figure 2 schematically illustrates the distinction between the variation patterns which usual models predict and those obtained by our data analyses. We can thus reject the hot spot model and Compton cloud model, because they predict rather long hardening timescale comparable to the shot timescales on a few sec. Rather, we suggest that the relatively long shot timescales could be related to accretion, while the short one may be due to local heating, such as magnetic reconnection. Such interpretations are consistent with the view of MHD accretion flow to be discussed in the next section.

2 Diversity of flows in adiabatic regimes

2.1 ADAF, ADIOS, and CDAF

Next, we turn the subject to the issue of hot accretion flow structure. Recently, multi-dimensional simulations have been intensively performed and a new paradigm of CDAF (convection-dominated accretion flow) has emerged in the adiabatic accretion regimes, in addition to the 'classical' ADAF (Ichimaru 1977) models (see Narayan 2001 for a review). The original optically thin ADAF model has been constructed basically in a vertically one-zone approximation. In other words, it was formulated in one (radial) dimension. Recently made 2D/3D hydrodynamical simulations, however, revealed distinct forms of accretion flows in the adiabatic (cooling-inefficient) regimes. Igumenshchev & Abramowicz (2000), for example, claimed that nearly one-dimensional ADAF appears only when the viscosity parameter α is moderate, $0.01 < \alpha \leq 0.1$. If α is very small, convection occurs and largely modify the flow structure (see table 1). Such flow is called as convection-dominated accretion flow (CDAF). If α is large ($\alpha \sim 1$), on the other hand, strong outflows result. Blandford & Begelman (1999) constructed a model of ADIOS (advection dominated inflows and outflows) based on this idea.

The critical test of these flows are to examine their spectral properties. Figure 3 displays the typical ADAF (left) and CDAF (right) spectra calculated in a similar way to that of Ball, Narayan, & Quataert (2001; see also Narayan & Yi 1995; Manmoto, Mineshige, & Kusunose 1997). The big distinction lies in their density profiles (see table 1): $\rho = \rho_0 (r/r_{\rm S})^{-p}$ with $p = 1.5$ in ADAF (with $r_{\rm S}$ being the Schwarzschild radius) and $p = 0.5$ in CDAF, respectively. The proportionality constant, ρ_0, was chosen to reproduce the given mass-accretion rate (see below), $\dot{M}$, for the case of ADAF. We took the same ρ_0 in CDAF, as well, to make comparison easier. Hence, figure 3 can be used only for demonstration purpose. The other adopted parameters are the same for both models: the inner-edge radius is 1.0 $r_{\rm S}$ the outer-edge radius is $10^{2.5}$ $r_{\rm S}$, the mass-flow rate is $\log \dot{m} \equiv \log(\dot{M}c^2/L_{\rm E}) = -3.29$ (with $L_{\rm E}$ being

Table 1: ADAF, CDAF, and MDAF

accretion mode	$\rho(r)$	$T(r)$	$v_r(r)$	$B^2(r)$
ADAF	$\propto r^{-3/2}$	$\propto r^{-1}$	$\propto r^{-1/2}$	...
outflow	$\propto r^{-1}$	$\propto r^{-1}$	$\propto r^{-1}$	...
CDAF	$\propto r^{-1/2}$	$\propto r^{-1}$	$\propto r^{-3/2}$	...
MDAF	$\propto r^{-0.5}$	$\propto r^{-1.0}$	$\propto r^{-1.5}$	$\propto r^{-1.5}$

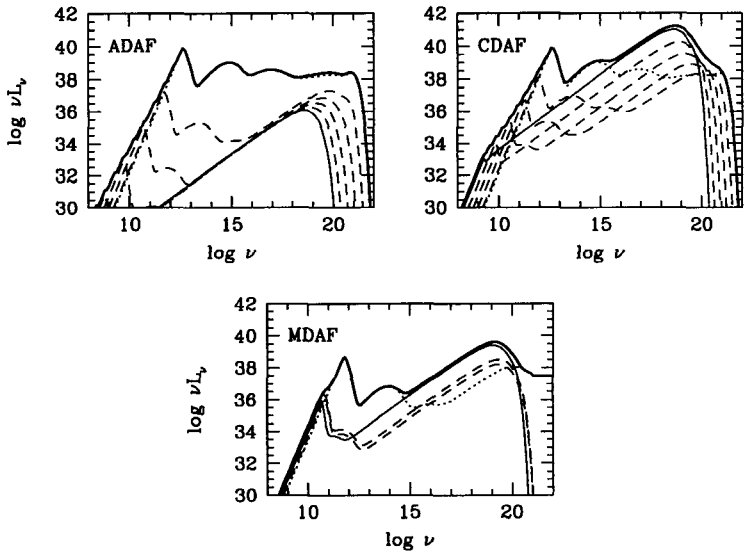


Figure 3: Typical ADAF (upper left) and CDAF (upper right) spectra, together with that of MHD flow (lower). For model parameters, see text. The thick solid line in each panel represents the total spectrum, while thin solid, dashed, and dotted lines represent the contributions from the outermost, middle, and innermost rings, respectively.

the Eddington luminosity), the mass of black hole is $m \equiv M_{\rm BH}/M_\odot = 10^8$, the ion temperature profile is $10^{12}(r/r_{\rm S})^{-1.0}$ K, the electron temperature profile is $10^{10}(r/r_{\rm S})^{-0.6}$ K, and magnetic field strengths are taken to be the equipartition values; i.e., $B^2 \propto (r/r_{\rm S})^{-(1+p)}$.

In the case of ADAF, emission from the innermost ring dominates over the contribution from the outer parts at all wavelengths because of its steep density profiles. Note that the density profile of $\rho \propto r^{-3/2}$ leads to the bremsstrahlung emissivity of $dE \propto \rho^2 T^{1/2} r^2 dr \propto r^{-1.5} dr$. In the case of CDAF, in contrast, X-ray emission comes from the outer parts due to less steep density profile ($\rho \propto r^{-1/2}$ yields $dE \propto r^{0.5} dr$), while radio to optical emissions are from the innermost part. The different natures of ADAF and CDAF are well demonstrated in figure 3. We also understand that bremsstrahlung is a dominant process in CDAF in X-rays, while Compton up-scattering of synchrotron emission is more important in ADAF. These distinctive features should be kept in mind when we compare with the observation.

2.2 MHD accretion flow

In the previous subsection, we treat the magnitude of the viscosity and magnetic field strengths to be free parameters. However, they cannot be a free parameter but should be determined in accordance with magnetic field amplifications within the disk. In other words, we need to perform MHD simulations. What do then the radiation spectra of MHD flow look like? Does MHD flow look more like ADAF or CDAF?

Multi-dimensional MHD simulations of accretion flow has also been extensively performed very recently, initiated by Machida, Hayashi, & Matsumoto (2000; see also Hawley & Krolik 2001; Hawley, Balbus, & Stone 2001). Machida, Matsumoto, & Mineshige (2001) have analyzed the flow patterns within the 3D MHD flow based on the simulation data by Machida, Hayashi, & Matsumoto (2000), finding large-scale convective motions dominating in accretion flows. Further, they also noted similar density and temperature profiles of MHD flow to those of CDAF (see table 2). Therefore, the spectra should be more like that of CDAF, which is exactly what we obtained (see figure 3).

However, one serious problem arises; namely, CDAF-like MHD flow models cannot account for the following observational facts of Cyg X–1 during its hard state:

1. CDAF predicts rather flat ($f_\nu \propto \nu^0$) spectrum, unless we assume significant electron heating as well as ion heating (Ball, Narayan, & Quataert 2001), but the observations clearly show a power-law decline (e.g. $f_\nu \propto \nu^{-0.7}$) in X-ray ranges.
2. Rapid variability seems to arise via time-dependent processes probably associated with magnetic flares in the innermost region (see §1), but in CDAF X-ray emission from the inner part is negligible (see figure 3).

In view of these arguments, ADAF looks more preferable, since then intermittent magnetic reconnection occurring in the innermost part produces seed synchrotron photons, which are to be Compton up-scattered to generate substantial X-ray emission in highly variable way.

Then, some of fundamental assumptions made in MHD simulations need to be re-considered. what is a loophole? The key to resolving this issue might be whether the dissipated magnetic energy by reconnection goes directly to plasma or radiation. So far, all the MHD simulations postulate no radiative cooling (since we are now concerned with the adiabatic accretion regimes), but if dissipated magnetic energy can be radiated away by enhanced emissivity as a consequence of magnetic reconnection, MHD flow structure should be more like that of ADAF, instead of CDAF, since entropy increase towards the BH is suppressed to some degree. This will explain both of the spectral shape and the presence of X-ray variability.

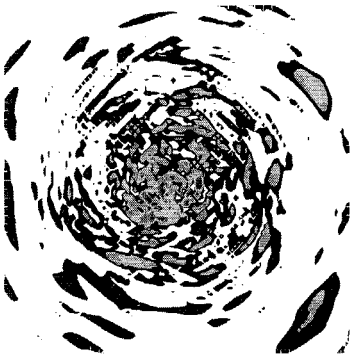


Figure 4: Contours of j/ρ (current density/matter density) values on the equatorial plane of the simulated MHD flow. Note that this quantity is proportional to electron drift velocity, and hence is a good indicator for occurrence of magnetic reconnection. (From Kawaguchi *et al.* 2000)

An alternative possibility is the effects of jet outflows. If we start the MHD calculations with poloidal magnetic fields, strong outflow (or jet) results (Kudoh et al. 2001?; Hawley, Balbus, & Stone 2001). Then, the enhanced energy production can be carried away to the infinity by jet material. Obviously, we need more studies to settle down the present issue.

2.3 Role of magnetic fields

Magnetic fields play many important roles in accretion disks, such as

- source of viscosity,
- disk corona (and ADAF) heating,
- cause of flares, producing variability,
- source of radiation via synchrotron,
- jet & outflow formation.

These processes have been successfully simulated through global MHD simulations of the inner accretion disk by several groups of researchers (see above). They have found that magnetic fields can be amplified sufficiently enough to explain the observations; the estimated viscosity parameters are in a range of $0.01 < \alpha < 1.0$. Large corona can be constructed by MHD processes. Also, it is possible to produce aperiodic variability with power-law decline in PSDs (Kawaguchi *et al.* 2000; Hawley & Krolik 2001). Figure 4 demonstrates spatial inhomogeneity pattern in magnetic-field distribution on the equatorial plane of MHD accretion flow. By performing percolation analysis, we have found

a beautiful power-law relation between the number of 'clusters' (those parts which possess large j/ρ values) and their volumes. Since j/ρ value is a good indicator for the preferential occurrence of magnetic reconnection, the natural consequence is production of a variety of flare amplitudes and durations. These were actually observed in Cygnus X–1 (Negoro *et al.* 1995). Our conclusion is that temporal aperiodic fluctuations could be closely related to spatial fractal structure created by magnetic fields (Mineshige, Takeuchi, & Nishimori 1994), as has been claimed by the concept of self-organized criticality (Bak, Tang, & Wiesenfeld 1987).

It is of great importance to stress that the roles of magnetic fields are more significant in ADAF/corona, than in standard-type disks, for the following reasons. In the standard-type disks, gravitational energy far exceeds internal energy of the disk gas because of efficient radiative cooling. Magnetic field energy is, on the other hand, suppressed below gas energy (but not by many orders below), since otherwise magnetic fields will leave the system in forms of bubbles. We thus finally obtain an equality,

$$E_{\rm mag} \leq E_{\rm gas} \ll E_{\rm grav} \sim E_{\rm rad}. \tag{1}$$

This situation is just like the photosphere of the Sun emitting optical flux. Therefore, magnetic fields are not dynamically important.

In ADAF/corona, in contrast, gravitational energy turns into fluid energy with little being radiated away. Since again magnetic field growth is regulated below a certain limit, we have an equality,

$$E_{\rm mag} \leq E_{\rm gas} \sim E_{\rm grav} \gg E_{\rm rad}. \tag{2}$$

The situation is like the solar coronae emitting X-rays. In fact, X-ray images of the Sun show violent activities on the solar atmosphere activated by magnetic fields. This is the main reason why magnetic fields are likely to produce observable effects in ADAF/coronae. It should be emphasized that all the MHD simulations so far made deal with this sort of hot flow and do not simulate the standard-type cool disk, in which radiative cooling is critical.

3 Relativistic Slim Disks

Finally, we touch on the recent interesting issue involved with intermediate-massive black holes (IMBHs) in ultra-luminous compact X-ray sources (ULXs) successively found in the arms of nearby spiral galaxies. Some of these seem to exhibit large luminosities, comparable to the Eddington luminosity. Then the theory predicts that photon trapping is substantial. Such a flow is well modeled

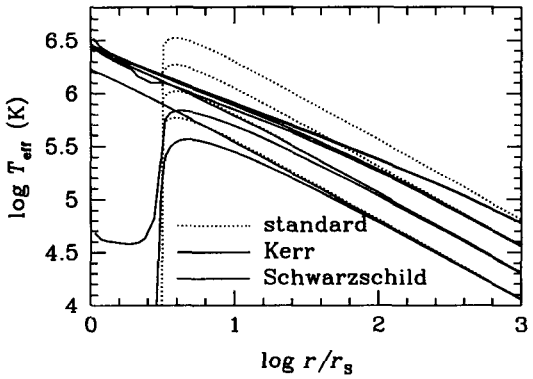


Figure 5: Effective temperature distributions of relativistic slim disks around a Schwarzschild hole (by blue lines), that around a Kerr hole with a spin parameter of $a = 0.998$ (by red lines), and that of a standard disk (by the blue lines), respectively. (From Manmoto & Mineshige 2001)

by the slim disk (Abramowicz *et al.* 1988). Now, we have good candidates to check the slim-disk theory.

We have constructed a relativistic slim disk model, in which fully relativistic treatments, including gravitational lensing, are considered (Manmoto & Mineshige 2001). Let us summarize the striking results obtained. There are three major points:

1. Radial temperature profiles become flatter as L increases, from $T \propto r^{-3/4}$ in the standard regime to $T \propto r^{-1/2}$ in the slim-disk regime.
2. For the disks around Schwarzschild (non-rotating) BHs the apparent inner-edge radius of the disk *decreases* from $3r_{\rm S}$ to $\sim 2\ r_{\rm S}$, as luminosity goes up to $L \sim L_{\rm E}$. Such trends are not clear when BHs are rapidly rotating, since the radius of the marginally stable last circular orbit is smaller than non-rotating BHs.
3. At $L \sim L_{\rm E}$, the disks around Schwarzschild and Kerr BHs look quite similar.

We show in figure 5 the radial temperature profiles of the optically thick disk for several accretion rates. Temperature profiles are flatter when advection dominates over radiative cooling, since then a part of energy (corresponding to the difference between the slim disk and the standard disk cases) is advective inward to the black hole. At low $\dot{M}$, there is a clear distinction between the Schwarzschild hole and Kerr hole cases, since the inner-edge radius is $3\ r_{\rm S}$ for the former, while it is $\sim r_{\rm S}$ for the latter. At high $\dot{M}$, however, the distinction

Table 2: Standard disk and slim disk ($m_1 \equiv M_{\rm BH}/10\ M_\odot$)

disk model	standard disk	slim disk
disk inner edge ($r_{\rm in}$)	$3\ r_{\rm S}$	$\leq 2\ r_{\rm S}$
maximum temperature ($T_{\rm in}$)	$\sim 1\ m_1^{-1/4}$ keV	$\sim 2\ m_1^{-1/4}$ keV
temperature profile	$\propto r^{-3/4}$	$\propto r^{-1/2}$

practically disappears, since substantial amount of radiation is expected to arise from inside $3r_{\rm S}$ even for the case with a Schwarzschild BH. This provides a marked difference between the low and high L accretion flows.

Accordingly, accretion disk spectra differ from those expected by the simple standard-disk theory. To make direct comparison with the observations, we performed spectral fitting to the theoretical spectra based on the multi-color disk model (MCD model; Mitsuda *et al.* 1984). The free parameters are the (apparent) inner-edge radius of the disk ($r_{\rm in}$) and the maximum temperature ($T_{\rm in}$). Basic tendencies are that for cases with a non- or slowly rotating BH $r_{\rm in}$ decreases as L increases. Unfortunately, however, there are many effects affecting the fitting results and it is not straightforward to derive the information as to the BH spin parameter independently at the present. Gravitational lensing and self-shielding of the radiation from the innermost parts by the outer parts tend to increase the size of the emission region, since for the latter the innermost part with smaller size is obscured and thus instead the surrounding part with a bit larger size is now most visible in X-rays. Conversely, projection, Doppler boosting, and Compton scattering tend to decrease $r_{\rm in}$. In large inclination systems, therefore, self-shielding and Doppler boosting operate in the opposite ways. Hence, the situations are very complex. Caution is needed when interpreting the fitting results of the observational spectra.

Another, but sometimes more useful fitting model is the extended multi-color disk model (or p-free model; Mineshige *et al.* 1994), in which radial temperature profile, $p \equiv -d\log T/d\log r$, is treated as a third fitting parameter (note that p is fixed to be 0.75 in the MCD model). Then, we will be able to discriminate more clearly the standard disk and slim disk, and thus get a good guess for the $L/L_{\rm E}$ ratio. For this fitting, however, we should have good X-ray photon statistics in soft X-ray bands. Extensive study with this model is urgently required.

Let us check what observations tell. Figure 6 shows the interesting results from Watarai, Mizuno, & Mineshige (2001). This is the $T_{\rm in}$-L diagram for BH candidates, including ultra-luminous compact X-ray sources (ULXs). We wish to pay particular attention to the temporal behavior of IC 342, which

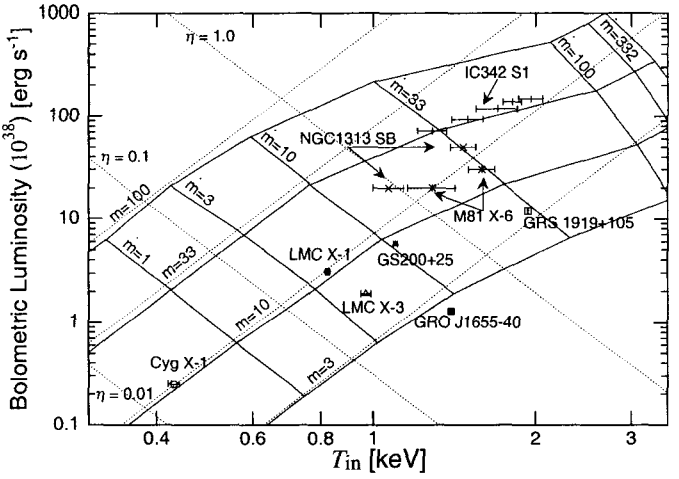


Figure 6: X-ray H-R diagram of X-ray sources. Solid lines represent the constant $m \equiv M_{\rm BH}/M_\odot$ and constant $\dot{m} \equiv \dot{M}c^2/L_{\rm E}$ loci according to our model, while dotted lines are the same but based on the standard accretion disk (SSD) theory. Some observational data of ULXs and Galactic BH candidates are also plotted. (From Watarai, Mizuno, & Mineshige 2001)

evolves along the line of the theoretically predicted track; namely, it clearly shows a decrease in $r_{\rm in}$ with an increase of L. For comparison, we also plot the constant-$r_{\rm in}$ lines with the straight lines. The fact of decreasing $r_{\rm in}$ (i.e. increasing $T_{\rm in}$) with increasing L indicates that the luminosity of IC 342 is comparable to the Eddington luminosity and thus its BH mass is around 30 solar masses. This also provides a good example that even if the estimated inner-edge radius is below 3 $r_{\rm S}$ the BH may not necessarily be an extreme Kerr hole.

The theoretical study of such near- or super-Eddington sources have just started. Interaction between radiation and matter has been a big concern and still remains an open question. We need fully 2D/3D radiation hydrodynamics to establish 2D/3D relativistic slim disk models. Internal Compton scattering in multi-dimension and evaluations of photon trapping effects are also important to consider.

Another interesting issue is the fate of radiation pressure-dominated disk. According to the standard-disk model, the hottest part which is dominated by radiation pressure is violently thermally unstable. How such a disk evolves is

still an open question. Possibilities which have been discussed so far include:

1. Relaxation oscillations leading to quasi-periodic bursting behavior (Honma, Matsumoto, & Kato 1991; Szuszkiewicz & Miller 1998).
2. A soft-to-hard transition which may occur when the viscosity parameter (α) is large (Takeuchi & Mineshige 1997).
3. Strongly clumped disk as is discussed by Krolik (1998).
4. Disk-corona structure which makes the disk body thermally stable (e.g. Nakamura & Osaki 1993; Svensson & Zdziarski 1994).

It is very interesting to note that a Galactic microquasar, GRS1915+105, shows bursting behavior. Recently, Yamaoka *et al.* identified the different phase of light curves and obtained the average spectra in the high (peak) state and in the low (valley) state. They found that the spectra in the high state is characterized by higher $T_{\rm in}(< 3\ r_{\rm S})$, smaller $r_{\rm in}$, and smaller $p(\sim 0.5)$, compared with the values in the low state. This supports the idea of the relaxation oscillations (possibility 1. above). It is also possible, on the other hand, that only an upper part of the disk undergoes transitions to produce coronae transiently as is suggested by Kubota (2001). We expect more observational evidence to accumulate in the next decade.

References

1. Abramowicz, M. A., Czerny, B., Lasota, J.-P., Szuszkiewicz, E. Astrophys. J., **332**, 646 (1988)
2. P. Bak, C. Tang, & K. Wiesenfeld, K. Phys. Rev., **59**, L381 (1987)
3. G.H. Ball, R. Narayan, E. Quataert: Astrophys. J. **552**, 221 (2001)
4. R.D. Blandford, M.C. Begelman: Mon. Not. Roy. Astron. Soc. MNRAS, **303**, L1 (1999)
5. Y.X. Fen, T.P. Li, L. Chen: Astrophys. J. **514**, 373 (1999)
6. A.A. Galeev, R. Rosner, & G.S. Vaiana, Astrophys. J. **229**, 318 (1979)
7. F. Honma, R. Matsumoto, & S. Kato, Pub. Astron. Soc. Japan, **43**, 147 (1991)
8. S. Ichimaru, Astrophys. J. **214**, 840 (1977)
9. V. Igumenshchev, & M.A. Abramowicz, Astrophys. J. Supple., **130**, 463 (2000)
10. V. Igumenshchev, M.A. Abramowicz, & R. Narayan Astrophys. J., **523**, L27 (2000)
11. S. Kato, J. Fukue, & S. Mineshige, Black-Hole Accretion Disks (Kyoto: Kyoto Univ. Press), Chap. 10 (1998)
12. T. Kawaguchi, & S. Mineshige, in *Active Galactic Nuclei and Related Phenomena*, ed. Y. Terzian, E. Khachikian, and D. Weedman (ASP Conf. Ser.), p.356 (1999)

13. T. Kawaguchi, S. Mineshige, M. Machida, R. Matsumoto, & K. Shibata, Pub. Astr. Soc. Japan **52**, L1 (2000)
14. A. Kubota, Ph.D. Thesis, University of Tokyo (2001)
15. T. Kudoh et al. Astrophys. J. ?? (2000?)
16. M. Machida, M.R. Hayashi, & R. Matsumoto, Astrophys. J., **532**, L67 (2000)
17. M. Machida, R. Matsumoto, S. Mineshige: Pub. Astr. Soc. Japan **53**, L1 (2001)
18. T. Manmoto, & S. Mineshige, Astrophys. J. (to be) submitted (2001)
19. T. Manmoto, S. Mineshige, & M. Kusunose, Astrophys. J., **489**, 791 (1997)
20. K. Mitsuda *et al.* Pub. Astron. Soc. Japan, **36**, 741 (1984)
21. S. Mineshige, M. Takeuchi, & H. Nishimori Astrophys. J. **435**, L125 (1994)
22. S. Mineshige, A. Hirano, S. Kitamoto, T.T. Yamada, & J. Fukue, Astrophys. J., **426**, 308 (1994)
23. R. Narayan, in Lighthouses of the Universe, in press (2001)
24. R. Narayan, I. Yi, Astrophys. J. **428**, L13 (1994)
25. R. Narayan, I. Yi, Astrophys. J. **452**, 710 (1995)
26. R. Narayan, R. Mahadevan, & E. Quataert, in The Theory of Black Hole Accretion Disks, ed. M. A. Abramowicz, G. Bjornsson, & J. E. Pringle (Cambridge: Cambridge Univ. Press), 148 (1998)
27. H. Negoro, S. Kitamoto, S. Mineshige, Astrophys. J. **554**, 528 (2001)
28. H. Negoro, S. Kitamoto, M. Takeuchi, & S. Mineshige, Astrophys. J. **452**, L49 (1995)
29. J.F. Hawley, S.A. Balbus, J.M. Stone: Astrophys. J. **554**, L49 (2001)
30. J.F. Hawley, & J.H. Krolik: Astrophys. J. **548**, 348 (2001)
31. E. Szuszkiewicz, J.C. Miller, Mon. Not. Roy. Astron. Soc.,**298**, 888 (1998)
32. F. Takahara, Prog.Theor.Phys. **62**, 629 (1979)
33. M. Takeuchi, & S. Mineshige, Astrophys. J., **486**, 160 (1997)
34. K. Watarai, T. Mizuno, & S. Mineshige, Astrophys. J., **549**, L77 (2001)
35. J.C. Wheeler, Astrophys. J., **214**, 560 (1977)
36. Yamaoka *et al.* in *New Century of X-ray astronomy*, ed. H. Kunieda (PASP conf. series), in press (2001).

ON ENERGETICS AND STRUCTURE OF SUB-PARSEC SCALE JETS IN QUASARS

M. SIKORA
N. Copernicus Astronomical Center, Polish Academy of Sciences,
Bartycka 18, 00-716 Warsaw, Poland
E-mail: sikora@camk.edu.pl

G. MADEJSKI
Stanford Linear Accelerator Center,
Menlo Park, CA 94025, USA
E-mail: madejski@slac.stanford.edu

In our review of sub-parsec scale jets in quasars, we discuss the following issues: observations of parsec and sub-parsec scale jets; energy dissipation and particle acceleration; radiative processes; magnetic fields, pair content and energetics; variability and its relation to the central engine activity. In particular, we describe how internal shocks can explain properties of gamma-ray flares and demonstrate that MeV blazars (those with luminosity peak in the 1-30 MeV range) can be unified with GeV blazars (those with luminosity peak at GeV energies) assuming that in GeV blazars the gamma-ray flares are produced via Comptonization of broad emission lines, whereas in the MeV blazars they result from Comptonization of infrared radiation of hot dust. We also make predictions about the radiative effects of bulk Compton process in the soft X–ray band and show how spectral and variability properties in that band can be used to constrain structure of jets near their bases.

1 Introduction

Observational data imply that jets in radio-loud quasars can extend over 8 decades of distance, from milliparsecs to hundreds of kiloparsecs. Such jets are launched in the vicinity of super-massive black holes and transport energy at a rate which is sometimes comparable with the accretion luminosity, i.e. $10^{45} - 10^{47}$ erg s^{-1}. [24,4] Jets are relativistic and therefore Doppler boosting often makes them observable only on one side of the quasar. However, the hot spots where jets terminate are usually seen on both sides of the central source, and emission on both sides shows similar flux. This suggests that the radiation from hot spots is nearly isotropic, and indicates that jets are "light", i.e. have lower mass densities than the external medium. The lightness is necessary to explain the nonrelativistic speeds of hot spots and formation of extended radio lobe structures. [1] Using the asymmetry of radio brightness of jets and counter-jets, Wardle and Aaron [36] derived the bulk Lorentz factor $\Gamma \sim 3$. Larger Lorentz factors are suggested by X–ray observations [5,33]; however, because X–

ray sensitivity is too low to observe counter-jets, such estimates are not direct and depend on the details of models of the X–ray production.

Quasar jets are imaged in the radio band down to parsec scales. There, they join the central cores, which in turn, due to the synchrotron self-absorption, produce very flat spectra, with the energy spectral index $\alpha < 0.5$ ($\alpha : F_\nu \propto \nu^{-\alpha}$) (see, e.g., review by Zensus.[37] In quasars with jets oriented close to the line of sight, the flux of the radio cores strongly dominates over that from radio lobes, and the total radio spectra are flat. Because of this, core-dominated radio quasars are often called FSRQs (flat spectrum radio quasars), as opposed to the lobe dominated radio quasars, which have steep radio spectra and are called SSRQs (steep spectrum radio quasars). Radio cores in FSRQs are often variable on monthly time scales, and variability of the radio flux is often accompanied by an appearance of ejecta propagating with apparently superluminal velocities.[34,16]

The physics of the parsec scale jets (energetics, magnetic fields, particle distribution) is determined by the use of images, spectra, variability, and polarization properties. In particular, the angle of linear polarization provides information about magnetic field orientation relative to the jet axis. Recent data suggest that such orientation changes from being perpendicular to the jet closer to the central source to the roughly parallel at somewhat larger distances.[32,19] This may indicate that at smaller distances, the jets are free (ballistic) and that their energy is dissipated mainly in perpendicular shocks resulting from collisions of the jet inhomogeneities moving with different radial velocities. At larger distances, on the other hand, most dissipation is likely to come from oblique, reconfinment shocks produced due to interaction of a jet with external medium. For a number of objects, there are also circular polarization measurements available.[12] If circular polarization is produced via the Faraday conversion mechanism, it provides information on the number of electrons at mildly relativistic energies. Combined with the information about the total power of a jet, this can be used to estimate the pair content. It should be emphasized that no information about number of mildly relativistic electrons is available from direct radio flux observations, because the synchrotron radiation by mildly relativistic electrons is self-absorbed.

The parts of jets located close to the cores can be studied also in other spectral bands besides radio, despite the fact that they are by many orders too compact to be resolved. In most FSRQs, typical thermal components of the quasar emission – such as the UV radiation from an accretion disc, the X–ray radiation from the disc corona, and the infrared radiation from dust – are overshined by the Doppler-boosted nonthermal jet radiation. Variability of this radiation is measured on time scales often shorter than those observed in

the radio band, and this strongly suggests its parsec/sub-parsec origin. High polarization in the optical and IR bands implies synchrotron mechanism, while the distinct high energy components, with luminosity peaks in the MeV – GeV range, [35] are presumably products of Comptonization of broad emission lines and IR radiation of hot dust. [28,3] If the EC (external-Compton) scenario is indeed responsible for the high energy components, then observations of the X–ray spectra can provide an exceptional opportunity to study the population of the lower energy population of the relativistic electrons directly. [29] Furthermore, due to narrowness of the electromagnetic spectrum bands covered by dominant broad emission lines and by infrared radiation of hot dust, any features such as breaks in the energy distribution of electrons should be sharply imprinted in the EC radiation components. In particular, as it will be demonstrated below, the shapes of the observed high energy spectra are consistent with the two-power-law injection function of electrons, and the latter can result from a two-step acceleration process of electrons as discussed below.

The two component nonthermal spectra are also observed in BL Lac objects, which differ from FSRQs as having very weak or undetectable emission lines. FSRQs and BL Lac objects form together a class of objects called *blazars*. It is shown by Fossati *et al.* [11] that broadband spectra of blazars form a sequence which can be parametrized by their total luminosities. In this sequence, FSRQs are the most luminous objects. Both their low and high energy spectral components appear to be least extended to the high energies, and their γ–ray luminosities during flares strongly dominate over synchrotron luminosities. The least luminous blazars are represented by the X–ray selected BL Lac objects. Their synchrotron spectra extend up to hard X–rays, and the γ–ray spectra reach TeV energies. γ–ray luminosities in the TeV-emitting BL Lac objects usually do not dominate over synchrotron luminosities. Low luminosity BL Lac objects are probably associated with nuclei of radio galaxies accreting at a low accretion rate, and γ–rays are very likely produced by the SSC (synchrotron-self-Compton) process. In the more luminous, radio selected BL Lac objects the EC process can be dominant, [20] in similarity to FSRQs.

In this presentation we focus on jets in quasars, where the interaction of the jet with external radiation is significant and the presence or lack of radiative effects of this interaction can be used to determine physical parameters and constraints on structure of jets on sub-parsec scales. This paper is organized as follows: in §2, we analyze (***) energetics of intrinsic collisions of material in the jet; in §3, we present the motivation for introduction of the two-step stochastic acceleration of electrons; in §4, we discuss the formation of spectral breaks in the high energy spectral components; in §5, §6, and §7, respectively, we derive the magnetic field intensity, the pair content, and the average power

of a jet. In §8, we formulate our predictions regarding the production of soft X–ray precursors via Comptonization of the external radiation by the cold electrons in a jet, and list the main conclusions in §9.

2 Energy dissipation

Cooling time scales of ultrarelativistic electrons, which in quasar jets produce synchrotron radiation in the optical band and inverse-Compton radiation in the MeV – GeV range, is much shorter than the dynamical/propagation time scale. Therefore, such electrons must be accelerated *in situ*, in sites where the jet loses part of its energy and the shocks are formed. The jet energy can be dissipated in internal shocks – formed in collisions between inhomogeneities in a jet [28,31] or following reconnection of magnetic fields, [25] – and/or in external shocks, formed via reconfinment of the jet by external medium [26,17] or due to collisions of a jet with external clouds [8]. Perpendicular orientation of magnetic fields to the jet axis, inferred from the polarization measurements in the optical band [13] and at the high radio frequencies [19] favors the scenario where at sub-parsec distances, the particles are accelerated in transverse shocks which are likely to be produced via collisions of inhomogeneities flowing down the jet with different radial velocities. A given collision event is then responsible for a given flare, and the time scale of such a collision, as measured in the comoving frame of the shocked plasma, is

$$t'_{coll} \simeq t_{fl}\mathcal{D}\,, \tag{1}$$

where t_{fl} is the observed time scale of the flare, and

$$\mathcal{D} \equiv \frac{1}{\Gamma(1-\beta\cos\theta_{obs})} \tag{2}$$

is the Doppler factor of the shocked plasma. The amount of energy dissipated during the collision is

$$E'_{diss} \simeq \frac{L'_{fl}t'_{coll}}{\eta_{rad}\eta_e} \simeq \frac{L_{fl}t_{fl}}{f\mathcal{D}^3\eta_{rad}\eta_e}\,, \tag{3}$$

where η_e is the fraction of dissipated energy used to accelerate electrons, η_{rad} is the (average) radiative efficiency of electrons, and

$$L_{fl} \equiv 4\pi\left(\frac{\partial L_{fl}}{\partial\Omega}\right) = 4\pi\left(\frac{\partial L'_{fl}}{\partial\Omega'}\right)\mathcal{D}^4 = L'_{fl}f\mathcal{D}^4 \tag{4}$$

is the apparent luminosity. All primed quantities are as measured in the source comoving frame. The factor f expresses the additional dependence of the apparent luminosity on the observed angle, if the intrinsic radiation is anisotropic. For the EC process, $f \simeq (\mathcal{D}/\Gamma)^2$, while for synchrotron and SSC radiation $f \simeq 1$.[7]

In the above relations it was assumed that during a collision, the shocked plasma moves with a constant bulk Lorentz factor. In general this is not true: depending on the initial parameters of inhomogeneities, the shock can accelerate or decelerate, and during the collision the shock structure can evolve from a double one to a single shock (forward or reverse). However, if inhomogeneities are intrinsically identical, then the forward-reverse shock structure is symmetric in the shocked plasma frame (= discontinuity contact surface frame), and the bulk Lorentz factor of the shock is constant. This case is adopted in our quantitative analysis, not only for simplicity, but also because it provides the highest efficiency of energy dissipation.

Assuming that the inertia of inhomogeneities is dominated by protons (i.e. $n_e/n_p \ll m_p/m_e$), and noting that

$$E_{diss} \equiv \eta_{diss}(E_1 + E_2) = \Gamma E'_{diss}\,, \tag{5}$$

where $E_1 \simeq (N_p/2)\Gamma_1 m_p c^2$ and $E_2 \simeq (N_p/2)\Gamma_2 m_p c^2$ are energies of inhomogeneities, we find that total number of protons involved in the collision is

$$N_p \simeq \frac{2\Gamma E'_{diss}}{\eta_{diss}(\Gamma_1+\Gamma_2)m_p c^2} \simeq \frac{2L_{fl}t_{fl}}{\kappa f \mathcal{D}^3 \eta_{rad}\eta_e m_p c^2}\,, \tag{6}$$

where

$$\kappa \equiv \frac{E'_{diss}}{N_p m_p c^2} = \frac{\eta_{diss}(\Gamma_1+\Gamma_2)}{2\Gamma} \tag{7}$$

is amount of energy dissipated per proton in units of $m_p c^2$. For $\Gamma_2 > \Gamma_1 \gg 1$, energy and momentum conservations give

$$\Gamma \simeq \sqrt{\Gamma_1\Gamma_2} \tag{8}$$

and

$$\eta_{diss} \simeq \frac{((\Gamma_2/\Gamma_1)^{1/2}-1)^2}{(\Gamma_2/\Gamma_1)+1}\,, \tag{9}$$

and then

$$\kappa \simeq \frac{((\Gamma_2/\Gamma_1)^{1/2}-1)^2}{2(\Gamma_2/\Gamma_1)^{1/2}}\,. \tag{10}$$

For $\Gamma_2/\Gamma_1 = 2.5$, which corresponds with $\eta_{diss} \simeq 0.1$, $\kappa \simeq 0.11$.

3 Particle acceleration

As it was demonstrated by Sikora and Madejski, [29] energy flux in powerful jets in quasars cannot be dominated by pair plasma. This is because such jets would produce much larger fluxes of soft X–ray radiation than is observed. Hence, we assume that the inertia of jets is dominated by protons. In this case, the structure of shocks and the structure of the turbulence around them are both determined by protons. Provided that the magnetic fields are amplified up to equipartition with protons, the protons are effectively accelerated both by 1st order and by 2nd order Fermi process. [15] Those which reach energies $> 10^9$ GeV, interact efficiently with ambient photons and trigger (mainly via photo-meson process) the synchrotron-supported pair cascades. [21] However, such a model fails to reproduce very hard X–ray spectra of FSRQs. [30] This may indicate that too few protons reach sufficiently high energies to power pair cascades.

Alternative possibility is that the high energy radiation in FSRQs is produced by directly accelerated electrons. However, for the acceleration of electrons to occur via the Fermi process, these electrons must be first preheated/preaccelerated up to energies γ_b, at which point, the magnetic rigidity of electrons becomes comparable with magnetic rigidity of thermal protons, i.e. when their momenta are equal:

$$m_e\sqrt{\gamma_b^2-1} \simeq m_p\sqrt{\gamma_{p,th}^2-1}\,, \tag{11}$$

where

$$\gamma_{p,th}-1 = \eta_{p,th}\kappa \tag{12}$$

is the average thermal proton energy in the shocked plasma, and $\eta_{p,th}$ is the fraction of the dissipated energy tapped to heat the protons. For the reasonable assumption that $\Gamma_2/\Gamma_1 < 10$, $\gamma_{p,th}-1 < 1$, i.e. the thermal proton plasma is at most mildly relativistic.

Noting that the average energy of injected electrons is

$$\bar{\gamma}_{inj}-1 \simeq \frac{n_p m_p}{n_e m_e}\eta_e\kappa = \frac{n_p m_p}{n_e m_e}\frac{\eta_e}{\eta_{p,th}}(\gamma_{p,th}-1)\,, \tag{13}$$

we find that for γ_b and $\bar{\gamma}_{inj} \gg 1$

$$\frac{\gamma_b}{\bar{\gamma}_{inj}} \sim \frac{n_e}{n_p}\frac{\eta_{p,th}}{\eta_e}\left(\frac{\gamma_{p,th}+1}{\gamma_{p,th}-1}\right)^{1/2}. \tag{14}$$

Hence, for non- or mildly relativistic shocks and for $\eta_{p,th} \sim \eta_e$, the threshold energy for the Fermi acceleration of electrons significantly exceeds the average

energy of electrons, even if $n_e = n_p$. That implies that the often-considered bulk preheating process is not able to provide an adequate number of electrons with $\gamma \geq \gamma_b$. But, as it was recently demonstrated in numerical PIC (particle-in-cell) simulations, the preheating/preacceleration mechanism has a stochastic character and a significant fraction of electrons can reach such energies. [9,27] Furthermore, the fact that in FSRQs, the X–ray spectral index α_X is never negative, implies that most of electrons occupy the lowest energies, i.e. the median Lorentz factor is $\ll \bar{\gamma}_{inj}$. Therefore, this scenario does not predict a thermal bump in the preaccelerated electron energy distribution. Instead, the preacceleration mechanism, in similarity to the Fermi process, produces a power-law energy distribution of electrons. It should be emphasized here that because the X–rays in FSRQs are produced in the slow cooling regime (see next Section), electrons with low energies are not populated by cooling effect, but are directly injected with such energies.

4 Electromagnetic spectra

The basic feature of the high energy spectra in FSRQs — a spectral break between the X–ray and the γ–ray bands — has a natural explanation in terms of the EC model. In this model, X–ray spectra are produced by electrons with radiative cooling time scale t'_{cool}, longer than the collision time scale t'_{coll} (*slow cooling regime*), whereas γ–rays are produced by electrons with $t'_{cool} < t'_{coll}$ (*fast cooling regime*). Noting that the angle averaged cooling rate of electrons, dominated by Comptonization of external radiation, is

$$|\dot{\gamma}| \simeq \frac{\sigma_T}{m_e c} u'_{ext} \gamma^2 \,, \tag{15}$$

we obtain that the angle averaged cooling time scale is

$$t'_{cool} \simeq \frac{\gamma}{|\dot{\gamma}|} \simeq \frac{m_e c}{\sigma_T} \frac{1}{\gamma u'_{ext}} \,, \tag{16}$$

where u'_{ext} is the energy density of an external radiation field. Then, from $t'_{cool} = t'_{coll}$, where t'_{coll} is given by Eq. (1), the break in the electron distribution is at energy

$$\gamma_c \simeq \frac{m_e c}{\sigma_T} \frac{1}{u'_{ext} t_{fl} \mathcal{D}} \,. \tag{17}$$

For $\gamma < \gamma_c$, the slope of the electron distribution is the same as the slope of the injection function; for $\gamma > \gamma_c$, the slope of the electron energy distribution is steeper by $\Delta s = 1$ ($s : N_\gamma \propto \gamma^{-s}$).

Since time scales of flares in FSRQs are rarely shorter than 1 day, the distances from the central source where they are produced,

$$r_{fl} \sim (r_{fl}/\Delta r_{coll}) c t_{fl} \mathcal{D} \Gamma \,, \tag{18}$$

are expected to be larger than 0.1 parsec. At such distances, contribution to u'_{ext} is dominated by the diffuse components of the broad emission line light and of the infrared radiation of hot dust. Noting that

$$u'_{ext} = \frac{1}{c} \int I'_{ext} d\Omega' = \frac{1}{c} \int I_{ext} \mathcal{D}_{in}^{-2} d\Omega \simeq u_{diff} \Gamma^2 \tag{19}$$

where

$$\mathcal{D}_{in} = \frac{1}{\Gamma(1 - \beta \cos\theta_{in})} \tag{20}$$

and θ_{in} is the angle between the photon direction and the jet axis, we predict that the break in an electron energy distribution at γ_c should be imprinted in the electromagnetic spectrum at frequency

$$\nu_{EC,c} \simeq \mathcal{D}^2 \gamma_c^2 \nu_{diff} \simeq \left(\frac{m_e c^2}{\sigma_T} \right)^2 \frac{\nu_{diff}}{u_{diff}^2 t_{fl}^2 \Gamma^4} \tag{21}$$

and that the spectrum should change the slope around $\nu_{EC,c}$ by $\Delta\alpha_{X,\gamma} \simeq 0.5$. For typical energy densities expected in the BELR (broad emission line region) as well as that provided by the hot dust radiation field, the break is predicted to be located in the range of $10^{20} - 10^{22}$ Hz (Sikora *et al.*), in agreement with observations.[35,11]

In most FSRQs, during high energy flares the slope of γ–ray spectra measured by EGRET [23] is $\alpha_\gamma \leq 1$, while the X–ray spectra show slopes $\alpha_X \simeq 0.6 - 0.7$.[6,10] The smaller than the predicted value 0.5 of $\Delta\alpha_{X,\gamma}$ can be explained by taking into account contribution to the soft and mid-X–ray bands of the SSC component [14,18]. More challenging are the so called "MeV blazars." In these objects, the spectra peak in the 300 keV - 30 MeV range and the γ–ray spectra are very soft ($\alpha_\gamma > 1.4$), while the X–ray spectra are very hard ($\alpha_X \leq 0.5$).[33] The values of $\Delta\alpha_{X,\gamma} \geq 1.0$ measured in those sources are much too large to be explained solely by the effect of cooling, and instead, it is necessary to postulate an additional break in the electron energy distribution besides γ_c. It is tempting to identify such a break with the break at γ_b, predicted by the two-step acceleration process discussed in the Section 3. External photons scattered by electrons with energies γ_b, are boosted to energies

$$\nu_{EC,b} \simeq \mathcal{D}^2 \gamma_b^2 \nu_{diff} \,. \tag{22}$$

For $2.5 < \Gamma_2/\Gamma_1 < 10$ and $\eta_{p,th} \sim 0.5$, the break in the electron injection function is $600 < \gamma_b < 1700$ (see Eqs. 10-12 and Table 1), and should appear in the electromagnetic spectrum at around 1 GeV if $\nu_{diff} = \nu_{BEL}$, and at lower energies (by a factor of $\sim 10 - 30$) if $\nu_{diff} = \nu_{IR}$. In the former case $\nu_{EC,b} \gg \nu_{EC,c}$ and the γ–ray spectra in the EGRET range should have slopes $\alpha_\gamma \leq 1$, while in the latter case $\nu_{EC,b} \sim \nu_{EC,c} \leq 30$ MeV, and the γ–ray spectra in the EGRET range should be soft. Hence, we conclude that for blazars producing most of their radiation in the GeV band (and hereafter called GeV blazars) the γ–ray flux is probably produced at smaller distances (closer to the central source), where the production of γ–rays is dominated by Comptonization of broad emission lines, while the MeV blazar phenomenon is likely to be produced at larger distances, where production of γ–rays is dominated by the Comptonization of infrared radiation from dust.

The above scheme can also explain two other differences between MeV blazars and GeV blazars. One, already mentioned above, is that in MeV blazars the X–ray spectra should be much harder than in GeV blazars. This can result from the lower contribution of the SSC component to the X–ray band in MeV blazars as compared to GeV blazars (Sikora *et al.*, in preparation). The other difference is that in the spectra of MeV blazars, in contrast to GeV blazars, the thermal UV bump is quite prominent. [33] This difference can be explained noting that magnetic fields are weaker at larger distances, and therefore in MeV blazars the synchrotron spectra are shifted to lower frequencies, revealing the UV bump.

5 Magnetic fields

Electrons with energies from the range $[\gamma; \gamma + d\gamma]$ produce synchrotron radiation at the rate

$$L_{\nu'_{syn}} d\nu'_{syn} \sim (N_\gamma d\gamma)|\dot{\gamma}_{syn}| m_e c^2 \sim (N_\gamma d\gamma) c\sigma_T \gamma^2 u'_B \tag{23}$$

and the angle averaged EC radiation at the rate

$$L_{\nu'_{EC}} d\nu'_{EC} \sim (N_\gamma d\gamma)|\dot{\gamma}_{EC}| m_e c^2 \sim (N_\gamma d\gamma) c\sigma_T \gamma^2 u'_{ext} \tag{24}$$

where

$$\nu'_{syn} \sim (2e/3\pi m_e c)\gamma^2 B' \tag{25}$$

$$\nu'_{EC} \sim \gamma^2 \nu'_{ext} \tag{26}$$

and $u'_B = B'^2/8\pi$ is the magnetic energy density. Noting that $L_{\nu_{syn}} d\nu_{syn} = L_{\nu'_{syn}} d\nu'_{syn} \mathcal{D}^4$ and $L_{\nu_{EC}} d\nu_{EC} = L_{\nu'_{EC}} d\nu'_{EC} (\mathcal{D}^6/\Gamma^2)$ we find that the ratio of

synchrotron to EC luminosity, both produced by the same population of electrons in the same energy range, radiating in the optically thin regime, is

$$\mathcal{R}_{syn/EC} \equiv \frac{\nu_{syn} L_{\nu_{syn}}}{\nu_{EC} L_{\nu_{EC}}} \sim \frac{u'_B}{u'_{ext}} \left(\frac{\Gamma}{\mathcal{D}}\right)^2 . \tag{27}$$

From Eqs. (17) and (21), we have

$$u'_{ext} \simeq \frac{m_e c}{\sigma_T} \frac{1}{t_{fl} (\nu_{EC,c}/\nu_{diff})^{1/2}} \tag{28}$$

and inserting this into Eq. (27), we obtain

$$u'_B \simeq u'_{ext} \mathcal{R}_{syn/EC} (\mathcal{D}/\Gamma)^2 \simeq \frac{m_e c}{\sigma_T} \frac{\mathcal{R}_{syn/EC} (\mathcal{D}/\Gamma)^2}{t_{fl} (\nu_{EC,c}/\nu_{diff})^{1/2}} . \tag{29}$$

For $t_{fl} \sim 3$ day, $\mathcal{R}_{syn/EC} = 0.1$, $\nu_{EC,c} = 10$ MeV and $\nu_{diff} = 10$ eV, and the observer located at $\theta_{obs} = 1/\Gamma$, Eq. (29) gives $B' \sim \sqrt{8\pi u'_B} \sim 0.6$ Gauss.

We can estimate the energy density of the relativistic electrons and determine whether they are in equipartition with magnetic fields. Relations $E'_{rad} = \eta_{rad} \eta_e E'_{diWess}$ and $E'_e = (1 - \eta_{rad}) \eta_e E'_{diss}$ give

$$E'_e = \frac{1 - \eta_{rad}}{\eta_{rad}} E'_{rad} , \tag{30}$$

where

$$E'_{rad} \sim L'_{fl} t'_{coll} \sim \frac{L_{fl} t_{fl}}{\mathcal{D}^3 (\mathcal{D}/\Gamma)^2} . \tag{31}$$

Hence, energy density of the relativistic electrons is

$$u'_e = \frac{E'_e}{\pi a^2 \Delta'_{sh}} \simeq \frac{1 - \eta_{rad}}{\eta_{rad}} \frac{L_{fl}}{2\pi c^3 (a/a_\Gamma)^2 t_{fl}^2 g \mathcal{D}^6 (\mathcal{D}/\Gamma)^2} , \tag{32}$$

where $a \leq a_\Gamma = c t_{fl} \mathcal{D}$ is the cross-sectional dimension (radius) of a shocked plasma sheet, and $\Delta'_{sh} = 2 c t_{fl} \mathcal{D} g$ is the thickness of the sheet. The scaling factor g, which relates the thickness of the shocked plasma sheet to the time scale of the flare/collision, depends on Γ_2/Γ_1 and on the adiabatic index. The analytical form of g is derived and presented in Moderski *et al.* (in preparation), and the example values of g are shown in the Table 1.

Combining Eqs. (29) and (32), we obtain

$$\frac{u'_B}{u'_e} \simeq \frac{2\pi m_e c^4}{\sigma_T} \frac{\eta_{rad}}{1 - \eta_{rad}} \frac{g (a/a_\Gamma)^2 \mathcal{R}_{syn/EC}}{(\nu_{EC,c}/\nu_{diff})^{1/2}} \frac{t_{fl}}{L_{fl}} \mathcal{D}^6 (\mathcal{D}/\Gamma)^4 . \tag{33}$$

Table 1: Dependence of the internal shock parameters on Γ_2/Γ_1. Values of g_0 and g are calculated for the adiabatic index of the shocked plasma $\hat{\gamma} = 5/3$. Analytical formulas for g_0 and g are presented in Moderski *et al.* (in preparation).

Γ_2/Γ_1	η_{diss}	κ	$\gamma_{p,th} - 1$	γ_b	$\bar{\gamma}_{inj}$	n_+/n_p	g_0	g
1.5	0.020	0.020	0.010	265	15.3	0.2	0.273	0.067
2.5	0.097	0.107	0.053	609	18.0	2.4	0.641	0.150
5.0	0.255	0.342	0.171	1120	20.4	7.6	1.237	0.255
10.0	0.425	0.739	0.370	1722	22.2	15.5	2.025	0.346

The equation demonstrates very strong dependence of u'_B/u'_e on Γ and on the observed angle θ_{obs}. Provided that blazars are observed at $\theta_{obs} \sim 1/\Gamma$, the equipartition condition, $u'_B \sim u'_e$, implies

$$\Gamma \simeq \left(\frac{\sigma_T}{2\pi m_e c^4} \frac{(\nu_{EC,c}/\nu_{diff})^{1/2}}{\mathcal{R}_{syn/EC}\, g\, (a/a_\Gamma)^2} \frac{1-\eta_{rad}}{\eta_{rad}} \frac{L_{fl}}{t_{fl}} \right)^{1/6} . \tag{34}$$

For flares with luminosities $L_{fl} \sim 10^{48-49}$ergs s^{-1} and time scales $t_{fl} \sim 3$d, and assuming $a = a_\Gamma$, $\Gamma_2/\Gamma_1 = 2.5$, and $\eta_{rad} = 0.5$, we obtain $\Gamma \sim 18 - 26$.

6 Pair content

The pair content is

$$\frac{n_+}{n_p} = (\frac{n_e}{n_p} - 1)/2 \tag{35}$$

where

$$\frac{n_e}{n_p} = \eta_e \frac{m_p \kappa}{m_e(\bar{\gamma}_{inj} - 1)} \tag{36}$$

$$\bar{\gamma}_{inj} = \frac{\int_1^{\gamma_{max}} \gamma Q\, d\gamma}{\int_1^{\gamma_{max}} Q\, d\gamma} \tag{37}$$

and Q is the electron injection function. For a double power-law injection function, $Q \propto \gamma^{-p_L}$ for $1 < \gamma < \gamma_b$, and $Q \propto \gamma^{-p_H}$ for $\gamma_b < \gamma < \gamma_{max}$, and observationally suggested indices, $1 < p_L < 2$ and $p_H > 2$,

$$\bar{\gamma}_{inj} \simeq \frac{(p_L - 1)(p_H - p_L)}{(2 - p_L)(p_H - 2)} \gamma_b^{2-p_L} . \tag{38}$$

For $p_L = 1.8$ ($\rightarrow \alpha_X = (p_L - 1)/2 \sim 0.45$ and $\alpha_\gamma = p_L/2 \sim 0.9$ in the EGRET band of GeV blazars), and $p_H = 2.8$ ($\rightarrow \alpha_\gamma = p_H/2 = 1.4$ in the EGRET band of MeV blazars), and $2.5 < \Gamma_2/\Gamma_1 < 10$, the number of pairs per proton is $2.4 < n_+/n_p < 15.5$ (see Table 1).

7 The power of a jet

Energy carried by the cold inhomogeneities prior to the collision is (see Eq. 6)

$$E_i \simeq \frac{N_p}{2} m_p c^2 \Gamma_i \simeq \frac{L_{fl} t_{fl}}{\kappa f \mathcal{D}^3 \eta_{rad} \eta_e} \Gamma_i \,, \tag{39}$$

where $i = 1, 2$. The time it takes to eject the inhomogeneities from the central source is

$$t_{ej,i} \simeq \frac{\lambda_0}{c \beta_i \Gamma_i} \simeq \frac{\mathcal{D}}{\Gamma_i} g_0 t_{fl} \tag{40}$$

where $\lambda_0 = g_0 c \mathcal{D} t_{fl}$ is the proper length of an inhomogeneity (the thickness of a shell segment), and g_0 is the factor depending on Γ_2/Γ_1 and on the adiabatic index of the shocked plasma (see Table 1). The time-averaged power of a jet is

$$P_j \sim \Psi \frac{E_1 + E_2}{t_{ej,1} + t_{ej,2}} \simeq \Psi \frac{N_p m_p c^3 \Gamma^2}{\lambda_0} \simeq \Psi \frac{L_{fl}}{2 \eta_e \eta_{rad} \kappa g_0} \frac{\Gamma^2}{f \mathcal{D}^4} \,, \tag{41}$$

where Ψ is the duty cycle factor. For $L_{fl} \sim 10^{48} \mathrm{erg\ s^{-1}}$, $\mathcal{D} \sim \Gamma \sim 15$, $\Gamma_2/\Gamma_1 \sim 2.5$, $\eta_e \sim \eta_{rad} \sim 0.5$ and $\Psi \sim 0.1$, Eq. (41) gives $P_j \sim 3 \times 10^{46}$ erg s^{-1}.

8 X–ray precursors

Prior to the collision, electrons in inhomogeneities are cold, but being pulled by streaming protons they scatter external photons and produce the beamed bulk-Compton radiation.[2,29] The spectrum of this radiation peaks around

$$\nu_i \simeq \mathcal{D}_i \Gamma_i \nu_{diff} \tag{42}$$

and has luminosity

$$L_{X,i} \simeq \delta N_{e,i} |\dot{\epsilon}_{e,i}| \mathcal{D}_i^4 \,, \tag{43}$$

where

$$\delta N_{e,i} = (N_e/2) \mathrm{Min}[1; r_{BC}/\lambda_0 \mathcal{D}_i] \,, \tag{44}$$

$$N_e \simeq \frac{E'_{e,inj}}{\bar{\gamma}_{inj} m_e c^2} \simeq \frac{\eta_e E'_{diss}}{\bar{\gamma}_{inj} m_e c^2} \simeq \frac{L_{fl} t_{fl}}{f \mathcal{D}^3 \eta_{rad} \bar{\gamma}_{inj} m_e c^2} \,, \tag{45}$$

$$|\dot{\epsilon}_{e,i}| \simeq \Gamma_i^2 c \sigma_T u_{ext} \,, \tag{46}$$

$$\mathcal{D}_i \equiv \frac{1}{\Gamma_i (1 - \beta_i \cos\theta_{obs})} \,, \tag{47}$$

r_{BC} is the distance within which most of the bulk Compton radiation is produced. The factor $\mathrm{Min}[1; r_{BC}/\lambda_0 \mathcal{D}_i]$ in Eq. (44) expresses the fact that in

general, due to light travel effects and the finite length of the source, not all electrons contribute to the radiation observed at a given instant.

Bulk-Compton flares from cold inhomogeneities are expected to precede the nonthermal flares by

$$\delta t_i \sim \frac{r_{fl}}{c\Gamma_i \mathcal{D}_i} \simeq \frac{\mathcal{D}\Gamma}{\mathcal{D}_i \Gamma_i} t_{fl} \frac{\Delta r_{coll}}{r_{fl}} \,. \tag{48}$$

Above formulae are very approximate, because they do not take into account that the observed bulk Compton radiation at any given instant includes contributions from a finite distance range over which u_{ext} can change significantly. However, making the conservative assumption that the jet is developed (accelerated and collimated) not earlier than at a distance where electron energy losses are already dominated by Compton interactions with the broad emission line flux rather than the direct radiation from an accretion disc, we can calculate the minimum bulk Compton luminosity, using an approximation that $u_{ext} \simeq u_{BEL} \sim$ const for $r < r_{BEL}$ and 0 for $r > r_{BEL}$. Noting that $r_{BEL} \sim 3 \times 10^{17} \sqrt{L_{UV,46}}$ cm and $u_{BEL} \sim 0.03$ erg cm^{-3} (Peterson [22]), we find that for $\theta_{obs} = 1/\Gamma = 1/15$, $\Gamma_2/\Gamma_1 = 2.5$, $L_{fl} = 10^{48}$ erg s^{-1} and $t_{fl} = 3$ day, the precursor produced by the faster-moving inhomogeneity should have luminosity $L_{X,2} \sim 3 \times 10^{46}$ ergs s^{-1}; should peak at around h$\nu_2 \simeq 3.2$ keV; and should precede the nonthermal flare by $\delta t_2 \simeq 0.7 t_{fl}$. The precursor produced by the slower moving inhomogeneity would have luminosity $L_{X,1} \sim 5 \times 10^{45}$ erg s^{-1}; would peak around h$\nu_1 \simeq 1.8$ keV; and would precede the nonthermal flare by $\delta t_1 \simeq 1.8 t_{fl}$. Since typical soft X–ray luminosities are of the order 10^{46} erg s^{-1}, we can infer from the above estimates that jets propagating through much denser quasi-isotropic radiation fields than would be provided by the BELR are predicted to produce much stronger soft/mid X–ray radiation than is observed. This puts severe constrains on the jet structure at $10^2 - 10^3$ gravitational radii from the BH, where quasi-isotropic radiation is strongly dominated by the very dense direct radiation of an accretion disc and its corona. Hence, the lack of very large soft X–ray excesses suggests that jets near their bases are very wide (meaning that they have relatively large opening angle) and/or are slower than at larger distances.

9 Summary

Multi-wavelength observations of blazars provide a valuable tool allowing us to learn about the structure and physical properties of sub-parsec-scale jets in quasars. The observations suggest:

• Rapid flares and transverse magnetic fields inferred from polarization data → *Streams of matter in small-scale jets are unsteady, and collisions between*

inhomogeneities propagating down the jet with different radial velocities can easily form transverse shocks;
• Large powers of jets and relatively low soft/mid X–ray luminosities → *Inertia of jets is dominated by protons and the number of pairs per proton is* < 15;
• Very hard X–ray spectra → *X–rays and* γ*–rays are produced by electrons accelerated directly, just as is the case for the synchrotron radiation, rather than being injected following the proton-induced pair cascades*;
• Differences between properties of GeV and MeV blazars → *EC model can explain these differences assuming a two-power-law electron injection function. The break in the injection function can be related to the threshold energy for the Fermi acceleration of electrons*;
• Luminous, short term γ–ray flares in FSRQs → *magnetic fields are much below equipartition with the relativistic plasma, unless* $\Gamma \geq 18 - 26$;
• No prominent spectral signatures of bulk Compton process in the soft/mid X–ray bands → *Collimation and/or acceleration of jets occurs beyond* $10^2 - 10^3$ *gravitational radii from the central source.*

Acknowledgments

This work was partially supported by Polish KBN grant 5 P03D 002 21 and NASA Chandra observing grant to Stanford University. M.S. thanks the Stanford Linear Accelerator Center for its hospitality during the completion of this work, and the Korea Institute for Advanced Study for hosting this interesting meeting.

References

1. Begelman, M.C., & Cioffi, D.C. 1989, ApJ, 345, L21.
2. Begelman, M.C., & Sikora, M. 1987, ApJ, 322, 650.
3. Błażejowski, M., Sikora, M., Moderski, R., & Madejski, G.M. 2000, ApJ, 545, 107.
4. Celotti, A., & Fabian, A.C. 1993, MNRAS, 228.
5. Celotti, A., Ghisellini, G., & Chiaberge, M. 2001, MNRAS, 321, L1.
6. Comastri, A., Fossati, G., Ghisellini, G., & Molendi, S. 1997, ApJ, 480, 534.
7. Dermer, C.D. 1995, ApJ, 446, L63.
8. Dermer, C.D., & Chiang, J. 1998, NewA, 3, 157.
9. Dieckmann, M.E., McClements, K.G., Chapman, S.C., Dendy, R.O., & Drury, L.O'C. 2000, A&A, 356, 377.
10. Donato, D., Ghisellini, G., Tagliaferri, G., & Fossati, G. 2001, A&A, 375, 739.

11. Fossati, G., Maraschi, L., Celotti, A., Comastri, A., & Ghisellini, G. 1998, MNRAS, 299, 433.
12. Homan, D.C., Attridge, J.M., & Wardle, J.F.C. 2001, 556, 113.
13. Impey, C.D., Lawrence, C.R., & Tapia, S. 1991, ApJ, 375, 46.
14. Inoue, S., & Takahara, F. 1996, ApJ, 463, 555.
15. Jones, F.C. 1994, ApJS, 90, 561.
16. Jorstad, S.G., Marscher, A., Mattox, J.R., Wehrle, A.E., Bloom, S.D., & Yurchenko, A.V. 2001, ApJS, 134, 181.
17. Komissarov, S.S., & Falle, S.A.E.G. 1997, MNRAS, 288, 833.
18. Kubo, H., *et al.* 1998, ApJ, 504, 693.
19. Lister, M.L. 2001, ApJ, in press [astro-ph/0107594].
20. Madejski, G.M., Sikora, M., Jaffe, T., Błażejowski, M., Jahoda, K., & Moderski, R. 1999, ApJ, 521, 145.
21. Mannheim, K., & Biermann, P.L. 1992, A&A, 253,, L21.
22. Peterson, B.M. 1993, PASP, 105, 247.
23. Pohl, M., Hartman, R.C., Jones, B.B., & Sreekumar, P. 1997, A&A, 326, 51.
24. Rawlings, S., & Saunders, R. 1991, Nature, 349, 138.
25. Romanova, M.M., & Lovelace, R.V.E. 1992, A&A, 262, 26.
26. Sanders, R.H. 1983, ApJ, 266, 73.
27. Shimada, N., & Hoshino, M. 2000, ApJ, 543, L67.
28. Sikora, M., Begelman, M.C., & Rees, M.J. 1994, ApJ, 421, 153.
29. Sikora, M., & Madejski, G.M. 2000, ApJ, 534, 109.
30. Sikora, M., & Madejski, G.M. 2001, AIP Conference Proceedings, 558, 275.
31. Spada, M., Ghisellini, G., Lazzati, D., Celotti, A. 2001, MNRAS, 325, 1559.
32. Stevens, J.A., Robson, E.I., & Holland, W.S. 1996, ApJ, 462, L23.
33. Tavecchio, F., Maraschi, L., Sambruna, R.M., & Urry, C.M. 2000, ApJ, 544, L23.
34. Vermeulen, R.C., & Cohen, M.H. 1994, ApJ, 430, 467.
35. von Montigny, C., *et al.* 1995, ApJ, 440, 525.
36. Wardle, J.F.C., & Aaron, S.E. 1997, MNRAS, 286, 425.
37. Zensus, J.A. 1997, ARAA, 35, 607.

LARGE SCALE JETS AND THE NUCLEAR ENGINE

A. CELOTTI
S.I.S.S.A., via Beirut 2-4,
34014 Trieste, Italy
E-mail: celotti@sissa.it

A few recent observational findings following from Chandra and HST observations on different AGN scales are highlighted and their implications on nuclear energy sources associated with radio loud objects and the accretion/jet connection are briefly discussed.

1 Introduction

Ejection of highly relativistic and highly collimated flows, i.e. jets, transporting a significant amount of the total power characterises the nuclear activity in a sizeable fraction of AGN and in at least some galactic black hole candidates in binary systems. Furthermore outflows might be commonly present even in 'radio-quiet' AGN as deep (radio) observation seem to suggest. It is therefore crucial to understand how the accretion and ejection of matter and energy take place in the vicinity of black holes and if/how these two processes are connected.

Although large scale jets are physically and temporally very distant from the engine itself, typically by seven orders of magnitude in size (and age), the understanding of their emission and interaction with any external medium can provide very important information on the energetics and composition of such flows, obviously crucial quantities to speculate on and understand their formation.

Furthermore, more direct inferences on the radiative power associated with dissipation in accreting flows can be drawn from the direct detection of nuclear emission and/or its effects on circumnuclear gas.

In the following I will briefly try to present some of the recent phenomenological findings which provide such pieces of information and their implications.

2 Interaction between lobes and the ICM: The case of Perseus

Signs already detected by ROSAT of the interaction between the radio emitting lobes associated with extragalactic jets and the external intracluster medium (ICM) have become impressively clearer thanks to Chandra imaging (e.g. Böhringer et al. 1993, Forman et al. 2001). Information on the interplay

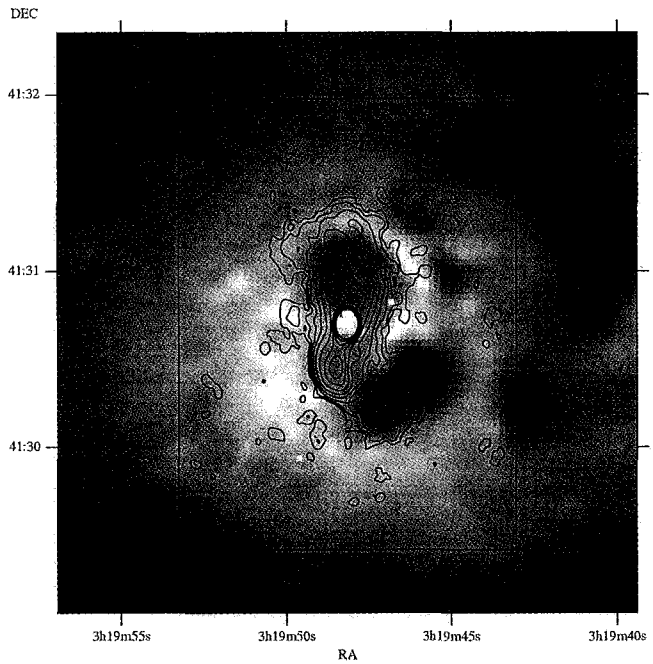


Figure 1: 1.4 GHz contours of the radio source 3C 84 overlaid on the Chandra image of the core of the Perseus cluster (from Fabian et al. 2000, 2002).

between these two components provides us with clues on the jet power as well as on the possible transfer of energy from jets to the external gas.

In the handful of objects detected so far the X–ray emission appears to be anticorrelated with the radio one, with depression of the X–rays within the lobes of the radio sources on scales of tens of kpc. Furthermore there is no evidence of formation of shocks at the interface between the two media – but rather the hot thermal gas appears to be displaced by the radio emitting plasma – thus indicating no direct heating of the ICM via shocks. Such phenomenological features can be seen e.g. in Fig. 1, representing the superposition of contours of the radio emitting source (at 1.4 GHz) onto the X–ray image of the Perseus cluster (Fabian et al. 2000, 2002).

In the specific case of Perseus the quality of the data has allowed us to infer more robust constraints on the energetics of the jet. In particular, limits on the physical parameters, deduced from both radiative cooling timescales and dynamics of the (plausibly buoyant) radio emitting plasma, have allowed

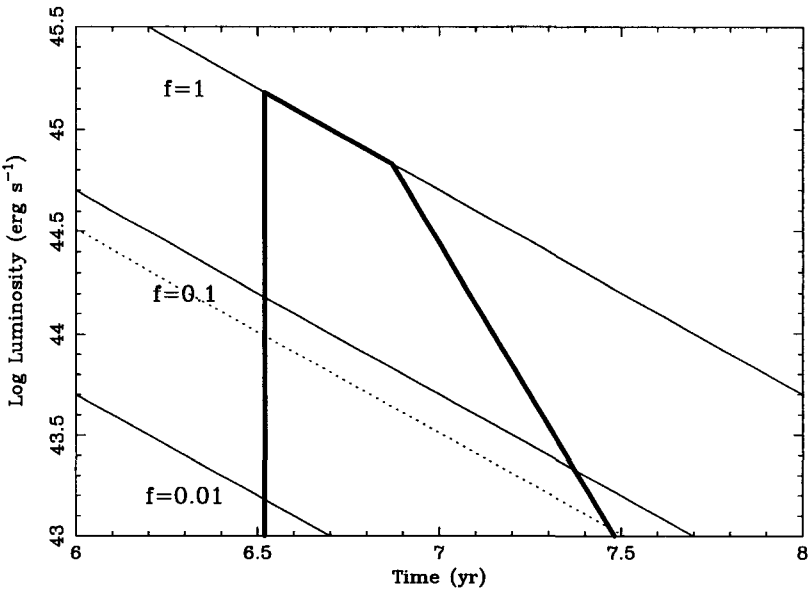


Figure 2: Dynamical constraints on the rate of power supplied to the N radio lobe of 3C 84 as a function of its lifetime. The three lines labelled with the filling factor f derive from the requirement to blow a hole equal in radius to that seen. To the left of the vertical line the hole expands at a velocity faster than the sound speed in the rims (no shocks are observed). To the right of the steepest diagonal line a buoyant bubble will rise faster than it expands, leading to it appearing detached from the center, contrary to observations. Finally the dotted line represents the minimum luminosity required to move gas to make the hole, $P_{thermal}V/t$ for $f = 1$ (from Fabian et al. 2002).

to conclude that: i) the relativistic plasma is not in a condition of minimum energy, ii) the filling factor of the non–thermal plasma is at least of order 10 % and iii) the energy in particles not directly 'observable' (via their synchrotron emission) amounts to a factor 10–100 more than the simplest minimal estimates (see Fig. 2 and its caption and Fabian et al. 2002). Note that the last two quantities, namely the radio plasma filling factor and the total energetics in particles, have been so far the major sources of uncertainty in the estimates of the energy accumulated in the lobes of radio galaxies over their lifetime, and thus of the power channelled into jets from the central engine.

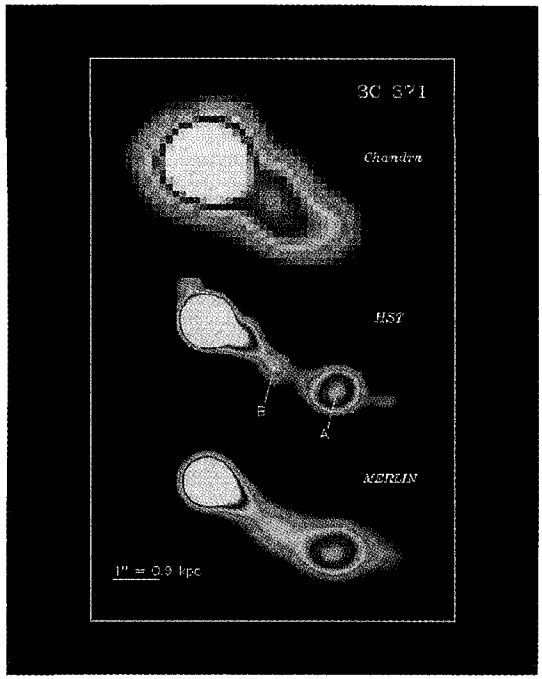


Figure 3: Multiwavelength images of the jet of 3C371. From top to bottom: Chandra ACIS-S (0.48 keV), HST WFPC2 F555W and MERLIN 1.4 GHz. The X-ray image was smoothed with a Gaussian achieving a resolution of 0"86 FWHM (from Pesce et al 2001).

3 Multi–frequency emission from large scale jets

Among the exciting results from Chandra is also the detection with unprecedented quality and resolution of X–ray emission from large scale jets (e.g. Schwartz et al. 2000, Chartas et al. 2000, Pesce et al. 2001, Harris & Krawczynski 2002). Interestingly such emission has been detected both in radio galaxies and quasars, i.e. sources which are believed (according to unification schemes) to be typically observed at significantly different angles with respect to the jet direction.

The high spatial resolution in X–ray has also allowed to study as never before any morphological difference in the emission at different frequencies (typically VLA, HST and Chandra have roughly a similar angular resolution), often revealing a decrease in the X-ray emission corresponding to an increase in the radio one with increasing distance from the central source, as well as

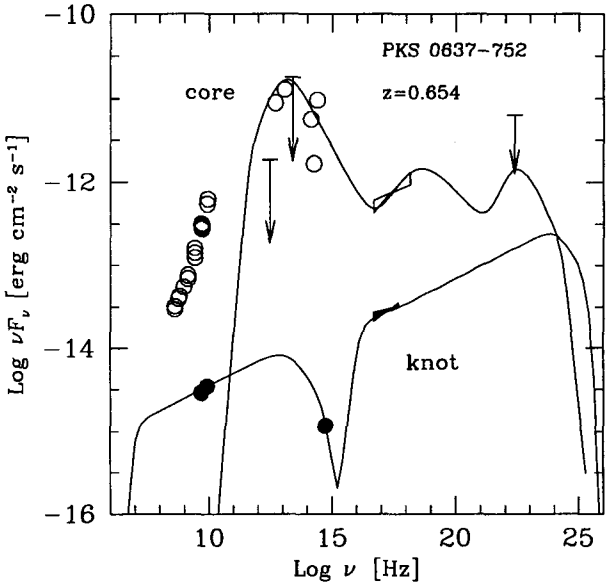


Figure 4: The Spectral Energy Distributions of the core and extended 7.8" knot of the blazar PKS 0637-752. The models (represented by solid lines) are calculated as detailed in Celotti et al. (2001).

displacements in the peak emission (of the same 'knot') at different frequencies (see e.g. Fig. 3, which reports the illustrative case of 3C 371, from Pesce et al. 2001).

Although the interpretation of the origin of the X-ray emission and its relation to that in other spectral bands are still open to discussion, it has been suggested that the bulk of such radiation is produced via inverse Compton scattering of the synchrotron emitting particles onto the microwave background soft photon field (e.g. Celotti et al. 2001; Harris & Krawczynski 2001). If this is the case two significant implications follow. a) The emitting plasma has to be still propagating at highly relativistic speeds even on such large scales and, more importantly, b) it is possible to infer a rather robust estimate of the kinetic power associated with the emitting particles. In fact, contrary to the analogous emission process likely occurring on sub-pc jet scales and responsible for the high energy emission in blazars (e.g. Sikora et al. 1994), in the case of

the 3 K seed photons both the external radiation energy density and the photon energy are precisely known. Therefore also the number of emitting particles responsible for the observed X–ray intensity is well constrained. As an example in Fig. 4 the spectral energy distributions of the quasar core and extended emission in PKS 0637-752 are shown together with the corresponding spectral models. The kinetic powers thus inferred are (intriguingly) consistently similar on the core and extended spatial scales.

4 Extended emission vs nuclear properties of radio galaxies

Observations by HST have also provided information on the connection between accreting flows and jets. Let us consider two aspects.

4.1 The core emission in radio galaxies

HST images have revealed the presence of a compact (unresolved at HST resolution) core in the center of a large fraction of low and high power radio galaxies (belonging to complete samples). In FR I sources the corresponding (optical) luminosity appears to correlate with the radio core one, suggesting that also the former one has a non-thermal origin and is likely associated with the emission from relativistic jets. For higher power sources (FR II) the situation is more complex, but the characteristic optical core intensity well relate to the properties of the optical spectra (more precisely to the typology of any optical emission line). In particular, above a luminosity of order $few \times 10^{42}$ erg s^{-1}, the optical core emission well exceeds the non–thermal level estimated from the radio core, indicating that a different radiation process might start dominating. For conciseness here we only stress a couple of implications of such findings: a) in low power objects there is evidence for the lack of obscuring material (e.g. in the form of a 'torus') with any significant covering factor; b) the excess emission in powerful FR II with respect to the non–thermal intensity might be due to the radiating accretion flow: if so, this appears to start contributing significantly to the emission above a certain power. For a black hole of $10^8 M_\odot$, an accretion flow with a radiative efficiency $\epsilon \sim 10\%$, and a 'suitable' bolometric correction, such a 'threshold' corresponds to an accretion rate in Eddington units $\dot{m} \sim 5 \times 10^{-2} M_8^{-1} \epsilon_{-1}^{-1}$ (Chiaberge et al. 2000, 2002).

4.2 The FR I/FR II power dichotomy

Among the strongest clues about the origin of the jet phenomenon is the finding by Fanaroff & Riley (1974) that most radio galaxies can be morphologically

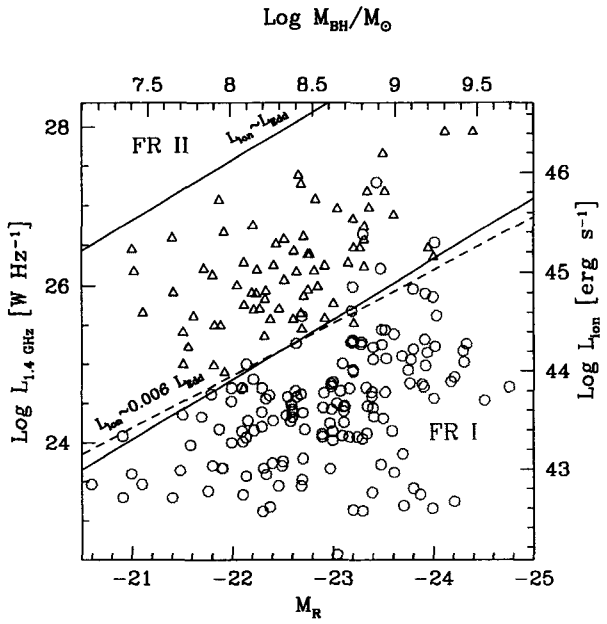


Figure 5: The radio jet power vs host magnitude plane with the (dashed) line sharply dividing FR I (circles) from FR II (triangles) radio galaxies (from Ledlow & Owen 1996). According to reasonably well established correlations this plot is equivalent to an accretion power vs black hole mass plane (as labelled on the right and upper axis, respectively). The central diagonal line corresponds to $L_{acc} \sim 6 \times 10^{-3} L_{Edd}$ (from Ghisellini & Celotti 2001).

classified in FR I or FR II and that this duality corresponds to a separation in radio power. Such division has been further sharpened by the findings of Ledlow & Owen (1996) that the critical value of radio power increases with increasing (optical) luminosity of the host galaxy.

However recently evidence has been accumulating on the association of the mass of black holes in the center of galaxies with both the luminosity in their bulge components (Kormendy & Richstone 1995, Magorrian et al. 1998) and the galaxy velocity dispersion (Ferrarese & Merritt 2000, see the contribution by Laura Ferrarese in these proc.). Furthermore independent indications show the existence of correlations between the observed large scale radio luminosity and both the intrinsic jet power and the radiation dissipated by the accreting flow (e.g. Willott et al. 1999 and references therein).

By combining these two sets of information (adopting the quantitative relations reported by the above authors), the dividing line between FR I and FR II can be re-express as a dividing line in the plane accretion luminosity (L_{acc}) vs black hole mass (see Fig. 5). In particular the dependence on black hole mass (i.e. on the luminosity of the host bulge) translates into a fixed value of the accretion rate (in Eddington units): the FR I/FR II dichotomy corresponds to sources accreting respectively below or above an inflow rate $\dot{m} \sim 6 \times 10^{-2} \epsilon_{-1}^{-1}$ (Ghisellini & Celotti 2001).

To conclude we stress the similar values of the accretion rate $\dot{m}$ inferred from the direct HST nuclear luminosity (see Section 4.1) and from the present correlations which are based on large scale properties. Obvious and tentative speculations on these findings might relate their origin to the existence of flows with different radiation efficiency, as in the (few) scenarios re–proposed and discussed in the last few years (see e.g. Menou's contribution to these proceedings).

Acknowledgements

It is a pleasure to thank the Organizers for inviting me to such an interesting meeting and for their very kind and generous hospitality.

References

1. H. Böringer *et al*, MNRAS **264**, L25 (1993)
2. A. Celotti *et al*, MNRAS **321**, L1 (2001)
3. G. Chartas *et al*, ApJ **542**, 655 (2000)
4. M. Chiaberge *et al*, A&A **355**, 873 (2000)
5. M. Chiaberge *et al*, A&A , (2002) submitted
6. A. C. Fabian *et al*, MNRAS **318**, L65 (2000)
7. A. C. Fabian *et al*, MNRAS , (2002) in press
8. B. L. Fanaroff and J. M. Riley, MNRAS **167**, 31p (1974)
9. L. Ferrarese and D. Merritt, ApJ **539**, L9 (2000)
10. W. Forman *et al*, astro-ph/0111526 (2001)
11. G. Ghisellini and A. Celotti, A&A **379**, L1 (2001)
12. D. E. Harris and H. Krawczynski, astro-ph/0109523 (2001)
13. J. Kormendy and D. Richstone, ARAA **33**, 581 (1995)
14. M. J. Ledlow and F. N. Owen, AJ **112**, 9 (1996)
15. J. Magorrian *et al*, ApJ **AJ**, 115 (2285)1998
16. J. E. Pesce *et al*, ApJ **556**, L79 (2001)
17. D. A. Schwartz *et al*, ApJ **540**, L69 (2000)
18. M. Sikora *et al*, ApJ **421**, 153 (1994)

19. C. J. Willott *et al*, MNRAS **309**, 1017 (1999)

MAGNETIC FIELD GENERATION IN ACCRETION DISKS

ETHAN T. VISHNIAC
Dept. of Physics and Astronomy, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218, USA
E-mail: ethan@pha.jhu.edu

In its usual form, mean-field dynamo theory ignores the constraints imposed by the conservation of magnetic helicity in conducting fluids. Here I discuss how dynamo theory can be modified to account for this important conservation law, and compare this new theoretical framework with numerical simulations of accretion disks. Among the implications of this work are that differential rotation and the magneto-rotational instability are sufficient to account for an efficient dynamo, so that all strongly ionized accretion disks will support dynamo activity. I conclude by discussing the effects of radiation pressure on the dynamo process and possibility of a relationship between the internal dynamo and the luminosity of the disk jet or wind.

1 Introduction

The radial transport of angular momentum in accretion disks occupies a role similar to that of nuclear fusion in stars. That is, it explains the generation of heat, which ultimately leads to the emission of radiation from the surface, and therefore plays a major role in determining the internal structure of accretion disks. However, angular momentum transport is unlike nuclear fusion in that it cannot be understood in terms of microphysical processes. In fact, if we attribute angular momentum transport to viscosity, then we find that particle viscosity fails to explain accretion disk phenomenology by perhaps ten or twelve orders of magnitude. For this reason it has been generally assumed that accretion disks transport angular momentum via some sort of collective process, which is usually referred to as 'turbulence', although magnetic effects have usually been subsumed in this term. The standard way of dealing with our ignorance is to characterize the transport process with a 'viscosity' which is defined as $\alpha_{SS} c_s H$, where c_s is the sound speed, H is the disk thickness, and α_{SS} is a number of order unity. A local definition is obtained by assuming that the viscous dissipation scales with the pressure.

Over the last ten years it has become generally clear that the most likely mechanism for angular momentum transport is a local instability of a magnetic field embedded in a differentially rotating flow [1,2], the 'magnetorotational' instability. The first direct application of this instability to accretion disks can be found in a pair of papers by Balbus and Hawley [3,4] which give the analytic theory for axisymmetric instabilities of a vertical field threading an

accretion disk, and provide a simplified numerical simulation of their effects. A tremendous amount of effort has gone into extending this work to cover the more realistic case of non-axisymmetric modes of an aximuthal magnetic field and to providing fully three dimensional simulations of accretion disks with vertical gravity. Here I will simply note some analytic results and generic features of these simulations.

First, the linear theory tells us that the fastest growing modes have a growth rate which is a large fraction of the rotation frequency, $\Omega(r) = (GM/r^3)^{1/2}$, where M is the mass of the accreting object. Typical wavelengths for the modes are $\sim V_A/\Omega$, where $V_A = B/(4\pi\rho)^{1/2}$ is the Alfvén velocity of the fluid. The modes are not isotropic, but should have $k_\phi < k_r < k_z$, in cylindrical coordinates. The average angular momentum transport is some fraction of V_A^2, and points outward. These analytic results are consistent with the results of the numerical simulations, in the sense that the turbulent eddies seen in these simulations are elongated azimuthally, and flattened vertically, with a correlation time which is on the order of an orbital period.

Second, the simulations seem to converge on a state with on average magnetic pressure which is some large fraction of the average gas pressure. This field is turbulent, but shows significant power on the largest scales in the simulations. To put it another way, the turbulence induced by the magnetorotational instability leads to the spontaneous appearance of a strong azimuthally symmetric field, along with the expected eddies. It is one of the few cases in which numerical simulations have produced a large scale dynamo, i.e. a magnetic field with a large fractional amplitude on scales much larger than the typical eddy scale. It is unique in that these simulations were not guided by any theoretical expectations regarding dynamos, and do not appear, at first glance, to satisfy the conditions for dynamo activity. In particular, conventional dynamo theory requires symmetry breaking in all three directions for a successful dynamo, but dynamo effects appear even in simulations with no vertical disk structure. On the other hand, the simulations show nonlocal effects suggesting the operation of some sort of mean field dynamo process. (For example, the time-averaged magnetic pressure is not a fixed fraction of the local gas pressure, but varies systematically as a function of height in the disk.)

In order to understand how accretion disks work, we need to understand the process that determines the amplitude of the large scale magnetic field in accretion disks. This should be a particular example of the dynamo process that produces the large scale magnetic fields of stars and galaxies. It differs from the dynamo process in planets in that strongly ionized astrophysical ob-

jects typically have negligible resistivity. That is, $\eta \ll V_A L$, where L is the size of the object. The general theory of astrophysical dynamos has been studied for several decades, but its fundamental principles have been severely criticized over the last ten years. Here I will discuss a new way of understanding dynamo theory, with particular reference to accretion disks. A more complete discussion of the fundamental issues can be found in Vishniac and Cho (2001)[5]. I will start by briefly summarizing the standard approach to mean-field dynamo theory.

2 Mean-Field Dynamo Theory

The evolution of a magnetic field in a resistive fluid is governed by the induction equation, i.e.

$$\partial_t \mathbf{B} = \nabla \times (\mathbf{v} \times \mathbf{B}) - \eta \nabla \times \nabla \times \mathbf{B}. \tag{1}$$

In order to follow the evolution of the large scale magnetic field, we write

$$\mathbf{B} \equiv \langle \mathbf{B} \rangle + \mathbf{b}. \tag{2}$$

The brackets here denote averaging over scales somewhat larger than the turbulent eddy size. In other words, they indicate a smoothing process which averages out all small scale features. The field $\langle \mathbf{B} \rangle$ is the 'mean field', whose dynamics are subject of this paper.

It is tempting to set $\eta = 0$, but this must be done with some caution. With $\eta = 0$ equation (1) implies that field lines are frozen in to the plasma, so that the field topology can never change. This presents major difficulties for *any* dynamo theory, and also is inconsistent with observations of reconnection in the solar corona[6,7] which indicate that under some circumstances field topology changes ("reconnection") can propagate at a speed close to V_A. Here we will assume that magnetic reconnection occurs at a speed close to V_A, but only on two dimensional surfaces. This is the minimum assumption required to explain observations, while more efficient reconnection would make it hard to explain the persistence of complicated structures in the solar corona. It is also consistent with at least some models of rapid reconnection[8].

Adopting this philosophy, how can we understand the dynamo process? The conventional way (see, for example, Moffatt 1978[9]) is to imagine that we have a set of large scale parallel field lines pointing in some arbitrary direction. If the underlying turbulence has a tendency to twist the field lines into spirals with a preferred handedness (helicity), then reconnection on two dimensional surfaces between adjacent spirals will produce a new field, at right angles to the old one, provided that there is a systematic gradient in the strength of the spirals. The new field component is at right angles to both the gradient

and the old field component. In a differentially rotating system, we can get a dynamo if the original field direction is in the $\hat{\phi}$ direction, and the dynamo process produces a radial field component. Differential shearing of B_r will then drive the azimuthal field component, closing the cycle. This is the '$\alpha - \Omega$ dynamo'. In the absence of global shear, we need a second round of dynamo action, which gives an α^2 dynamo.

In mathematical terms, this follows from approximating the evolution of the the small scale field component, $\mathbf{b}$, as

$$\partial_t \mathbf{b} \approx \nabla \times \mathbf{v} \times \langle \mathbf{B} \rangle, \tag{3}$$

and substituting the result into the evolution equation for the large scale field,

$$\partial_t \langle \mathbf{B} \rangle = \nabla \times \langle \mathbf{v} \times \mathbf{b} \rangle. \tag{4}$$

In a turbulent, incompressible and homogeneous plasma this implies

$$\partial_t \langle \mathbf{B} \rangle = \nabla \times (\alpha \cdot \mathbf{B}) + \nabla \cdot (\mathbf{D_T} \cdot \nabla) \langle \mathbf{B} \rangle. \tag{5}$$

Here α, the kinetic helicity, and $\mathbf{D_T}$, the turbulent diffusion tensor, are dyads given by

$$\alpha_{il} \equiv \epsilon_{ijk} \langle v_j \partial_l v_k - v_k \partial_l v_j \rangle \tau_c, \tag{6}$$

and

$$D_{T,ij} \equiv \langle v_i v_j \rangle \tau_c, \tag{7}$$

where τ_c is the eddy correlation time. The component of the electromotive force along the large scale field direction, $\langle \hat{\mathbf{B}} \rangle \cdot \langle \mathbf{v} \times \mathbf{b} \rangle$, is the piece that drives an increase in the volume-integrated magnetic field energy. In symmetric turbulence α vanishes, but $\mathbf{D_T}$ does not. In fact, since a successful dynamo requires non-vanishing diagonal components for α, we can see from this expression that a successful dynamo should require symmetry breaking in all three directions.

Accretion disks are strongly shearing systems, so we expect them to work according to an $\alpha - \Omega$ dynamo. For approximately isotropic turbulence this is

$$\partial_t B_r \approx -\partial_z (\alpha_{\phi\phi} B_\phi) + \partial_z D_T \partial_z B_r, \tag{8}$$

$$\partial_t B_\phi \approx -\frac{3}{2} \Omega B_r + \partial_z D_T \partial_z B_\phi. \tag{9}$$

Here we have retained only the vertical structure of the field, so that D_T can be taken as $\langle v_z^2 \rangle \tau_c$. When this gives a successful dynamo, it implies an exponential growth rate, Γ_d, which is roughly the geometric mean between the shear and $\alpha_{\phi\phi}/L_z$, where L_z is the vertical scale height of the field.

Despite the straightforward nature of this derivation, it doesn't actually give a useable theory of astrophysical dynamos. Simulations of the magnetorotational instability show dynamo effects even in the absence of any vertical symmetry breaking, when we would expect $\alpha_{\phi\phi} = 0$. Somewhat more troubling is the fact that attempts to simulate α^2 dynamos produce unimpressive results. In fact, such systems show dynamo growth times which increase sharply with decreasing effective resistivity, implying that the time required to grow a magnetic field in a realistic astrophysical system would be absurdly long [10,11]. This is consistent with theoretical arguments which show that the mean-field dynamo theory described above neglects important effects, which act to suppress dynamo action on time scales shorter than the resistive time, L^2/η [12,13,14].

We can get some idea of the nature of this problem by considering the first step in the mean-field dynamo process, twisting field lines into spirals. This is intuitively appealing if we consider field lines as isolated strings of infinitesimal radius. More realistically, the field occupies a non-zero volume. Twisting a tube into a spiral shape requires that we either allow the ends to slip, or allow parts of the tube to twist in the opposite sense. There is a geometrical constraint which is ignored in the standard picture.

3 Magnetic Helicity Conservation and Dynamo Theory

3.1 Magnetic Helicity and the Electromotive Force

A simple mathematical description of the problem with dynamo theory follows from considering the magnetic helicity

$$H \equiv \mathbf{A} \cdot \mathbf{B} = \mathbf{A} \cdot (\nabla \times \mathbf{A}). \tag{10}$$

Since $\mathbf{A}$ is gauge dependent, so is H. Here I will use the coulomb gauge

$$\nabla \cdot \mathbf{A} = 0. \tag{11}$$

Any gauge choice gives an H which is a conserved quantity in the limit $\eta = 0$, but some choices are useful and lead to clear predictions for dynamo theory (always equivalent to the results we get from the Coulomb gauge), and the rest give no clear predictions at all. The physically interesting thing about H is that it is not only a conserved quantity, but one that carries topological information. In particular, a spiral magnetic field line has an associated magnetic helicity.

For any gauge choice we have

$$\partial_t H = -\nabla \cdot [\mathbf{A} \times (\mathbf{v} \times \mathbf{B} + \nabla \Phi)] - \eta \mathbf{B} \cdot \nabla \times \mathbf{B}, \tag{12}$$

where Φ is an arbitrary function of space and time. However, for the Coulomb gauge we require

$$\nabla^2\Phi = \nabla\cdot(\mathbf{v}\times\mathbf{B}). \tag{13}$$

Since H is of order LB^2, where L is a characteristic scale of the field, it takes less energy to hold magnetic helicity on large scales than on eddy scales, and a divergent amount on infinitesimal scales. In the limit of vanishing resistivity the resistive term in equation (12) is unimportant for any reasonable model of MHD turbulence. It can even be neglected in the presence of fast reconnection, as long as reconnection only occurs in an infinitesimal fraction of the plasma volume. Consequently, magnetic helicity conservation is very good for small resistivity, a point that was originally stressed by Taylor [15] in connection with laboratory plasmas.

How does this affect dynamo theory? The large scale distribution of magnetic helicity can be divided into a piece carried by large scale magnetic structures and a piece carried by small scale structures, or

$$\langle H\rangle = \langle\mathbf{A}\rangle\cdot\langle\mathbf{B}\rangle + \langle\mathbf{a}\cdot\mathbf{b}\rangle. \tag{14}$$

Henceforth we will use $h \equiv \langle\mathbf{a}\cdot\mathbf{b}\rangle$. The evolution of the first piece, in a perfectly conducting fluid, is

$$\partial_t(\langle\mathbf{A}\rangle\cdot\langle\mathbf{B}\rangle) = 2\langle\mathbf{B}\rangle\cdot\langle\mathbf{v}\times\mathbf{b}\rangle - \nabla\cdot[\langle\mathbf{A}\rangle\times(\langle\mathbf{v}\times\mathbf{b}\rangle + \nabla\langle\Phi\rangle)]. \tag{15}$$

The second term on the right hand side is the magnetic helicity transport driven by mean-field terms. The first represents the exchange of magnetic helicity between large and small scales. This term is proportional to the component of the electromotive force which drives the dynamo process. In other words, the generation of a large scale magnetic field is a direct consequence of the transfer of magnetic helicity between large and small scales.

We can estimate the rate at which h is transferred to large scale magnetic field structures by considering its role in biasing the value of the electromotive force. The inverse cascade rate is

$$\tau_{cascade}^{-1} \sim \frac{V_A^2}{\langle v^2\rangle}\tau_c^{-1}. \tag{16}$$

For an a large scale magnetic field in equipartition with the turbulent cascade this implies that magnetic helicity is transferred to the large scale field in one eddy turn over time. This suggests that unless the large scale field is very weak it is reasonable to take

$$H \approx \langle\mathbf{A}\rangle\cdot\langle\mathbf{B}\rangle. \tag{17}$$

Then combining equations (15) and (17) we see that

$$2\langle \mathbf{B} \rangle \cdot \langle \mathbf{v} \times \mathbf{b} \rangle = -\nabla \cdot [\langle \mathbf{a} \times (\mathbf{v} \times \mathbf{B} + \nabla \phi) \rangle] \equiv -\nabla \cdot \mathbf{J}_H, \tag{18}$$

where J_H is defined as the magnetic helicity current carried by small scale structures, or the anomalous magnetic helicity current. That is, the component of the electromotive force parallel to the large scale magnetic field is given by the divergence of the magnetic helicity current carried by eddy scale structures. Equations (4) and (18) yield

$$\partial_t \langle \mathbf{B} \rangle = \nabla \times \left[\frac{-\langle \mathbf{B} \rangle}{2\langle B \rangle^2} \nabla \cdot \mathbf{J}_H + \langle \mathbf{v} \times \mathbf{b} \rangle_\perp \right], \tag{19}$$

where the second term on the right hand side is the component of the electromotive force perpendicular to the large scale field direction. Evaluating $\mathbf{J}_H$ is necessary to understand the dynamo process. By contrast, attempts to estimate the kinetic helicity only tell us about the dynamo process when the large scale magnetic field is so weak that the transfer of magnetic helicity between scales is unaffected by the extremely limited capacity of the turbulent eddies to store magnetic helicity.

3.2 The Anomalous Magnetic Helicity Current

The most direct way to estimate the anomalous magnetic helicity current is to write $\mathbf{a}$ in terms of the action of the turbulent velocity field on the large scale magnetic field, or

$$\mathbf{a} \approx (\mathbf{v} \times \mathbf{B} - \nabla \phi) \tau_c, \tag{20}$$

where

$$\nabla^2 \phi = \nabla \cdot (\mathbf{v} \times \mathbf{B}). \tag{21}$$

If we substitute this into the definition of the magnetic helicity current we find, after some manipulation, that

$$\mathbf{J}_H = -\tau_c \int \frac{d^3\mathbf{r}}{4\pi r} \epsilon_{lmn} \langle B_k \rangle \langle B_l \rangle \partial_k \partial_m \langle v_i(\mathbf{x}) v_n(\mathbf{x} + \mathbf{r}) \rangle. \tag{22}$$

We see that $\mathbf{J}_H$ is parity-invariant, unlike α. In completely isotropic turbulence it will also vanish, but the degree of symmetry breaking necessary for a dynamo effect is smaller than in the conventional picture.

Equation (22) is not a particularly enlightening expression, but we can gain somewhat more insight by rewriting it as

$$\mathbf{J}_H \approx -\lambda_c^2 \tau_c \langle (\mathbf{B}) \cdot \omega ((\mathbf{B}) \cdot \nabla) \mathbf{v} \rangle, \tag{23}$$

where λ_c is some suitably averaged eddy size and $\omega \equiv \nabla \times \mathbf{v}$ is the fluid vorticity. This corresponds to twisting a field line in both directions, but then systematically moving right and left handed spiral segments in opposite directions.

In this model, we generate left (or right) handed spirals by separating segments of the same field line with different helicities, moving them in opposite directions, and then reconnecting them (along two dimensional) surfaces, into new field lines. If the flow of magnetic helicity has a non-zero divergence, then the new field lines will have a preferred sense of twisting, and the first step in the usual cartoon for the dynamo process will have been completed without violating magnetic helicity conservation.

3.3 The Dynamo in Accretion Disks

What does this mean for accretion disks? Equation (8) for the $\alpha - \Omega$ dynamo becomes

$$\partial_t B_r \approx -\partial_z B_\phi^{-1} \partial_z \left(\frac{\lambda_c^2}{2} B_\phi^2 \langle \omega_\phi \frac{1}{r} \partial_\phi v_z \rangle \tau_c \right) + \partial_z D_T \partial_z B_r. \tag{24}$$

We see that the success, or failure, of the dynamo depends entirely on the sign of the correlation $\langle \omega_\phi \partial_\phi v_z \rangle$. Combining equations (9) and (24) we see that a successful dynamo in a simple vertically periodic system requires

$$\frac{1}{r} \langle \omega_\phi \partial_\phi v_z \rangle < 0. \tag{25}$$

Integrating this expression by parts, invoking approximate incompressibility, and characterizing spatial derivatives by wavenumbers, we find that

$$\frac{1}{r} \langle \omega_\phi \partial_\phi v_z \rangle \approx -k_r k_\phi \langle v_r^2 + v_z^2 \rangle - k_\phi^2 \langle v_r v_\phi \rangle. \tag{26}$$

In a strongly shearing system eddies tend to assume a spiral shape and the sign of $k_\phi k_r$ reflects the sign of $\partial_r \Omega$. For an accretion disk, where the angular velocity decreases outward, we have $k_r k_\phi > 0$. In systems where the angular velocity increases outward we would obtain the opposite sign, but the condition for a successful dynamo would also switch sign. The first term on the right hand side of equation (26) always helps drive a successful dynamo. The sign of the second term depends only on the sense of radial angular momentum transport mediated by Reynolds stresses. For the magnetorotational instability this is positive, and this term also contributes to a successful dynamo. We

conclude that the turbulence driven by the magnetorotational instability always generates an anomalous magnetic helicity current with the sign required to drive a successful dynamo.

Oddly enough, in this simple model the large scale field has no net magnetic helicity. Instead the magnetic helicity current is constant, with J_H carrying a fraction which varies periodically from 0 to 1. This current is positive and must eventually leave the system.

We can gain some further insight into the nature of this model by considering how it works in different systems. Suppose we consider instead a simple shearing system, like an accretion disk but without the centrifugal force. This system can generate turbulence via the Kelvin-Helmholtz instability, but in such a system the sense of momentum transport is different. The turbulence will move momentum in such a way as to reduce the shear. In other words, $\langle v_r v_\phi \rangle$ will change signs and a successful dynamo is no longer guaranteed. In fact, depending on the relative size of the two terms in equation (26), the system can act as an anti-dynamo, suppressing magnetic fields on all scales dominated by shear. This simulation has been performed[16] with precisely this result.

Suppose we reversed the sign of $\partial_r \Omega(r)$? This is obviously not meant to be a realistic model of an accretion disk, but may be applicable to other systems. For example, the Sun has $\partial_r \Omega(r) > 0$ throughout the convection zone. The Sun is poorly modeled as a cylindrical system, but ignoring that point, we note that this model suggests that there should be global magnetic helicity current in the $-\hat{z}$ direction. In fact, this is seen, in the sense that the Sun is ejecting positive magnetic helicity from its southern polar cap and negative magnetic helicity from its northern polar cap throughout the solar cycle.

4 Limits, Problems, and Speculations

This approach to dynamo theory is idealized in several ways. Here we point out a few of the questions raised by this line of work.

4.1 Magnetic Field Fluctuations

In the conventional approach to dynamo theory, the electromotive force, $\langle \mathbf{v} \times \mathbf{b} \rangle$, is evaluated by substituting for $\mathbf{b}$ an expression from the induction equation. Another approach is to take

$$\mathbf{v} \approx \frac{(\mathbf{B} \cdot \nabla)\mathbf{b}}{4\pi\rho} \tau_c. \tag{27}$$

We note that this expression is *not* appropriate in a strongly sheared plasma. However, when this substitution can be made the resulting expression is usually added to the previous one, on the theory that in a turbulent plasma the two estimators are independent and cumulative. This means that the electromotive force becomes

$$\langle \mathbf{v} \times \mathbf{b} \rangle_i \approx \epsilon_{ijk} \left(\langle v_j \partial_l v_k \rangle - \frac{1}{4\pi\rho} \langle b_j \partial_l b_k \rangle \right) \tau_c B_l. \tag{28}$$

A similar argument can be advanced for our expression for the anomalous magnetic helicity current, in which case we should replace equation (22) with

$$\mathbf{J_H}(\mathbf{x}) = \int \frac{d^3r}{4\pi r} \left[\langle (\mathbf{B} \cdot \nabla \mathbf{v}(\mathbf{x}) \mathbf{B} \cdot \omega(\mathbf{x}+\mathbf{r}) \rangle - \frac{1}{4\pi\rho} \langle (\mathbf{B} \cdot \nabla \mathbf{b}(\mathbf{x}) \mathbf{B} \cdot \mathbf{j}(\mathbf{x}+\mathbf{r}) \rangle \right]. \tag{29}$$

However, since accretion disks are strongly sheared, in the sense that $\Omega(r)\tau_c > 1$, this expression is not clearly better than equation (22). A direct calculation of the anomalous magnetic helicity current induced by the MRI would be extremely useful.

4.2 Scale Transfer, Seed Fields, and the Conventional Dynamo

The process by which large scale magnetic helicity embedded in small scale structures is moved to large scale structures, i.e.

$$\langle \mathbf{a} \cdot \mathbf{b} \rangle \rightarrow \langle \mathbf{A} \rangle \cdot \langle \mathbf{B} \rangle, \tag{30}$$

is not instantaneous. In order to assume that it is, we need an inverse cascade rate which is larger than the turbulent dissipation rate, or

$$V_A^2 \tau_c k_{eddy}^2 > \frac{\langle v^2 \rangle \tau_c}{L_B^2}, \tag{31}$$

where L_B is the scale length of the large scale magnetic field. Here we have used the *coherent* inverse cascade rate, which depends on the strength of the large scale field. After some manipulation we find that this condition is equivalent to

$$V_A > \langle v^2 \rangle^{1/2} \frac{\lambda_{eddy}}{L_B}. \tag{32}$$

In other words, in order to ensure that large scale fluctuations in the magnetic helicity are transferred to large scale structures, rather than being dissipated by turbulent transport, we need a significant large scale seed field.

An obvious source for a large scale seed field would be the random accumulation of fields in a turbulent medium. Turbulence in conducting fluid is, by itself, sufficient to produce a disordered magnetic field with an energy density roughly comparable to the turbulent energy density and a typical scale comparable to the largest eddy scales. At larger scales the conserved quantity, the magnetic helicity, should show a Poisson distribution, which implies a large wavelength magnetic field scaling as $L_B^{-5/4}$. This is slightly steeper than the seed field required for the dynamo process. Evidently an adequate seed field is *not* guaranteed.

On the other hand, the conventional dynamo process is not completely suppressed, even in the limit of ideal MHD. The eddies can contain a small, but nonzero, amount of helicity, so the kinematic helicity will drive a large scale field until the inverse cascade overwhelms its effects. This limit will be reached when

$$\rho V_A^2 \alpha_k \leq V_A^2 \tau_c k_{eddy}^2 \langle \mathbf{a} \cdot \mathbf{b} \rangle, \tag{33}$$

where α_k is the kinematic helicity. Since in the absence of transport effects the magnetic helicity buried in eddy scale structures must be balanced by the magnetic helicity contained in large scale structures, we have

$$\langle \mathbf{a} \cdot \mathbf{b} \rangle = -\langle \mathbf{A} \cdot \mathbf{B} \rangle \sim \rho V_A^2 L_B. \tag{34}$$

This implies that the limit on the conventional dynamo process can be expressed as an upper limit on the large scale field, i.e.

$$V_A^2 L_B < \alpha_k \langle v^2 \rangle \tau_c. \tag{35}$$

We conclude that the conventional dynamo process can successfully seed a dynamo if

$$\alpha_k > \frac{\langle v^2 \rangle \tau_c}{L_B} \sim v_{rms} \frac{\lambda_{eddy}}{L_B}. \tag{36}$$

The implication is that we should view the production of a strong, large scale magnetic field as a three step process. The first is driven purely by the internal dynamics of the turbulent cascade, and produces disordered field with a weak long wavelength tail on eddy turn over time scales. The second operates according to conventional mean-field dynamo theory and amplifies certain modes in the tail by a modest, but important factor. The third works through the eddy-scale transport of magnetic helicity, and can drive the large scale magnetic field up to equipartition with the turbulent energy density. The second stage of this process is the most restrictive, in terms of the kind of symmetry breaking required. However, it's only the last part of this process brings us to the point of direct comparison with observations of astrophysical magnetic fields.

4.3 The Ejection of Magnetic Helicity

The simple dynamo described in §3.3 has periodic vertical boundary conditions, which automatically avoids the global implications of the magnetic helicity current. What can we say about the operation of this mechanism in a real accretion disk?

The magnetic helicity flux will be of order

$$J_z \sim B^2 \langle v^2 \rangle \tau_c R^2, \tag{37}$$

where R is the disk radius. The minimum corresponding energy flux can be obtained by assuming that the magnetic helicity flux will be carried by magnetic structures of size R, in other words

$$E_z \sim B^2 \langle v^2 \rangle \tau_c R. \tag{38}$$

For an accretion disk where the angular momentum transport is mediated by the MRI, we have

$$B^2 \sim \alpha_{SS} P, \tag{39}$$

$$\tau_c \sim \Omega(r)^{-1}, \tag{40}$$

and

$$\langle v^2 \rangle \sim \alpha_{SS} c_s^2. \tag{41}$$

Here P is the pressure in the disk and c_s is the sound speed. Since the mass flux through the disk is

$$\dot{M} \sim \alpha_{SS} P \frac{c_s}{\Omega^2}, \tag{42}$$

this implies

$$E_z \sim \alpha_{SS} \frac{c_s}{r\Omega(r)} [\dot{M} r^2 \Omega^2]. \tag{43}$$

The last factor is just the total accretion energy, while the ratio of the sound speed to the orbital speed is also the ratio of the disk thickness to its radius.

We see from this that the minimum energy transmitted into the halo is typically a small fraction of the accretion energy. It is also smaller than the energy invested in the dynamo process, by a factor of $c_s/(r\Omega)$. This factor is due to our assumption that the typical scale of magnetic field structures in the halo is r, rather than the disk thickness. The ultimate fate of this energy is uncertain. It is tempting to equate it with a minimum wind or jet energy. I note, however, that in the only global simulations currently available (see J. Krolik in this proceedings) the magnetic helicity current that emerges from the accretion disk makes a return through the middle of the disk simulation.

5 Conclusions and Summary

The work presented here is only an outline of how realistic astrophysical dynamos work. Nevertheless, the following points seem fairly secure.

First, conventional mean-field dynamo theory, which relies on a source of kinematical helicity to drive dynamo activity, is inadequate. At best, it describes an early phase in the evolution of an astrophysical magnetic field, when the field energy density is much smaller than the turbulent energy density. This does not depend on whether or not the large scale magnetic field has a large magnetic helicity density. Kinematic helicity transfers magnetic helicity between scales, and its operation entails a build-up of magnetic helicity on small scales, and a consequent suppression of its ability to continue generating magnetic flux.

Second, magnetic helicity conservation is an important aspect of astrophysical dynamo theory. Numerical simulations that cannot conserve magnetic helicity are not good models for astrophysical plasmas. In a turbulent plasma, this is more important than how kinetic energy is handled. Errors in the kinetic energy density cascade to small scales and are transformed into heat. Errors in magnetic helicity cascade to large scales and affect the evolution of the magnetic field.

Third, the existence of a dynamo in homogeneous turbulence depends on an asymmetry in the turbulence which leads to a systematic flow of magnetic helicity. This depends on a velocity correlation tensor which is invariant under parity reversal. This is strikingly different from conventional mean-field dynamo theory in which parity invariance automatically implies the failure of the dynamo process.

Fourth, the only simple example of a fast dynamo requires turbulence in a differentially rotating system. In this case the outward transport of angular momentum guarantees a successful dynamo. This seems to explain the ubiquity of dynamos in accretion disk simulations, in which the magneto-rotational instability is both the consequence of the existence of a magnetic field, and its cause. The reverse situation, in which angular momentum is transported inward, can be simulated by eliminating coriolis forces. As expected, this produces an *anti-dynamo*, in which all magnetic activity is suppressed.

Fifth, dynamos of this kind should show a global magnetic helicity current aligned with $-[\partial_r \Omega(r)]\hat{z}$. In the case of accretion disks, this implies a strong, vertical magnetic helicity current. The fate of this current, once it has left the accretion disk, may be critical to the formation of any large scale poloidal magnetic field. Current simulations are probably not adequate to address this issue.

In a real astrophysical disk the assumption of local homogeneity is only marginally satisfied, that is, the density and pressure scale heights should be greater than the vertical eddy scale by a factor of several near the disk midplane, but they may be nearly equal several density scale heights above or below the midplane. This implies corrections to the magnetic helicity current that will be quantitatively significant everywhere, and large as $V_A \to c_s$. Ultimately, despite the insights afforded by the analytical approach given here, we will need to rely on numerical simulations for quantitative models of the global structure of disk magnetic fields. However, I am cautiously optimistic that the two approaches can be combined to yield a mean-field theory of disk structure, much as the mixing-length approach to convection allows us to construct useful stellar models without solving the hydrodynamic equations at every time step.

Acknowledgments

I am happy to acknowledge helpful discussions with Amitava Bhattacharjee, Eric Blackman, Jungyeon Cho, Patrick Diamond, Eun-Jin Kim, Julian Krolik and Alex Lazarian. The work presented here has been supported in part by the US National Science Foundation through Grant AST-0098615. Finally, I am pleased to thank the Korea Institute for Advanced Study for its hospitality during my visit to Korea.

References

1. Velikhov, E.P.Soviet Phys. JETP Lett. **35**, 1398 (1959)
2. Chandrasekhar, S. *Hydrodynamic and Magnetohydrodynamic Stability*, (Oxford: Oxford University Press, 1961)
3. Balbus, S.A., & Hawley, J.F.Ap.J. **376**, 214 (1991)
4. Hawley, J.F., & Balbus, S.A. Ap.J. **376**, 223 (1991)
5. Vishniac, E.T. & Cho, J. Ap.J. **550**, 752 (2001)
6. Dere, K.P. Ap.J. **472**, 864 (1996)
7. Innes, D.E., Inhester, B., Axford, W.I., & Wilhelm, K. Nature **386**, 811 (1997)
8. Lazarian, A. & Vishniac, E.T. Ap.J. **517**, 700 (1999)
9. Moffatt, H.K. *Magnetic Field Generation in Electrically Conducting Fluids*, (Cambridge: Cambridge University Press, 1978)
10. Cattaneo, F., & Hughes, D.W.Phys. Rev. E **54**, 4532 (1996)
11. Brandenburg, A.Ap.J. **550**, 824 (2001)
12. Vainshtein, S.I. and Cattaneo, F. Ap.J. **393**, 165 (1992)
13. Gruzinov, A.V. & Diamond, P.D. Phys. Rev. Letters **72**, 1651 (1994)
14. Gruzinov, A.V. & Diamond, P.H. Phys. Plasmas **3**, 1853 (1996)

15. Taylor, J.B. Phys. Rev. Letters **33**, 1139 (1974)
16. Hawley, J.F., Gammie, C.F & Balbus , S.A. Ap.J. **464**, 690 (1996)

RADIATION DRIVEN WIND FROM HOT ACCRETION FLOW

Myeong-Gu PARK
Kyungpook National University, Taegu 702-701, KOREA
E-mail: mgp@knu.ac.kr

Radiation driven wind from hot accretion flow is discussed. Hot accretion flows like ADAF and CDAF are more likely to develop Compton heated outflows compared to the low-temperature accretion flow due to the poloidal density gradient and high radiation temperature. We find that CDAFs are more suitable to produce radiation driven winds than ADAFs. Specific conditions for the existence of outflow and their properties are explored for CDAF.

1 Introduction

1.1 Spherical Accretion and Disk Accretion

Generic accretion flow onto black holes may be classified by its geometrical shape into either spherical (or quasi-spherical) flow or disk-like flow. Accretion flow becomes spherical when the angular momentum of the infalling matter is zero or very small. Due to the geometrical nature of the spherical flow radiation produced at inner part of the flow must go through the infalling matter. Spherical flow generally has a very low radiation efficiency $e \equiv L/\dot{M}c^2 \leq 10^{-7}$: Most of the gravitational energy is retained in the particles in the form of kinetic energy, which particles carry into the black hole, and only a small amount is released to the outside world via radiation. The temperature of the flow may be hot due to the adiabatic compression and inefficient cooling at low density, or cold when the density is large enough to allow efficient cooling. This bifurcation roughly overlaps with the transition from optically thin flow to optically thick (in scattering) flow around $\dot{m} \simeq 1$ where the dimensionless mass accretion rate is defined as $\dot{m} \equiv \dot{M}/\dot{M}_E$ and $M_E \equiv L_E/c^2$ is the Eddington accretion rate. Without consideration for the interaction between outgoing photons and infalling matter, spherical accretion is simply hot and optically thin for $\dot{m} \ll 1$ or cold and optically thick (in absorption) for $\dot{m} > 1$. However, when the radiative interaction is considered, there exists optically thin (in absorption) hot accretion flow with $\dot{m} > 1$ that has higher entropy, maintained by global Compton preheating [1,2]. This branch of accretion has higher radiation efficiency than the corresponding cold accretion, $e \simeq 10^{-4}$, though this is still much smaller than that of disk-like accretion, $e \simeq 0.1$. Spherical accretion is attractive because it can be solved without any free parameters or ad-hoc assumptions, but it has had only limited applications due to the restriction

that the angular momentum of the flow has to be very small.

In contrast, disk-like accretion occurs when the angular momentum is large enough to form a disk outside the horizon. The classic disk flow of Shakura[3] and Shakura and Sunyaev[4] is geometrically thin and optically thick in absorption. The disk is in LTE state and its temperature is low. Its angular momentum is very close to that of the Keplerian rotation and, therefore, gas moves very slowly inward only by the outward viscous transport of the angular momentum. Binding energy of the circular orbit has to be released via radiation, thereby, making the radiation efficiency high, $e \simeq 0.1$. It has been successfully applied to a number of luminous accreting systems.

However, it turns out that disk flow would become geometrically thick when the mass accretion rate is above the Eddington rate, $\dot{m} \gg 1$. Thermal state of this geometrically thick disk is equal to the cool thin disk in which radiation and matter are locked to each other in LTE. But the pressure within the disk is not negligible anymore, and the flow picks up a significant radial velocity from the radial pressure gradient and general relativistic effect (see e.g. Paczyński & Wiita[5]). In thin disk, viscous heating is locally balanced by the radiative cooling through the disk surface. In thick disk a large fraction of the viscous dissipation energy is retained in the gas which advects into the hole as in spherical flow. Now, heating is balanced by local radiative cooling plus advection of thermal energy in particles. Sometimes, thick or rather not-thin disk flow is approximated by vertically integrated thin disk, and this approach is called slim disk[6]. This type of disk flow resembles cool spherical accretion flow.

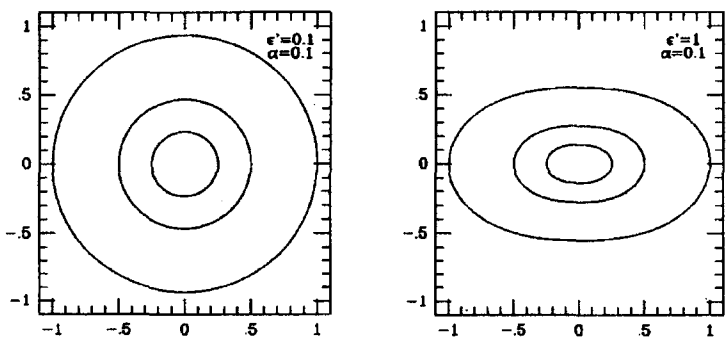


Figure 1: Typical isodensity contours of two-dimensional, self-similar ADAF in the meridional plane (Narayan & Yi 1995a).

1.2 Advection Dominated Accretion Flow

Quite recently, a new type of geometrically thick disk is rediscovered: it is hot and has low radiation efficiency[7,8,9,10]. Ions are kept at near virial temperature while electrons are at much lower temperature due to the inefficient Coulomb coupling. Therefore the disk has to be geometrically thick. As in the cool thick disk, the pressure in the disk produces significant radial motion that makes almost all of viscous dissipation energy radially advected into the hole without releasing much radiation. For these reasons, this type of accretion has been called ion torus or advection-dominated accretion disk (ADAF). For example, Figure 1 shows typical isodensity contours of self-similar ADAFs[11]. They are not disk-like and the flow fills up the whole three-dimensional volume.

Park & Ostriker[12] recognized that this hot disk accretion, specifically ADAF, has substantial resemblance to the hot spherical accretion flow in terms of dynamics, thermal characteristics and radiation transport. Although most works on ADAF have been done on the approximations that it can be treated as vertically integrated thin disk with radiation escaping freely through the surface, ADAF is more like spherical accretion with rotation rather than disk with radial motion. They especially focused on the thermal properties and radiation transport in two-dimensional ADAF solutions and find that the usual properties of ADAF from one-dimensional study apply only to the equatorial part of ADAF. High-temperature state is maintained only in the toroidal region around the rotation axis while the polar region will cool down to low-temperature state. The other, generally neglected, aspect of ADAF is that the geometrical shape of the flow guarantees the interaction of radiation with matter as in spherical accretion flow. Any photons produced at the inner part of ADAF should go through the outer part of the flow. Radiative interaction may or may not affect the dynamical and thermal properties of the flow. In spherical flow, radiative interaction is not important when $\dot{m} \ll 1$ while it fundamentally changes the character of the flow when $\dot{m} \geq 1$[1]. Hence, in ADAF too, the radiative interaction may significantly modify the flow structure.

In subsequent work, Park & Ostriker[13] studied the global thermal nature of ADAF with special consideration given to radiative preheating by Compton scattering. They find that without allowance for Compton preheating, a very restricted domain of ADAF solution is permitted and with Compton preheating included, a new branch, preheated ADAF, appears in the solution space. These solutions correspond to the high-temperature, high $\dot{m}$ spherical accretion flow maintained by preheating. In addition, they also find solutions where the flow near the equatorial plane accretes normally while the flow near the pole is overheated by Compton preheating, probably becoming a polar

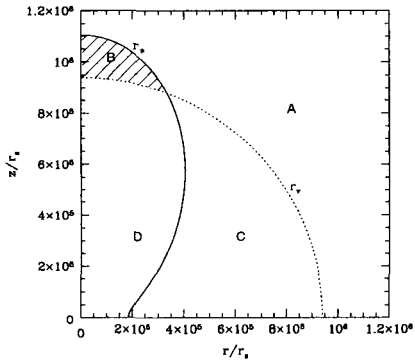


Figure 2: Shaded region B denotes the overheated part of ADAF.

wind. Shaded region in Figure 2 shows the overheated region of ADAF for certain combination of radiation efficiency and radiation temperature[12]. However, these ADAFs with radiation driven wind are allowed only in a limited parameter space.

Accretion flow with outflow is quite desirable. Large fraction of accretion powered sources, for example AGNs whose central engines are powered by accretion onto black holes, show ubiquitous outflows or jets. In disk accretion, outflows are generally linked with geometrically thick disk. One recent example is the possibility of outflow from ADAF: Narayan & Yi [9,11] noted the positivity of Bernoulli parameter in ADAF that can produce outflows driven by hydrodynamics. Blandford & Begelman[14] further suggest that only a small fraction of the gas supplied actually falls into the black hole, and that the binding energy it releases is transported radially outward by the torque so as to drive away the remainder in the form of a wind. However, details of the physical flow structure is not specified in this work: ADAF equations are arbitrarily generalized to produce the specific mass loss at a given radius. So there seems to be more than one ways to produce outflow from hot accretion flow like ADAF albeit with some limitations.

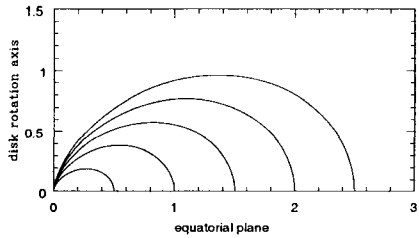


Figure 3: Isodensity contours of self-similar CDAF (Quataert & Gruzinov 2000).

1.3 Convection Dominated Accretion Flow

Convection is another global energy transport mechanism possible in three-dimensional flow. Analytic studies of ADAF showed the convectively unstable nature of the flow[11] (see Begelman & Meier[15] for radiation-dominated ADAF). This has been subsequently proved in a series of multi-dimensional numerical simulations[16,17,18,19,20]. Especially, the numerical studies of ADAF by Igumenshchev & Abramowicz[17,18] show that ADAF becomes convection-dominated accretion flow (CDAF) whenever the viscosity parameter $\alpha \leq 0.1$.

Further analyses of self-similar CDAF show the unique properties of CDAF[21,22]: Density varies as $\rho \propto r^{1/2}$ (r is the radius in unit of the Schwarzschild radius, $R_S \equiv 2GM/c^2 = 3.0 \times 10^5 m$ cm, and m is the black hole mass in solar unit), which is much flatter than the usual $r^{-3/2}$ in ADAF or spherical accretion flow. From the conservation of mass, the mean radial velocity, therefore, varies as $v \propto r^{-3/2}$, compared to $r^{-1/2}$ in ADAF or in the spherical flow. Another distinct characteristic of CDAF is that the energy produced at the inner hot part of the flow is mainly transported outward by convection. Hence, CDAF has as much resemblance to the rotating stellar envelope as to the conventional accretion flow.

The total Luminosity radiated from the flow would be whole or part of the energy transported by convection, which we may write as $L_c \equiv e_c \dot{M} c^2$. The typical convective "efficiency" $e_c \sim 10^{-2} - 10^{-3}$, thus making CDAF relatively high efficiency accretion flow compared to ADAF or spherical accretion. The temperature structure of CDAF is basically equal to that of ADAF: ions at near virial temperature and electron temperature flattening at relativistic regime.

The two-dimensional density structure of self-similar CDAF in the meridional plane is shown in Figure 3. The isodensity contours clearly show that the density gradient from the equatorial plane to polar axis is much steeper than that in ADAF. Coupled with higher radiation efficiency and higher temperature this makes CDAF prone to produce outflows than any other accretion flow.

So in this work, we explore the specific conditions in the parameter space for developing radiation driven outflows in CDAF. We consider only the non-magnetic CDAF.

2 Flow Properties

2.1 Density and Temperature

Multi-dimensional numerical simulations and analytic analyses show that density and temperature profiles of CDAF follow self-similar form in radius[21,22,23]:

$$\rho(r) \propto r^{-1/2}; \quad T_i(r) \propto r^{-1}. \tag{1}$$

Quataert & Gruzinov[22] further show that since CDAF should stay marginally stable to convection, the density should follow a power law in $\sin\theta$ while the temperature is constant on spherical shells. So we assume that the density is given by

$$\rho(r,\theta) = \rho_0 r^{-1/2} \sin^2\theta \tag{2}$$

$$= \rho_{out}(r/r_{out})^{-1/2} \sin^2\theta. \tag{3}$$

The flow extends from r_{out} down to $r_{in} = 1$.

The electron temperature, T_e, in accretion flow is determined by the balance between various cooling and heating processes. In general, detailed studies of hot accretion flow[24,25,23] show that the electron temperature is equal to the ion temperature for $T_e \lesssim 10^9$ K, and then at smaller radii it flattens somewhere between 10^9 K to 10^{11} K due to highly efficient relativistic bremsstrahlung and synchrotron. Hence, in this paper, we approximate the electron temperature

as

$$T_e(r,\theta) = \frac{T_0}{r} \quad \text{for} \quad r > r_1 \tag{4}$$

$$= T_1 \quad \text{for} \quad r \leq r_1, \tag{5}$$

where $T_0 = 10^{12}$ K, T_1 is a constant parameter between 10^9 K and 10^{11} K, and $r_1 \equiv T_0/T_1$. Detailed spectrum calculation for CDAF shows $T_1 \lesssim 10^{11}$ K when the direct viscous heating to electron is significant and $T_1 \lesssim 10^{10}$ K when it is minimal[23].

2.2 Luminosity and Radiation Temperature

We assume that the outflow is generated around the polar axis of opening angle θ_c while the equatorial part of the flow is described by the self-similar CDAF.

Photons from inner part of CDAF are assumed to be dominated by relativistic bremsstrahlung emission. We calculate the radiation temperature and luminosity of these photons and use them as the Compton heating radiation.

3 Condition for Outflow

3.1 Overheating

The thermal equilibrium temperature of electron is determined by the balance between Compton heating and bremsstrahlung cooling. When the equilibrium temperature T_{eq} is higher than the virial temperature, defined as $T_{vir} \equiv (2/5)(m_p c^2/k)(GM/r)$, the gas is overheated and will make an outflow. Hence the condition for onset of outflow is

$$T_{eq}(r,\theta) > T_{vir}(r). \tag{6}$$

Other conditions for outflow include the radiation temperature at a given radius has to be higher than T_{eq}. Also the contribution of low-energy bremsstrahlung photons from the outer part has to be smaller than that of hot Comptonizing photons. And the overheated part should be within the outer boundary of CDAF determined by the global energy balance.[23]

3.2 Shape of the Outflow

We define the critical radius $r_X(\theta)$ that satisfies

$$T_{eq}(r_X) = T_{vir}(r_X). \tag{7}$$

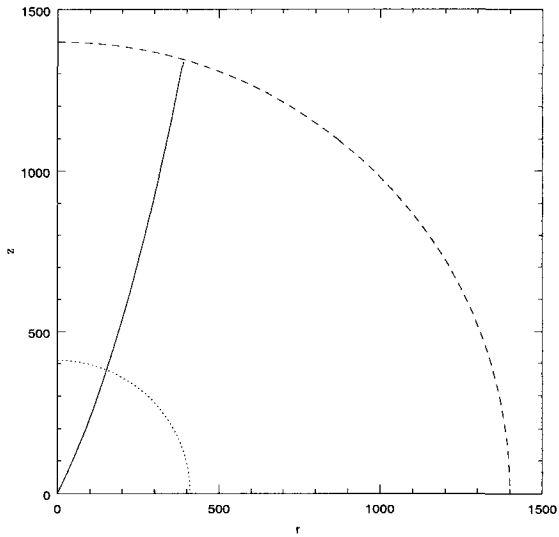


Figure 4: Polar region within the solid curve is overheated in CDAF for $l_{out} = 10^{-4}$ and $T_1 = 10^{11}$ K. Radius in the unit of the Schwarzschild radius.

For total luminosity of $l_{out} = 10^{-4}$ and the inner electron temperature of $T_1 = 10^{11}$ K, $r_X(\theta)$ is the solid curve in Figure 4. The conic region inside the curve toward the polar axis is the overheated producing outflow. The opening angle θ_c, the angle between the curve and the polar axis at $r = 0$, is 27° in this case. Lower luminosity flow has smaller opening angle and the outflow looks more like a jet.

4 Summary and Discussion

We have investigated the condition for Compton heated outflow in self-similar CDAF. Polar outflows or jets are expected under reasonable parameters. The outflow is collimated with an opening angle less than about 20°. CDAF is also more suitable for producing outflow than conventional ADAF.

Acknowledgments

This work was supported by Korea Research Foundation Grant (KRF-2001-015-DP0610).

References

1. M.-G. Park, *ApJ* **354**, 64 (1990)
2. L. Nobili, R. Turolla and L. Zampieri, *ApJ* **383**, 250 (1991)
3. N.I. Shakura *Ast. Zhur.* **49**, 921 (1972)
4. N.I. Shakura and R.A. Sunyaev, *A&A* **24**, 337 (1973)
5. B. Paczyński and P. Wiita, *A&A* **88**, 23 (1980)
6. M. Abramowicz, B. Czerny, J.-P. Lasota and E. Szuszkiewicz, *ApJ* **332**, 646 (1988)
7. S. Ichimaru, *ApJ* **214**, 840 (1977)
8. M.J. Rees, M.-C. Begelman, R.D. Blandford and E.S. Phinney, Nature **295**, 17 (1982)
9. R. Narayan and I. Yi, *ApJ* **428**, L13 (1994)
10. M. Abramowicz, X. Chen, S. Kato, J.-P. Lasota and O. Regev, *ApJ* **438**, L37 (1995)
11. R. Narayan and I. Yi, *ApJ* **444**, 231 (1995)
12. M.-G. Park and J.P. Ostriker, *ApJ* **527**, 247 (1999)
13. M.-G. Park and J.P. Ostriker, *ApJ* **549**, 100 (2001)
14. R.D. Blandford and M.C. Begelman, *MNRAS* **303**, L1 (1999)
15. M.C. Begelman and D.L. Meier, *ApJ* **253**, 873 (1982)
16. I.V. Igumenshchev, X. Chen, and M.A. Abramowicz, *MNRAS* **278**, 236 (1996)
17. I.V. Igumenshchev and M.A. Abramowicz, *MNRAS* **303**, 309 (1999)
18. I.V. Igumenshchev and M.A. Abramowicz, *ApJ* **130**, S463 (2000)
19. J.M. Stone, J. E. Pringle and M.C. Begelman, *MNRAS* **310**, 1002 (1999)
20. I.V. Igumenshchev, M.A. Abramowicz and R. Narayan, *ApJ* **537**, L27 (2000)
21. R. Narayan, I.V. Igumenshchev and M.A. Abramowicz, *ApJ* **539**, 798 (2000)
22. E. Quataert and A. Gruzinov, *ApJ* **539**, 809 (2000)
23. G.H. Ball, R. Narayan and E. Quataert, *ApJ* **552**, 221 (2001)
24. R. Narayan and I. Yi, *ApJ* **452**, 710 (1995)
25. R. Narayan, D. Barret and J.E. McClintock, *ApJ* **482**, 448 (1997)

MODES OF ACCRETION IN X-RAY SOURCES

A. M. BELOBORODOV
Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, M5S 3H8 Ontario, CANADA
andrei@astro.su.se

Three classical modes of accretion are briefly discussed: wind-fed, spherical, and disk. The three modes are illustrated with the mass transfer onto black holes in high-mass X-ray binaries. Then a new regime of mini-disk accretion is described and it is argued that observed wind-fed X-ray sources are likely to accrete in this regime. Switching from one accretion mode to another can cause the observed spectral state transitions.

1 Introduction

Accretion became an important subject of astrophysics when it was realized to feed the brightest X-ray sources in the sky — X-ray binaries. These sources are associated with the most compact astrophysical objects: neutron stars and black holes. Accretion onto such objects releases energy in a relativistically deep potential well, $\phi \sim -c^2$, and can have a high radiative efficiency. It produces radiation in a small region of radius r comparable to the Schwarzschild radius $r_g = 2GM/c^2 = 0.3 \times 10^6 (M/M_\odot)$ where M is the object mass. The observed luminosities L are often comparable to the Eddington limit $L_{\rm E} = 2\pi r_g m_p c^3/\sigma_{\rm T} \approx 1.3 \times 10^{38} (M/M_\odot)$ erg/s. The high L combined with the small size imply a high temperature of the source: a lower bound on the temperature can be estimated assuming blackbody emission, $kT_{\rm bb} \approx (L/\sigma r^2)^{1/4} \sim$ keV. This emission is in the X-ray band.

There are three textbook regimes of accretion: (1) wind-fed (Hoyle & Lyttleton [1]; Bondi & Hoyle [2]) (2) spherical (Bondi [3]), and (3) disk (Shakura & Sunyaev [4]). In all regimes gas falls onto the accretor and releases its gravitational energy, yet the specific accretion mechanisms and the flow patterns differ. In this review we briefly discuss these classical modes and illustrate them with black-hole accretion in high-mass X-ray binaries. Then we describe a new regime of mini-disk accretion.

2 Wind-fed X-ray Binaries

Mass transfer in a binary system with a compact object strongly depends on the type of the normal companion (donor). If the donor is massive (OB star) it emits a powerful wind which is partially captured by the compact companion.

The mass transfer thus occurs even if the donor does not fill its Roche lobe. We will focus on the wind-fed accretion here because it provides nice illustrations of different accretion modes. Also, historically, the first accretion scenario (Bondi-Hoyle-Lyttleton) was wind-fed.

The wind material is captured from a cylinder around the line that connects the two companions. It is called the accretion cylinder and its radius is

$$R_a = \frac{2GM}{w^2} \approx 3 \times 10^{10} \left(\frac{M}{M_\odot}\right) \left(\frac{w}{10^8}\right)^{-2} \text{ cm}, \quad (1)$$

where w is the wind velocity and M is the mass of the accretor. Winds from massive OB stars have $w \approx 10^8$ cm s^{-1}, approximately equal to the escape velocity from the star surface. The separation of the observed X-ray binaries, $A \sim 10^{12}$ cm, is much larger than R_a. This has two consequences: (1) only a small fraction $\sim (R_a/A)^2$ of the wind is captured and (2) the flow inside the accretion cylinder is nearly plane-parallel before it gets trapped by the gravitational field of the accretor.

The wind crosses the distance A and accretes on a timescale $A/w \sim 10^4$ s. It is much shorter than the orbital period of the binary, $P \sim 10^5 - 10^6$ s and one could expect that the orbital rotation weakly affects the accretion flow; in particular, the net angular momentum $\bar{l}$ of the trapped gas must be close to zero. In fact, even a slow orbital rotation and a corresponding small $\bar{l} \neq 0$ can strongly impact the mechanism of accretion: the trapped gas falls many decades in radius onto the compact object ($R_a/r_g \approx 10^5$) and the approximation $\bar{l} = 0$, which is good at $R \sim R_a$, can fail before the gas reaches r_g. A very small $\bar{l} \sim r_g c$ would be sufficient for a disk formation around the compact object and one has to compare the actually trapped $\bar{l}$ with $r_g c$.

The trapped $\bar{l}$ was estimated and compared with $r_g c$ by Illarionov & Sunyaev [5] and Shapiro & Lightman [6]. Consider the accretion cylinder as a sequence of moving pancakes of radius R_a (perpendicular slices of the cylinder). When viewed from the accretor the pancakes have the orbital angular velocity $\Omega = 2\pi/P$ and their specific angular momentum is $\bar{l} = (1/4)\Omega R_a^2$. Using equation (1) one gets $\bar{l}/(r_g c) \approx (P/\text{day})^{-1}(M/\text{M}_\odot)(w/10^8)^{-4}$. Three X-ray binaries are classified as massive ones with black hole companions: Cyg X-1, LMC X-1, and LMC X-3 (see Tanaka & Lewin [7]). They have orbital periods $P = 5.6$, 4.2, and 1.7 d, respectively, and masses M of order of $10M_\odot$. One thus obtains $\bar{l} \sim r_g c$ for these objects, which is marginal for formation of a small-scale disk. Whether the disk forms or not depends on the precise value of $\bar{l}$ which is difficult to calculate because the details of wind accretion are not fully understood (note the strong dependence of $\bar{l}$ on the wind velocity, w).

At distances $R \gg r_g$ from the accretor one can assume $\bar{l} = 0$ and we first

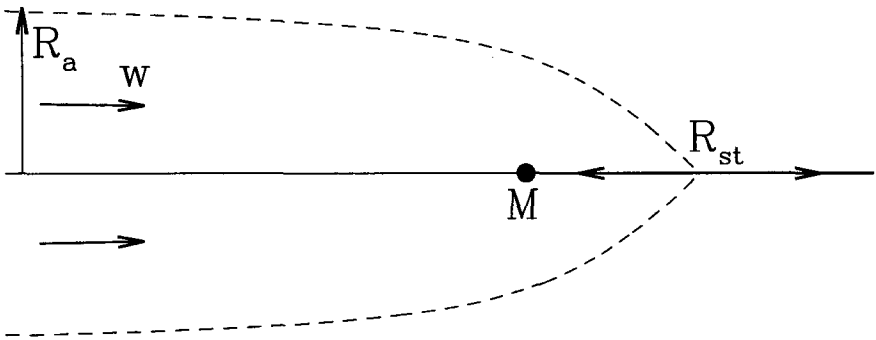


Figure 1: Bondi-Hoyle-Lyttleton accretion. An initially plane-parallel flow is deflected by the accretor and focused onto the axis of symmetry (the accretion line). Further accretion proceeds along the line inside a stagnation radius $R_{\rm st}$.

discuss models with zero net angular momentum (Sect. 3 and 4).

3 Bondi-Hoyle-Lyttleton Accretion

The original model of Bondi & Hoyle [2] envisions an initially plane-parallel flow interacting with an accretor of mass M. The accretor deflects the gas streamlines from the straight lines and captures a part of the flow. The model neglects the gas pressure, so that the streamlines follow ballistic (free-fall) trajectories of test particles in the gravitational field of the accretor. The problem has a symmetry axis and the deflected trajectories intersect on the axis behind the accretor (Fig. 1). This one-dimensional caustic (called the accretion line) is assumed to be sticky. The gas in the caustic has a velocity u^r which differ from the the radial component $\hat{u}^r$ of the free fall that impinges on the caustic. The gas motion in the caustic obeys the momentum conservation law that reads

$$\frac{{\rm d}u^r}{{\rm d}r} = \frac{{\rm d}\dot{M}}{{\rm d}r}\frac{(\hat{u}^r - u^r)}{\dot{M}(r)} - \frac{GM}{r^2 u^r}. \qquad (2)$$

Here $\dot{M}(r)$ is the flux of mass in the caustic, and ${\rm d}\dot{M}/{\rm d}r$ is its feeding rate, which is easily calculated given the ballistic trajectories of the infall. The infall momentum is directed away from the accretor and its inertia competes with gravity. Gravity dominates at small r and here $u^r < 0$ (gas flows in) while at large radii $u^r > 0$ (gas flows out). Hence there must be a stagnation point on the line at a radius $R_{\rm st} \sim R_a$. The $R_{\rm st}$ is not well defined in the

model — infinite number of steady solutions with different $R_{\rm st}$ exist. For a given solution $u^r(r)$ one easily calculates the dissipated energy: $\mathrm{d}E_{\rm diss}/\mathrm{d}r = (\mathrm{d}\dot{M}/\mathrm{d}r)[\hat{u}_\perp^2 + (\hat{u}^r - u^r)^2]/2$ where $\hat{u}_\perp$ is the non-radial velocity component that is canceled where the ballistic flow reaches the caustic.

This simple and beautiful model does not, however, apply to real X-ray sources. The model is based on the assumption that the colliding gas on the accretion line immediately emits the dissipated energy. In fact, the radiative losses are small on the R_a/w timescale and the stored heat creates a high pressure that forces the shocked gas to expand away from the caustic. As a result a steady bow shock sets in at a distance $\sim R_a$ in front of the accretor. Hunt [8] first studied the subsonic gas dynamics behind the shock and showed that a spherically symmetric inflow forms at distances $R < R_a$ from the accretor. Subsequent numerical simulations (e.g. Blondin *et al.* [9]; Ruffert [10,11]) found that the flow is unstable and variable if the bow shock is strong, i.e. if its Mach number $\mathcal{M} = w/c_s$ is well above unity [here $c_s = (10kT/3m_p)^{1/2}$ is the upstream sound speed]. In the case of a modest $\mathcal{M} \gtrsim 1$ the variations are weak and the flow is approximately laminar. This is the most likely case if accretion occurs in the radiation field of a luminous X-ray source which preheats the upstream flow (Illarionov & Beloborodov [12]). The laminar transformation of an initially plane-parallel flow into a spherical infall is schematically shown in Fig. 2.

4 Spherical Accretion

First detailed study of the spherically symmetric (radial) accretion was done by Bondi [3]. The steady gas dynamics is governed by simple equations of baryon conservation and momentum conservation,

$$4\pi R^2 \rho u = \dot{M} = \mathrm{const}, \qquad u\frac{\mathrm{d}u}{\mathrm{d}R} = -\frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}R} - \frac{GM}{R^2}. \tag{3}$$

The problem has three unknowns: baryon density ρ, pressure p, and accretion velocity u. Hence one more equation is needed. In the Bondi model the set of equations is closed by assuming a polytropic equation of state $p = K\rho^\Gamma$. One can then find a family of solutions specified by an accretion rate $\dot{M}$ and an inner boundary condition. The inner boundary differs for neutron stars and black holes; in the latter case the flow must pass through a sonic point since the gas supersonically falls in at the event horizon. The standard theory of polytropic spherical accretion is explained in details in textbooks (see e.g. Shapiro & Teukolsky [13]; Frank, King & Raine [14]).

This theory, however, does not apply to luminous X-ray sources, as was pointed out by Zel'dovich & Shakura [15]. The accreting gas exchanges energy

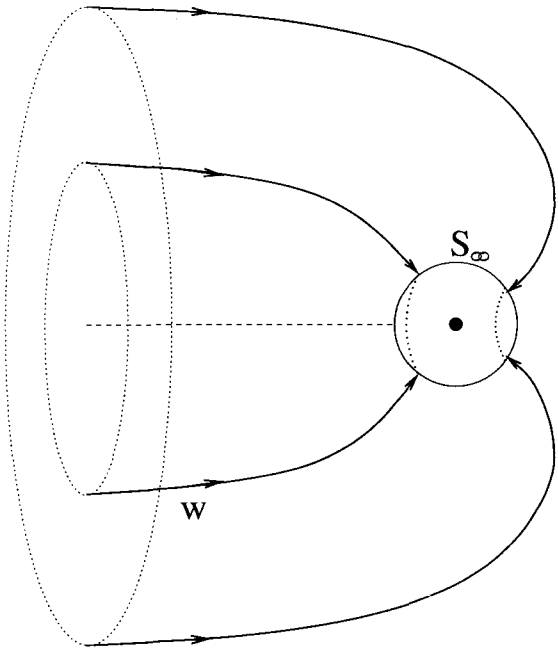


Figure 2: Transformation of a plane-parallel flow into a spherical infall. S_∞ denotes the sphere of radius $R \sim R_C$ (eq. 6). Inside S_∞, the flow is Compton cooled and falling freely.

with the radiation field of the source through Compton scattering and the effective $\Gamma \neq const.$ Compton scattering dominates over other radiative processes and the thermal balance of the accreting gas reads

$$\frac{dT}{dt} = -\frac{2}{3} T \operatorname{div} \mathbf{v} + \frac{T_C - T}{t_C}. \tag{4}$$

The first term on the right-hand side is the compressive heating, and the second term describes the energy exchange with the radiation on the Compton timescale $t_C = 3\pi m_e c^2 R^2/(\sigma_T L)$. The T_C is the Compton temperature of the radiation field. If the radiation scatters off a plasma with $T > T_C$, the photons on average gain energy i.e. cool the plasma, while at $T < T_C$ the photons loose energy (because of the quantum recoil effect in Compton scattering) and heat the plasma. The T_C is defined so that a plasma with $T = T_C$ is in energy equilibrium with the radiation field. It is determined by the shape of the

radiation spectrum F_ν,

$$kT_C = \frac{\int F_\nu h\nu \mathrm{d}\nu}{4\int F_\nu \mathrm{d}\nu}. \quad (5)$$

In hard X-ray sources $T_C \sim 10^8$ K.

The temperature of the wind far from the accretor is relatively low, $T_0 \sim 10^5$ K $\ll T_C$, and where the flow approaches the X-ray source it is heated by Compton scattering. Inside R_a, just where the spherical inflow is forming, the gas is still heated, so that the effective $\Gamma > 5/3$ and changes with R. Compton heating strongly affects the flow, breaks the spherical symmetry, and leads to formation of an inflow-outflow pattern of accretion (Ostriker *et al.*[16]; Illarionov & Kompaneets[17]; Igumenshchev, Illarionov, & Kompaneets[18]). The situation changes inward of the so-called Compton radius

$$R_C = \frac{GMm_p}{5kT_C} \approx 0.1R_a. \quad (6)$$

Here the escape temperature $T_{esc}(R) = GMm_p/(5kR)$ exceeds T_C and a stable spherical inflow forms with $\mathcal{M} > 1$. The plasma is now cooled by the X-rays rather than heated. The cooling timescale t_C gets shorter than the compressive heating timescale R/v, and the gas is cooled well below T_{esc}. It implies that the pressure term in the momentum equation (3) is small compared to the gravitational term and gas falls almost freely, with a high Mach number. Magnetic fields trapped from the donor wind are amplified and dissipated in the converging infall (Bisnovatyi-Kogan & Ruzmaikin[19]; Mészáros[20]). Their energy should not, however, exceed the gas thermal energy, and therefore the fields should not affect free-fall.

The Compton-cooled radial free-fall at $R \ll R_C$ emits little radiation. Its low pressure implies that the compressive heating is weak and the gravitational energy of the gas transforms largerly into the ballistic kinetic energy. If the compact object is a black hole, this energy is swallowed along with the flow, leading to a low radiative efficiency of accretion. This regime can hardly take place in the observed bright sources like Cyg X-1. An appropriate mode of accretion for these sources should be able to dissipate efficiently the infall energy. Dissipation can naturally occur if the infall has a non-zero angular momentum and transforms into a disk before the gas plunges into the event horizon. We now turn to accretion with non-zero angular momentum.

5 Viscous Disk

Accretion with angular momentum l is stopped by the centrifugal barrier at a radius $r = l^2/GM$. Here the gas forms a ring that rotates at about Keplerian

angular velocity $\Omega_K = (GM/r^3)^{1/2}$. In this section we focus on models with $l \gg r_g c$. This is the condition of the standard disk model where steady accretion is possible only if some mechanism redistributes angular momentum and allows the gas to spiral toward the centre. A plausible viscosity mechanism is provided by MHD turbulence that develops due to the differential character of Keplerian rotation (see Balbus & Hawley[21] for a review). Viscosity dissipates the orbital energy of the rotating gas and the gas diffuses to more bound circular orbits. A steady accretion disk dissipates energy with a rate $F_{\rm diss}$ per unit area,

$$F_{\rm diss}(r) = \frac{3\dot{M}}{4\pi}\Omega_K^2 S, \tag{7}$$

where $S(r)$ is a numerical factor determined by the inner boundary condition. For a Schwarzschild black hole $S = 1 - (3r_g/R)^{1/2}$ so that $S = 0$ at the inner boundary of the viscous disk, $r = 3r_g$, where the gas starts to fall freely into the hole. This standard disk model was successfully applied to X-ray binaries and active galactic nuclei (AGN), the latter being just a scaled version of the standard model that corresponds to a massive black hole, $M = 10^6 - 10^9 M_\odot$.

If the dissipated energy is immediately radiated away, the emerging radiation flux from the disk surface is $F(r) = F_{\rm diss}(r)/2$ (the factor 1/2 accounts for the two faces of the disk). The total emitted flux is easily calculated, however, its spectrum F_ν is difficult to predict. The simplest model assumes that the disk emits as a black body and then one finds the radial profile of its surface temperature, $T_{\rm bb} = (F/\sigma)^{1/4} \propto r^{-3/4}$. The integrated disk spectrum (multicolor blackbody) has an exponential cutoff at $\sim$ keV in X-ray binaries and $\sim$ 10 eV in AGN. Such emission (with the right temperature) is indeed observed in many sources. However, in addition, one observes a second spectral component that peaks at about 100 keV. Example spectra of the famous black hole candidate Cyg X-1 are shown in Fig. 3. Most of the time this object spends in the 'hard' state when the 100 keV component is dominant. Sometimes it switches to the 'soft' state where the hard emission is weak and the soft blackbody component is strong.

Phenomenologically, a two-peak emission spectrum is well explained. There must be two gaseous components that are heated in the accretion flow: one, dense, emits as a black body (presumably an optically thick accretion disk), and the other is rarefied and hot, with $kT \sim 100$ keV. A fraction of blackbody photons from the dense component get upscattered in the hot plasma, become hard X-rays and form the second spectral peak (e.g. Poutanen[26], Zdziarski[27]). This Comptonization process is well understood and one can derive the optical depth τ_T (to electron scattering) of the hot plasma from the spectral data. Interestingly, one finds $\tau_T \approx 1$ along with $kT \approx 100$ keV in various sources

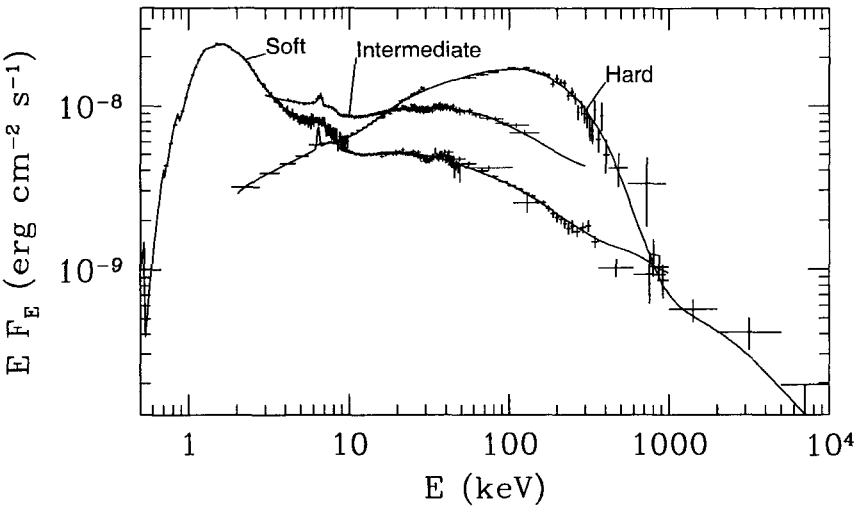


Figure 3: Spectrum of Cyg X-1 in the hard, soft, and intermediate states. The hard-state spectrum is obtained in 1991 by CGRO (Gierliński *et al.* [22]. McConnell *et al.* [23]. The intermediate-state spectrum was obtained by RXTE on May 23, 1996. The soft-state spectrum was obtained by ASCA and RXTE on May 30, 1996 and by OSSE between June 14-25, 1996 (Gierliński *et al.* [24]). The solid curves show the best fits to the data by the EQPAIR models. (From Coppi [24].)

which include both X-ray binaries and AGN. However the geometry of the hot+cold gas is still unknown and it is unclear why the accretion flow chooses to heat gas with $\tau_T \approx 1$. Two basic suggestions are (see Beloborodov [28] for a review): (1) buoyant magnetic loops are generated in the turbulent disk and form an active corona (Galeev, Rosner & Vaiana [29]), and (2) the cold disk is transformed into a hot two-temperature torus at a radius r_{tr}; the electron temperature in the torus is $kT_e \approx 100$ keV (Shapiro, Lightman & Eardley [30]). The two modifications of the standard disk look likely, yet both are strongly dependent on the poorly understood viscosity/heating mechanism.

6 Mini-Disk

As it was discussed in Sect. 2, the estimated angular momentum of the wind-fed accretion flows in X-ray binaries is marginally enough to form a disk. Is a disk actually formed there? — Probably yes, because otherwise accretion would occur in the quasi-spherical regime with little dissipation, which is hard to reconcile with the observed high luminosities. If so, at what radius r_0 does

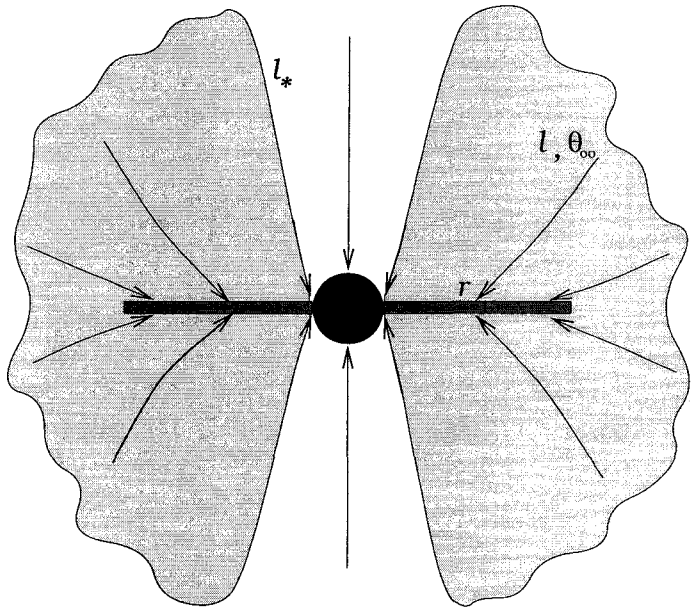


Figure 4: Schematic picture of the disk formation. The inflow has angular momentum $l(\theta_\infty)$ increasing toward the equatorial plane, $\mathrm{d}l/\mathrm{d}\sin\theta_\infty > 0$. The collision radius, which is a growing function of l, monotonically increases as θ_∞ approaches $\pi/2$. The shadowed parts of the accretion flow collide outside r_g and form a couple of radiative shocks which sandwich the caustic. The other part of the flow (at small polar angles where $l < l_*$) is directly swallowed by the black hole with a low radiative efficiency.

the disk form and how does it compare with $3r_g$ (the inner edge of the standard disk)? What is the critical radius $r_{\rm cr}$ for the *viscous* disk formation?

The critical radius turns out quite large, $r_{\rm cr} \approx 14r_g$, and the accretion regime with $r_g < r_0 < r_{\rm cr}$ needs to be addressed (Beloborodov & Illarionov[31]). Such a 'mini-disk' is very different from the standard viscous disk as regards both the dynamics and the emission mechanism.

6.1 The Caustic

Fig. 4 schematically shows the disk formation in the inner region of a quasi-spherical inflow with angular momentum. The inflow is almost radial at large radii, it is Compton cooled at $R_{\rm C} \sim 10^4 r_g$ and falls freely toward the center. Near the black hole, the rotation deflects the infall from the radial direction and a disk-like caustic appears in the defocused flow. The flow is assumed to be

symmetric around the spin axis and also symmetric about the equatorial plane. In this plane a ballistic streamline collides with its symmetric counterpart. The radius of collision is determined by the angular momentum of the streamline, l,

$$r = \frac{l^2}{GM} - \frac{r_g}{7}, \tag{8}$$

where $r_g/7$ is a relativistic correction to the Newtonian formula. The angular momentum of a streamline depends on its initial polar angle at $R \gg r_g$, which we denote θ_∞. A likely distribution $l(\theta_\infty)$ has $l = 0$ on the polar axis ($\theta_\infty = 0, \pi$) and a maximum l_0 at the equatorial plane ($\theta_\infty = \pi/2$). A simple example is a 'solid body' rotation, $l(\theta_\infty) = l_0 \sin\theta_\infty$. From equation (8) one can see that streamlines with larger l collide at larger r. The polar part of the inflow with $l < l_* = 0.75 r_g c$ crosses the event horizon before it can collide, while the equatorial part with $l > l_*$ forms a caustic outside r_g (Fig. 4). The outer edge of the caustic is defined by the streamlines with $\theta_\infty \to \pi/2$ and its radius is $r_0 = l_0^2/GM - r_g/7$. The condition for the disk formation outside the black hole ($r_0 > r_g$) reads $l_0 > l_*$.

When the ballistic flow approaches the caustic, it passes through a shock and then continues to accrete through the disk. A collisionless shock forms on each side of the caustic. The accretion picture shown in Fig. 4 assumes that the shock is radiative and pinned to the equatorial plane. This is indeed the case if the source luminosity exceeds $10^{-2} L_{\rm E}$ (Beloborodov & Illarionov [31]). The shocked ions pass their energy via Coulomb collisions to the radiatively efficient electrons on a timescale

$$t_{\rm Coul} = \frac{\sqrt{\pi/2}\, m_p \theta_e^{3/2}}{m_e c \sigma_{\rm T} z n \ln\Lambda} \approx 17 \frac{T_e^{3/2}}{zn}\,{\rm s}, \qquad \theta_e \equiv \frac{kT_e}{m_e c^2}. \tag{9}$$

Here n is the postshock proton density, $\ln\Lambda \approx 15$ is a Coulomb logarithm, and $z > 1$ accounts for the possible $e^\pm$ production. The postshock electrons have a much smaller temperature than the protons, $T_e \ll T_p$. They immediately radiate the received heat by Compton scattering, and T_e is controlled by the cooling=heating balance,

$$\frac{T_e}{t_{\rm C}} = \frac{T_p}{t_{\rm Coul}}, \tag{10}$$

where $t_{\rm C} = 3 m_e c^2/(8\sigma_{\rm T} F)$ is the Compton cooling time-scale and F is the local radiation flux. For a radiative shock $F \sim (3/2) n k T_p v_{\rm ff}/\xi$ where $v_{\rm ff} = (2GM/r)^{1/2}$ is the free-fall velocity at a radius r and $\xi \geq 4$ is the compression at the shock front. Then $t_{\rm C} \approx m_e c^2 \xi/(8\sigma_{\rm T} n k T_p v_{\rm ff})$ and the energy balance (10)

yields

$$\frac{kT_e}{m_e c^2} \approx \left(\frac{\ln \Lambda \, m_e}{\sqrt{2\pi} m_p} \frac{\xi c}{4 v_{\rm ff}} \right)^{2/5} \approx 0.1 \left(\frac{\xi c}{4 v_{\rm ff}} \right)^{2/5} . \tag{11}$$

Density scaled out from this estimate. Given $\xi \approx 4$ and $v_{\rm ff} \approx c$ one finds $kT \approx 100$ keV.

The hot postshock layer that radiates most of the proton heat has optical depth $\tau_{\rm T} \approx (v_{\rm ff}/\xi) t_{\rm Coul} z n \sigma_{\rm T}$ and satisfies the relation $\tau_{\rm T} \theta_e \approx 0.1$; one thus gets $\tau_{\rm T} \sim 1$ for typical $\theta_e \sim 0.1$. In deeper layers, at $\tau_{\rm T} > 1$, T_p and T_e fall off sharply and the shocked plasma condenses to a thin dense disk. A similar radiative shock on a surface of a neutron star was studied by Zel'dovich & Shakura[15] and Shapiro & Salpeter[32].

The pattern of accretion with shocks pinned to the caustic requires a sufficiently high accretion rate $\dot{M}$ corresponding to the disk luminosity $L > 10^{-2} L_{\rm E}$. At smaller $\dot{M}$ the shocked plasma is unable to cool rapidly and it supports a quasi-spherical shock front far away from the disk. Such a flow was simulated numerically by Igumenshchev, Illarionov & Abramowicz[33]. The simulations show that the shock tends to shrink toward the disk of radius l_0^2/GM when $\dot{M}$ increases.

6.2 Accretion in the Mini-Disk

The disk is composed of the shocked matter that entered the disk with different horizontal velocities. A strong turbulence develops under such conditions and efficiently mixes the disk in the vertical direction. As a result, the disk shares the horizontal momentum with the feeding infall and accretes with a vertically-averaged horizontal velocity. This situation can be described as a 'sticky' caustic, and it resembles the accretion line in the Bondi-Hoyle-Lyttleton (BHL) problem. The infall absorption by the sticky caustic implies energy dissipation, and the released heat is radiated away, contributing to the total luminosity. The process of disk-infall inelastic collision is governed by the corresponding law of momentum conservation. The relativistic conservation law $\nabla_k T_i^k = 0$ in the Schwarzschild metric yields

$$\frac{{\rm d}u^r}{{\rm d}r} = \frac{{\rm d}\dot{M}}{{\rm d}r} \frac{(\hat{u}^r - u^r)}{\dot{M}(r)} - \frac{GM}{r^2 u^r} \left(1 - \frac{j^2}{j_{\rm K}^2} \right) , \tag{12}$$

where $j_{\rm K}^2(r) = r^2 r_g c^2/(2r - 3r_g)$ is the squared angular momentum of circular Keplerian orbits. The first term on the right-hand-side is a result of the disk-infall interaction. Equation (12) is similar to the BHL equation (2) that describes gas motion along the accretion line. The only difference is the

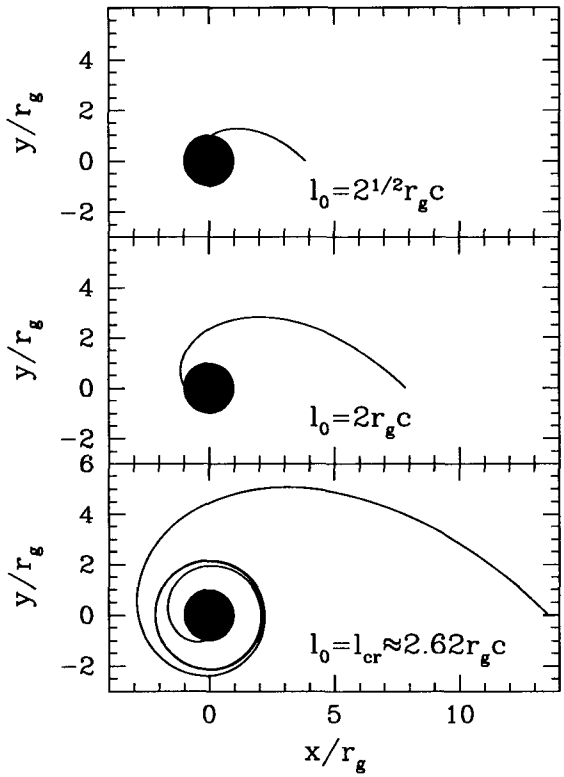


Figure 5: The trajectory of gas motion in the disk, from the outer edge into the black hole. In the critical case (bottom panel) the trajectory makes an infinite number of revolutions at the critical radius $r_{\rm cr} \approx 2.14 r_g$ before falling into the black hole.

additional term proportional to $j^2/j_{\rm K}^2$. It appears because now the accreting matter rotates and experiences a centrifugal acceleration. The repulsive centrifugal force dominates over gravity if $j > j_{\rm K}$ and it can stop accretion. Therefore, there exists a critical angular momentum $j_{\rm cr}$ above which the free-fall accretion is stuck and requires viscosity like the standard disk does.

The critical angular momentum can be calculated. For illustration, consider a flow with $l(\theta_\infty) = l_0 \sin\theta_\infty$; l_0 is the only parameter of the problem. The calculated streamlines of the mini-disk are shown in Fig. 5. A disk with $l_0 < j_{\rm cr} = 2.62 r_g c$ accretes on the free-fall timescale, with negligible horizontal viscous stresses. Its radius can be as large as $14 r_g$. A streamline of the disk starts at the outer edge r_0 with a relatively large angular momentum $j = l_0$ and

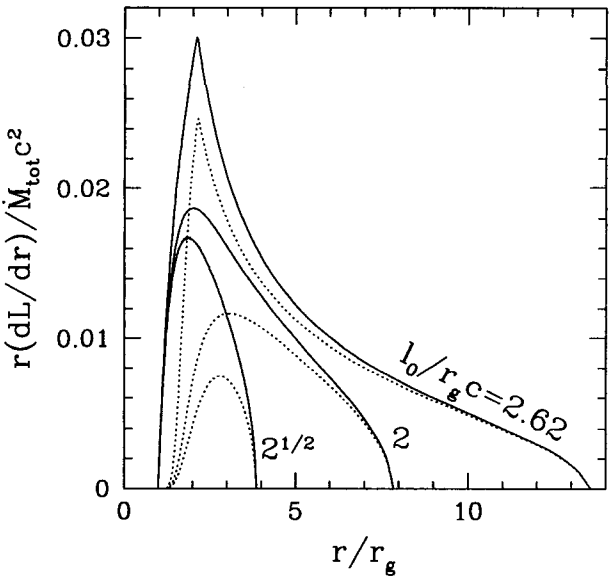


Figure 6: The radial distribution of the disk luminosity: solid curves – the released luminosity, dotted curves – the observed luminosity corrected for the radiation capture by the black hole. Three models are shown with $l_0 = j_{cr} \approx 2.62 r_g c$, $l_0 = 2 r_g c$, and $l_0 = \sqrt{2} r_g c$.

then, as the streamline approaches the black hole, j decreases. This happens because at smaller r the disk absorbs the infall with smaller angular momentum; it allows the disk to overcome the centrifugal barrier without horizontal viscous stresses and accrete even with an initially large l_0. The bottleneck for accretion is at $2r_g$ and here accretion can occur only if $j < 2r_g c$. One can thus derive a necessary condition for the inviscid accretion regime that reads $\bar{l} < 2r_g c$ where $\bar{l}$ is the mean angular momentum of the whole flow.

6.3 Emission from the Mini-Disk

The energy dissipated in the inelastic caustic-infall interaction is easily calculated for a given dynamic solution, like it was done in the BHL model. The radial distribution of the resulting luminosity is shown in Fig. 6. It peaks at $2r_g$ where the centrifugal deceleration of the inflow is strongest (the bottleneck for accretion). Here the relative velocity between the disk and the infall increases and the inelastic collision liberates more energy. The small size of the

main emitting region implies that the released luminosity is partially captured into the black hole, and the corresponding correction is shown in Fig. 6.

The net radiative efficiency of accretion $\eta = L/\dot{M}c^2$ depends on l_0: it rises from $\eta = 0$ at $l_0 < l_*$ to a maximum $\eta \approx 0.0372$ at $l_0 = j_{\rm cr}$. This maximum is comparable to the efficiency of the standard disk around a Schwarzschild black hole, $\eta = 1 - 2\sqrt{2}/3 \approx 0.0572$.

The energy is released in two steps: (1) at the shock front where the vertical momentum of the infall is canceled and (2) inside the disk where the infall inelastically shares its horizontal momentum with the disk. The first step heats a postshock layer with $\tau_{\rm T} \approx 1$ and $kT_e \approx 100$ keV, and the subsequent dissipation in the geometrically thin dense disk can emit optically thick cool radiation. The two gaseous components thus emit a two-peak radiation spectrum that is typical for black-hole sources.

6.4 Comparison with Other Accretion Models

The mini-disk is a sticky caustic in a freely falling flow and in this respect it resembles the BHL accretion. In their case, accretion proceeds along a one-dimensional caustic, the accretion line. The accretion line is fed at each point by matter from an initially *plane-parallel* flow which is axisymmetric and has *zero* net angular momentum. The mini-disk is a two-dimensional caustic in an asymptotically *isotropic* inflow with a *non-zero* net angular momentum. In contrast to the BHL accretion line, the radial momentum of matter impinging on the disk is directed inward. As a result, the mini-disk does not have the degenerate stagnation radius that exists in the BHL problem.

For large angular momenta, the disk accretion is stopped by the centrifugal barrier and then the mini-disk regime switches to the standard viscous regime. The major differences of the mini-disk from its standard counterpart are: (1) it is inviscid and accretes supersonically, (2) the energy release in the disk is caused by the infall absorption, and (3) most of the energy is released inside $3r_g$.

7 Discussion

Angular momentum l is a major parameter of an accretion flow. Two classical modes of accretion — spherical and viscous disk — are valid in the limits $l \ll r_g c$ and $l \gg r_g c$, respectively. The two limits are bridged by an intermediate, but qualitatively different, mini-disk regime. It takes place for the range of the disk radii $r_g < r_0 < 14 r_g$.

Motivation for the theoretical study of the mini-disk comes from the observed high-mass X-ray binaries with black holes: the wind-fed accretion flows

there have $l \sim r_g c$. This fact is interesting: why should wind-fed black holes accrete near the threshold for disk formation? One possible explanation is as follows. Suppose that the dominant majority of black holes with massive companions accrete gas with small $l < l_*$. It should be difficult to observe such objects because of their low luminosity. When l exceeds l_* and the disk forms, the luminosity rises dramatically and the black hole switches on as an X-ray source. One therefore expects to see preferentially objects with l above the threshold for disk formation. If the l-distribution of objects falls steeply toward high l, most of the bright sources should be near the threshold, i.e. the regime $\bar{l} \sim r_g c$ can be widespread among observed *bright* sources.

One of the most intriguing properties of accreting black holes is their spectral state transitions. An example of such transitions in Cyg X-1 is shown in Fig. 3. The transitions may be caused by modest changes of the angular momentum of the accretion flow. Such changes should naturally occur since the trapped l is very small and sensitive to the fluctuations of the donor wind. Suppose a black hole accretes in the mini-disk regime and emits a strong 100 keV emission. If the angular momentum increases by a factor of 2, the disk has to switch to the standard viscous regime with a higher luminosity and softer spectrum. This results in a hard-to-soft transition. If the angular momentum decreases by a factor of 2, the source luminosity would fall down. Similar spectral state transitions were recently observed in LMC X-3 (Wilms *et al.*[34]).

References

1. F. Hoyle and R.A. Lyttleton, *Proc. Cam. Phil. Soc.* **35**, 405 (1939).
2. H. Bondi and F. Hoyle, *MNRAS* **104**, 273 (1944).
3. H. Bondi, *MNRAS* **112**, 195 (1952).
4. N.I. Shakura and R.A. Sunyaev, *A&A* **24**, 337 (1973).
5. A.F. Illarionov and R.A. Sunyaev, *A&A* **39**, 185 (1975).
6. S. L. Shapiro and A.P. Lightman, *ApJ* **204**, 555 (1976).
7. Y. Tanaka and W.H.G. Lewin, in *X-ray Binaries*, eds. W.H.G., J. van Paradijs, and E.P.J. van den Heuvel, 126 (Cambridge Univ. Press, Cambridge, 1995).
8. R. Hunt, *MNRAS* **154**, 141 (1971).
9. J.M. Blondin, T.R. Kallman, B.A. Fryxell, and R.E. Taam, *ApJ* **356**, 591 (1990).
10. M. Ruffert, *A&A* **317**, 793 (1997).
11. M. Ruffert, *A&A* **346**, 861 (1999).
12. A.F. Illarionov and A.M. Beloborodov, *MNRAS* **323**, 159 (2001).
13. S.L. Shapiro and S.A. Teukolsky, *Black Holes, White Dwarfs, and Neu-*

tron Stars (Wiley-Interscience, New York, 1983).
14. J. Frank, A. King, and D. Raine, *Accretion Power in Astrophysics* (Cambridge Univ. Press, Cambridge, 1992).
15. Ya.B. Zel'dovich and N.I. Shakura, *Soviet Astron.* **13**, 175 (1969).
16. J.P. Ostriker, R. McCray, R. Weaver, and A. Yahil, *ApJ* **208**, L61 (1976).
17. A.F. Illarionov and D.A. Kompaneets, *MNRAS* **247**, 219 (1990).
18. I.V. Igumenshchev, A.F. Illarionov, and D.A. Kompaneets, *MNRAS* **260**, 727 (1993).
19. G.S. Bisnovatyi-Kogan, A.A. Ruzmaikin, *Ap&SS* **28**, 45 (1974).
20. P. Mészáros, *A&A* **44**, 59 (1975).
21. S.A. Balbus and J.F. Hawley, *Rev. Mod. Phys.* **70**, 1 (1998).
22. M. Gierliński et al., *MNRAS* **288**, 958 (1997).
23. M.L. McConnell et al., *ApJ* **424**, 933 (1994).
24. M. Gierliński et al., *MNRAS* **309**, 496 (1999).
25. P.S. Coppi, in *High Energy Processes in Accreting Black Holes*, eds. J. Poutanen and R. Svensson, ASP Conf. Series, **161**, 375 (1999).
26. J. Poutanen, in *Theory of Black Hole Accretion Disks*, eds. M.A. Abramowicz, G. Björnsson, and J.E. Pringle, 100 (Cambridge Univ. Press, Cambridge, 1998).
27. A.A. Zdziarski, 1999, in *High Energy Processes in Accreting Black Holes*, eds. J. Poutanen and R. Svensson, ASP Conf. Series, **161**, 16 (1999).
28. A.M. Beloborodov, in *High Energy Processes in Accreting Black Holes*, eds. J. Poutanen and R. Svensson, ASP Conf. Series, **161**, 295 (1999).
29. A.A. Galeev, R. Rosner, and G.S. Vaiana, *ApJ* **229**, 318 (1979).
30. S.L. Shapiro, A.P. Lightman, and D.M. Eardley, *ApJ* **204**, 187 (1976).
31. A.M. Beloborodov and A.F. Illarionov, *MNRAS* **323**, 167 (2001).
32. S.L. Shapiro and E.E. Salpeter, *ApJ* **198**, 671 (1975).
33. I.V. Igumenshchev, A.F. Illarionov, and M.A. Abramowicz, *ApJ* **517**, L55 (1999).
34. J. Wilms et al., *MNRAS* **320**, 327 (2001).

This page is intentionally left blank

PART 3

Energy Extraction from Rotating Black Holes

This page is intentionally left blank

CURRENT HIGH ENERGY EMISSION FROM BLACK HOLES

R. D. BLANDFORD
130-33 Caltech
Pasadena, CA 91125, USA
E-mail: rdb@caltech.edu

Two related topics are discussed. 1. Accretion onto black holes at low and high (though not very high) rates is believed to proceed adiabatically (*ie* non-radiatively). It is argued that the liberated energy is carried off by an outflow, probably involving almost all of the gas that is supplied. Two dimensional, fluid, accretion disks, in which mass, angular momentum and energy are transported to the disk surface, are summarized. It is conjectured that relatively minor changes are needed to describe magnetised disks. By contrast, the disk surface physics is argued to dictate the character of the outflow. 2. Ultrarelativistic jets appear to be produced by active galactic nuclei (AGN), pulsars and γ-ray bursts (GRB). In all three cases, it is argued that the power is generated electromagnetically by a magnetic rotator, (in a DC not AC form), and transported in this manner to the emission site. A model of a relativistically expanding electromagnetic shell is described and used to provide a simple model of a GRB in which the γ-rays are produced by unstable electrical currents flowing along the rotation axis. The shell drives a relativistic blast wave into the surrounding medium with a speed that varies with latitude and whose afterglow emission may exhibit achromatic breaks. Similar processes may be at work in non-relativistic plerions like the Crab Nebula and, possibly, AGN jets. The observational implications of these two classes of model and the prospects for performing instructive, numerical experiments to elucidate them further are briefly outlined.

1 Introduction

I would like to take literally the title of this meeting and summarize two, related, theories of how black holes (and, to some extent, magnetised neutron stars) behave. The first, which draws upon research carried out with Mitch Begelman, is a particular viewpoint on how accretion proceeds at rates that range from low to high. The second, which is summarised in greater depth elsewhere[6], involves collaboration with Max Lyutikov. It contains some new approaches to describing how electromagnetic Poynting flux is released by a black hole (or indeed any compact spinning, magnetised body) and how it propagates to a remote site from which observable radiation is emitted. I should caution the reader that these are both active and controversial research areas and there are many other points of view that I shall not review. However, I will try to highlight one or two areas where I think that the differences between models are sharply delineated and where I anticipate that progress is imminent.

2 Ondol Dynamics

2.1 Modes of Accretion

Accretion [a] onto a black hole is approximately scale-free, in the sense that the mode of the accretion is most strongly determined by the ratio of the mass *supply* rate to the Eddington rate ($\dot{M}_{\rm E} = 4\pi GM/\kappa_T c$) and the product of the angular velocity of the hole, Ω, with its mass M [Ferrarese; henceforth, other contributors to these proceedings will be designated using square brackets]. (It is assumed that the gas has sufficient angular momentum to form a disk.) There will be some sensitivity to the mass through the effective temperature. The mass supply rate is usually estimated by the Bondi rate $\dot{M}_B \sim \pi r_B^2 \rho_B s_B \sim \dot{M}_E \tau_B c/s_B$, where ρ_B, s_B, τ_B are the density, sound speed and Thomson optical depth at the Bondi radius, $r_B \sim 2GM/s_B^2$. This should be true for black holes with masses that range from ~ 5 $M_\odot$ stellar holes to the many billion solar mass holes that are thought to power quasars [10 8].

Four accretion states can be distinguished, although it is hard to be quantitative without a better understanding of the underlying fluid mechanics and plasma physics.

- *Drought ($\dot{M} << \dot{M}_E$)* The accreting gas is *adiabatic*. That is to say it is unable to cool and radiate away its binding energy within some *transition* radius $r_{\rm trans}$. This requires that the internal energy be carried mostly by hot ions which do not heat the electrons enough to enable them to radiate efficiently on the inflow timescale [13]. Instead, the surplus energy, which is transported outward by magnetic stress, off in an outflow. This may account for most of the mass supply so that the hole accretion rate is much less than the mass supply [7].

  Possible examples include low power, galactic nuclei, (including Sgr A*), AGN with powerful radio sources (including quasars) and binary X-ray sources in their low hard states [Menou, Mineshige].

- *Rain ($\dot{M} \sim \dot{M}_E$* The gas can radiate efficiently all the way down to the marginally stable circular orbit $r_{\rm ms}$. The flow is conservative so that the accretion rate is similar to $\dot{M}$ [14]. Most of the radiation will be emitted quasi-thermally by the disk, although there is likely to be considerable, nonthermal, coronal activity as well. The inner disk is radiation- and electron scattering-dominated with thickness $H \sim \dot{M}\kappa/4\pi c$. These disks are expected and observed to be unstable and are probably quite inho-

[a] Ondol is a special Korean system of underfloor heating.

mogeneous. [b] The high gas density and the intense radiation field may prevent the formation of relativistic jets.

Possible examples include Seyfert galaxies, radio-quiet quasars and binary X-ray sources in their high, soft states.

- *Deluge* ($\dot{M}_E \lesssim \dot{M} \lesssim \dot{M}_E(c/s_B)^2$) Radiation is trapped within the trapping radius $r_{\rm trap} \sim \dot{M}\kappa_T/4\pi c$ within which the accreting gas flows faster than the photons can diffuse. The flow is again adiabatic and the energy that is released is carried off by a mildly relativistic, radiation-dominated wind [7] [Park]. The hole accretes at the roughly the Eddington rate.

  Possible examples include broad absorption line quasars and Galactic superluminal sources like GRS1915+105.

- *Inundation* ($\dot{M} \gtrsim \dot{M}_E(c/s_B)^2$) The trapping radius lies beyond r_B. Strongly super-Eddington accretion is possible because outflows from the inner disk are unable to escape.

  Possible examples include massive black holes during their rapid growth phases and some GRB models.

The reason why outflows are inevitable when the flow is adiabatic is that the torque G in the disk automatically transports energy radially outward at a rate $G\Omega$, where Ω is the angular velcocity [7]. There is a local heating rate $G\nabla\Omega$ which can unbind gas in a nearly Keplerian disk. [c] The energy liberated by each proton that (altruistically) cross the horizon drives a powerful wind from larger radii which allows many (up to 10^5) other protons to escape. [d]

2.2 The Importance of Magnetic Field

Early models of accretion disks [14] were non-specific about the nature of the viscous torque allowing it to be either fluid or hydromagnetic in character. We now know that, under almost all circumstances, the latter is correct. The magnetorotational instability [1] ensures that magnetic fields grow to dynamically important strength on the orbital timescale [Vishniac]. The nonlinear evolution of this instability is still far from clear. However, it is likely that it involves coronal heating, jet formation, and possibly emission line cloud dynamics.

[b] This may permit radiation to escape with a luminosity in excess of the Eddington limit[3].
[c] This is not necessarily true for a disk that extends beyond r_B where the potential is much softer.
[d] It is possible that the energy is carried off hydromagnetically, in which case, there need be little mass loss.

What is clear is that we are still quite ignorant of the true laws of MHD. The global evolution of magnetic field, as witnessed much closer to home in the terrestrial magnetosphere and the solar corona, is controlled by principles of global stability and physical processes like reconnection, conduction and equilibration, which we do not understand well enough to generalize to astrophysical environments. However, the prospects are generally quite good for improving our understanding of these matters through more detailed space physical observation, terrestrial experiments using giant pinches, lasers and particle beams and numerical simulation.

2.3 Two Dimensional Adiabatic Disks

Despite these remarks it is still instructive (and easier) to consider fluid disks. These are inevitably at least two dimensional [Lee]. A useful approximation is that the viscosity is small enough that they can be treated as being in approximate hydrostatic equilibrium. This requires specifying how the density varies along isobars which, in turn, depends upon the manner in which energy, mass and angular momentum are transported through the disk.

Two dimensional adiabatic disks are likely to become linearly unstable according to the Høiland criterion, which is a linear combination of the more familiar Schwarzschild and Rayleigh prescriptions for instability and appropriate for hot, rotating flows. There are actually two criteria. Most accretion disks are quite stable to the first of these which essentially refers to radial interchanges. It is the second criterion that is more relevant. It essentially states that fluid disks are marginally unstable to the interchange of slender rings when they are gyrentropic; that is to say when the surfaces of constant entropy and specific angular momentum coincide. (In fact this implies that the surfaces of constant Bernoulli function are also coincident.) Furthermore, the nature of the unstable modes is such that if a disk is only slightly unstable, the motions of fluid elements will transport mass, angular momentum and energy along these gyrentropes to the disk surface where they can be transferred to an outflow.

The precise properties of the outflow depend upon the detailed nature of the dissipation and momentum transfer that takes place at the disk surface and these processes are not even understood in the case of the sun. If the wind is thermally driven, then there must be an entropy jump at its base as the gas in the disk changes from being bound to unbound. However it also possible that the gas remains relatively cold and that magnetic stress create a momentum-driven wind as appears to happens in the case of the solar wind, where two million degree gas acquires a velocity at infinity of ~ 800 km s^{-1}. In

addition, there is the possibility of launching a cold, magnetocentrifugal wind from a near Keplerian accretion disk – a possibility that does not exist in the case of the sun. The key deduction is that the structure of adiabatic, accretion disks depends upon the unknown physics of mass outflow.

However, even this is not the whole story. The gas that remains in the disk must flow inward on a viscous timescale and the manner in which this happens depends upon the prescription for the viscosity. Indeed, in the presence of convective motion, there is a natural quadrupolar circulation in the disk which will be established and this must be superposed upon the net flow to the disk surface and smaller radii. Just as is the case with Eddington-Sweet circulation, the flow adjusts to ensure that there is no local accumulation of mass angular momentum and energy. [e]

It is possible to construct self-consistent, self-similar models that exhibit convection, inflow, outflow and circulation after specifying a functional form for the viscous torque and outflow launching mechanism[6]. These models are generic and their general features are displayed by some numerical simulations of two dimensional disks[15]. Self-similar disks are subject to the criticism that they cannot describe the flow close to the hole, where most of the energy is released. This criticism can be met by constructing an explicit non self-similar, relativistic torus model around $r_{\rm ms}$. The physical assumptions that must be introduced to create such a model are necessarily arbitrary but they suffice to demonstrate that there is no difficulty of principle associated with terminating a self-similar solution at small radius. Similar concerns have been raised concerning the outer radius. Understanding this transition has turned out to be a much harder problem.

2.4 Magnetized Disks

We have just argued that fluid disks are gyrentropic. However, this is unlikely to be true of magnetised disks. The dynamics is clearly complex [Krolik]. For example, if it turns out that the magnetic flux preferentially settles into isorotational surfaces and distributes the angular momentum rapidly on these surfaces, then the disk structure will be barytropic. This prescription has some numerical support [16]. Whatever the correct prescription, the self-similar models described above can be modifed with only minor changes. It is the prescription for generating the outflow at the disk surface that is more important for determining the disk structure.

[e] If we do not admit the possibility of an outflow from the disk surface, then the global circulation will transport heat from small radii, where it is mostly liberated, to $r_{\rm trans}$ where it must either be radiated away or inflate the disk so that the flow becomes unsteady.

3 Seungmu Dynamics

3.1 Ultrarelativistic Jets

Ultrarelativistic [f] jets are found in a variety of locales. They were first seen in extragalactic radio sources [Celotti, Sikora]. Indirect arguments, together with direct measurement of superluminal expansion show that the emitting material can move with a Lorentz factor $\Gamma \sim 10$ and be collimated into a cone with opening angle $\sim 5°$. There is now evidence that their sources are spinning, massive black holes with their attendant accretion disks and that they are collimated within $\sim 100m$. Most jets appear to be associated with accretion disks that radiate well below their Eddington limits and carry as much, or even more, power than is radiated by the disk. [g]

The second example of ultrarelativistic jets is provided by plerions. These are supernova remnants, exempified by the Crab Nebula, with central, rapidly spinning, magnetised, neutron stars. The rotational energy of the neutron star appears to be carried off by a relativistic outflow. Recent, observations by the Chandra Observatory (now replicated in other plerions) show a pair of jet-like features, together with an equatorial disk of X-ray emission. This was surprising because it was thought that jets required an accretion disk to form. The region, that is observed directly, is clearly not moving with ultrarelativistic speed as they would then be beamed away from us. However, it is reasonable supposition that they contain invisible, ultrarelativistic cores.

The third example is provided by GRBs which are known to be cosmologically distant and, consequently, extremely energetic. It has been inferred that the outflow is ultrarelativistic, with Lorentz factors $\Gamma \sim 300$ as γ-rays with energy in excess of ~ 0.5 MeV have to escape without creating pairs [12]. They are popularly associated with stellar compact objects and it was guessed that they might be beamed, in order that the burst energies not be unreasonably large. Reports of achromatic breaks in afterglow spectra from some long duration bursts, which can be formed when Γ decreases to of order the reciprocal of the jet opening angle, support this view.

3.2 Magnetic Rotators

A reasonable guess is that all of these ultrarelativistic outflows (if not the more general class of cosmic jets) are due to similar physical processes. However,

[f] Seungmu dancers are able to beat a drum and, simultaneously, spin long tassles attached to their hats.

[g] It is well worth searching for the stellar counterparts of ultrarelativistic jets - Galactic blazars that are beamed towards us.

although popular models of AGN and plerionic jets are essentially electro- or hydromagnetic, most GRB models invoke a "hot big bang" that produces a fireball with a very high entropy per baryon, just like the early universe. I would like to explore the alternative hypothesis that all three types of source are really quite similar and derive their power from the continuous extraction of rotational energy from a compact object by electromagnetic stress - a "cold, steady state" model instead of a hot, big bang! [h] Furthermore, I suppose that the energy remains mostly in an electromagnetic form as it is transported into the emission region.

This magnetic rotator model is best developed in the case of plerions, where it is supposed that the central spinning neutron star possesses an inclined dipole moment and that it is surrounded by a force-free magnetosphere through which currents flow and space charge is maintained [Hirotani]. The complete electrodynamical description of this magnetosphere remains an unsolved problem. However, it seems likely that somewhere beyond the light cylinder, the electromagnetic field becomes essentially axisymmetric and that variation on the scale of a wavelength dies away. There are at least three ways by which this can occur. There can be steady reconnection in the outflowing, "striped" wind. Alternatively, the waves can decay through parametric instability into higher frequency waves. These two processes are essentially dissipative. Finally, the minority of magnetic field lines that emanate from the neutron star's southern magnetic pole, and which can be traced into the northern hemisphere, may gradually be pulled by magnetic tension across the equatorial plane back into the southern hemisphere (and *vice versa*). This can happen non-dissipatively. In summary, I will presume that only the DC, not the AC component of the electromagnetic field survives.

If this simplification of the magnetic geometry actually takes place, we will be left with a relativistic wind containing a predominantly toroidal magnetic field spun off by the central body. Associated with this toroidal magnetic field will be a poloidal electrical field that is almost equal in magnitude. (Many authors have argued that this wind quickly becomes plasma-dominated and terminates in a relativistic, fluid shock front in the inner nebula. However, it is very hard to see how the DC magnetic field can be erased so quickly and, although there are moving "wisp" features, these appear to lie in the equatorial plane. There is really no evidence for a relativistic shock in tne nebula.)

There are several suggestions as to how the jets associated with extra-

[h] It is amusing that, if we are correct in identifying long duration GRBs with compact objects of size $\sim$ 10 km, then their sources are observed to be active for a million light crossing times – an order of magnitude greater than the number of crossing times that we have observed a typical quasar!

galactic radio sources are launched. One of the simplest is that the power for the ultrarelativistic outflow is extracted as Poynting flux from the spin of the hole by magnetic field lines that are supported by external currents flowing the disk. It is also possible to extract energy electromagnetically from a surrounding accretion disk assuming that enough of its area is threaded by open magnetic field. However, this is most likely to produce a hydromagnetic outflow where the terminal velocity is no more than mildly relativistic although it may be responsible for collimating the much faster flow from the black hole.

When we consider electromagnetic models of GRBs, we find several advantages. The most fundamental is that the electromagnetic stress tensor, is anisotropic, in contrast to an isotropic fluid pressure tensor. This, implies that electromagnetic outflows can be naturally self-collimating. In addition, the presence of a dominant electromagnetic field implies that the effective internal sound speed is that of the fast mode which is close to the speed of light. Electromagnetic jets, in contrast to fluid jets are no more than mildly hypersonic. [i]

They are few direct clues as to the prime movers of GRBs [12]. The closest model to that of AGN jets has a spinning, stellar black hole surrounding by a stellar mass torus - possibly a tidally destroyed neutron star - that confines a $\sim 10^{14}$ G magnetic field [van Putten]. The emission lasts as long as the torus survives. The model that is closest to plerions involves a rapidly-spinning magnetar that has just been formed by accretion-induced collapse of a white dwarf. Alternatively, it might be possible for an accreting magnetised neutron star that is collapsing under rotational support to form a black hole or for two merging pulsars to act as magnetic rotators [j]

The common feature of all of these putative sources is that they spin off a toroidal magnetic field and an associated, electromagnetic Poynting flux that is unburdened by baryons.

3.3 Electromagnetic Black Holes

The idea that the spin energy of a black hole can be extracted electromagnetically has received an observational boost from the discovery that black holes

[i] Contrast this with the hypernova model of GRBs [12] [Maeda, Lee] where it is supposed that a high entropy per baryon fluid is collimated by a vortex inside a star so that the pairs and γ-rays eventually transfer their energy to protons and the ratio of the momentum flux to the pressure is $> 3\Gamma^2 \sim 3 \times 10^5$. This seems very unlikely to be true of a naturally occuring explosion, especially in the region where internal shocks are supposed to be operating.

[j] In principle, it is possible that an electromagnetic jet can be formed inside a collapsing star, though the star is not needed to provide the collimation and it seems hard to believe that plasma can be excluded from the outflow as efficiently as required.

are commonplace on both the stellar and the massive scale (as well as, perhaps, on the intermediate scale) and that the second parameter, the spin is large so as to allow gas to orbit close to the horizon and to form strongly redshifted emission lines [20].

There are several ways through which the rotational energy associated with the spinning spacetime can be tapped electromagnetically [5]. The particular choice that I have emphasised, because I believe that it represents the dominant energy channel, is that the horizon is threaded by a large flux of open magnetic field [Park]. A continuous, electromagnetic Penrose process operates in the ergosphere of the black hole which results in Poynting flux flowing inward across the horizon and, simultaneously, propagating away from the hole to infinity. [k] The source of the power is ultimately the reducible mass of the hole, against which the electromagnetic field in the ergosphere is doing work.

However, not all the field lines that thread the horizon need connect with the outflow. Some low latitude field lines may connect directly to the accretion disk and provide a supplementary power source for the disk as well as a possible driving torque for exciting quasi-periodic oscillations [Dotani]. This energy channel could be important, especially if the disk is thick. However, it is unlikely to lead to an ultrarelativistic outflow. [l]

The process that I have just described is distinct from (though can operate simultaneously with) an alternative process, the extraction of binding energy from the accreting gas both in the disk and in the plunging region between the inner edge of the disk and the horizon. The extra power that this process produces can be charged to the spin energy of the hole, which increases at a slower rate than it would do so in the absence of magnetic stress. However the intermediate working substance that effects this transformation is the inertia of the plasma not the electromagnetic field. There are three reasons for emphasising direct extraction of energy from the hole to extraction from the infalling gas at least for a rapidly spinning hole. The first is that the event horizon has a larger effective area than the annular ring between the hole and the disk. The second is that any gas-driven outflow is likely to be contaminated with baryons and consequently is unlikely to achieve an ultrarelativistic outflow velocity required. The third is that holes probably rotate much faster

[k] Energy flux is conserved in Boyer-Lindquist coordinates which becomes singular at the horizon and so power appears to emerge from the horizon in the Boyer-Lindquist frame. However, physical observers must orbit with respect to this coordinate system. Doppler boosting the energy flux into a frame moving with a physical observer produces an inwardly directed energy flux.

[l] The magnetic connection of the disk to the plunging gas seems to be a less promising source of power because the magnetic field lines quickly reconnect leaving the gas effectively disconnected from the disk.

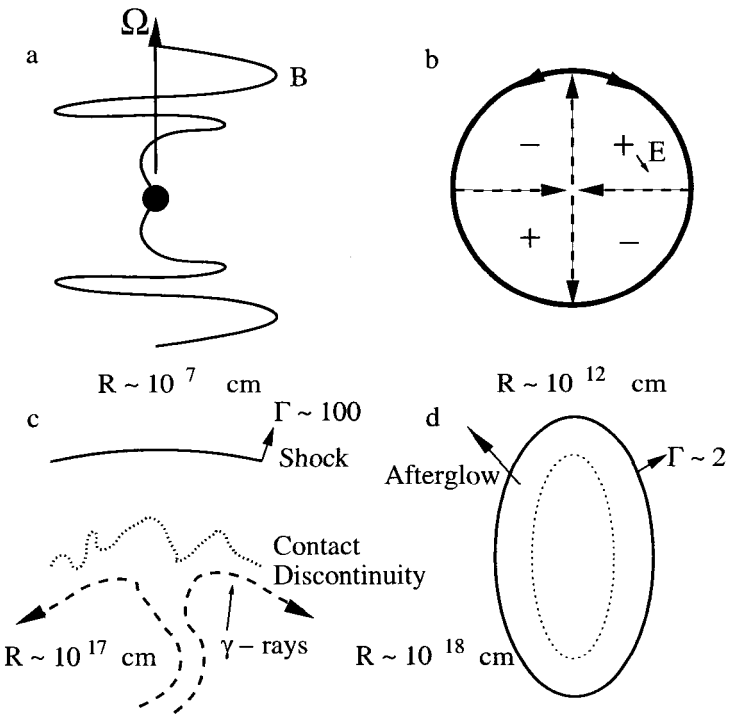


Figure 1: Four stages in the expansion of a magnetic shell with scales appropriate to a long duration GRB. a) The magnetic field $\vec{B}$ changes from poloidal to toroidal close to the outgoing light surface of the magnetic rotator at a radius $R \sim 10^6$ cm. The alternating component of the electromagnetic field decays relative to the DC toroidal field. b) The source is active for ~ 100 s. By this time, it will have inflated a magnetic bubble with radius $R \sim 3 \times 10^12$ cm, expanding with Lorentz factor $\Gamma \sim 3 \times 10^4$. The magnetic field is mostly toroidal, with the signs shown, while the electric field $\vec{E}$ is poloidal. The quadrupolar current flow is shown dashed. The shocked circumstellar medium is compresed into a thin shell of thickness $\sim 10^3$ cm. c) By the time the shell has expanded to $R \sim 10^17$ cm $\Gamma \sim 100$ and most of the electromagnetic pulse has caught up with the blast wave. This phase is observed a time ~ 100 s after the initial explosion. The current flow is still largely quadrupolar, though it is unstable along the axis and the equator and this drives an electromagnetic turbulent cascade. which ultimately creates electrical resistance and dissipation in the form of pair production, particle acceleration and intermittent, γ-ray emission. These instablities also promote corrugation of the contact discontinuity and incorporation of the magnetic field into the shocked interstellar medium where it can mix with relativistic electrons accelerated at the bounding shock front. d) When the blast wave has expanded by a further factor ten, its speed is only mildly relativistic. Its shape will be quite prolate as the expansion is fastest along the axis. Most of the energy released by the central, spinning, magnetic rotator is now carried by the shocked interstellar medium.

than orbiting gas, except quite close to the horizon, from where the exrtraction of energy will be quite inefficient.

3.4 Electromagnetic Shells

Suppose that a magnetic rotator spins off magnetic flux into the far field for a time $t_{\rm source}$ and that this creates a relativistically expanding shell of electromagnetic field that drives a blast wave into the surrounding medium. The blast wave is supposed to be bounded on its outside by a strong shock front that moves with Lorentz factor Γ and, on its inside, by a contact discontinuity, separating it from the shell, that moves with Lorentz factor Γ_c. Suppose further, for simplicity, that the current flow well beyond the light cylinder along the axes. then flows along the contact discontinuity and finally returns to it source through the equatorial plane. If we ignore the poloidal component of the magnetic field, (and consequently the flux of angular momentum), then the relevant solution of the force-free equations [Park] associated with this current flow has the form

$$B_\phi = \frac{f_+(t-r) + f_-(t+r)}{r\sin\theta} \tag{1}$$

$$E_\theta = \frac{f_+(t-r) - f_-(t+r)}{r\sin\theta} \tag{2}$$

The two terms in each expression are fast modes propagating outward and inward. The charge and current density vanish in the interior of the shell. We can determine the functions f_+, f_- by specifying the electromagnetic field at some small radius beyond the light surface and by matching to an ultrarelativistic blast wave expanding into the surrounding medium. This last requires that the outer surface of the shell move at the same speed as the inner surface of the blast wave and that the magnetic stress normal to this surface match the pressure in the blast wave.

The simplest assumption to make is that the strength of the magnetic rotator is constant in time, ($f_+ = \text{const}$) and that the external density is constant in radius. These assumptions imply that the Lorentz factor of the blast wave's outer shock front varies with radius R according to $\Gamma \propto \csc\theta R^{-1/2}$ [4]. (The Lorentz factor of the contact discontinuity, Γ_c, exhibits a similar variation.) The electromagnetic velocity (the velocity of the frame in which the electric field vanishes) in the body of the shell $\vec{\beta} = \vec{E}\times\vec{B}/B^2$ is radial and equal in magnitude to $(f_+ - f_-)/(f_+ + f_-)$ and the magnetic stress in a frame moving with this velocity is $\propto f_+ f_- \csc^2\theta$. The expansion of the blast wave is anisotropic, being faster along the poles, giving an electromagnetic power

per steradian $L_\Omega \propto \csc^2\theta$. Note that information is propagated inward by the fast mode which does modify the solution and allows it to react to changes in the load including changes in its effective impedance. Note also that it takes a very long time for a wave to be reflected by the blast wave and return to the origin. This is generally true of ultrarelativistic flows and stationary solutions which take a long time to be established can be quite misleading.

A comparison with a non-relativistic plerion like the Crab Nebula, is instructive. The magnetic bubble expands with speed of only $\dot{R} << c$ and, consequently, the pulsar must be producing magnetic flux at a rate that is roughly $(c/\dot{R})^{1/2}$ times too large to account for the strength of the magnetic field in the nebula. Therefore, most (95 percent in the case of the Crab Nebula) of the flux must be destroyed. On topological grounds, the natural places for this destruction to occur are on the axis and the equatorial plane. These regions are, in any case, formally unstable to pinch and tearing mode instabilities [m] respectively [2]. [n] The X-ray emission from the Crab Nebula, observed by Chandra [19] , may well be a consequence of particle acceleration associated with electrical resistance / flux destruction / generation of electromagnetic turbulence / particle acceleration tracing out the current flow in the inner part of the nebula. If we consider the flow of electromagnetic field in the nebula, we find that there must be a steady flow of magnetic flux towards the rotation axis and the equatorial plane [2]. In addition, there will be a compensatory backflow of relativistic electrons and positrons into the interior of the plerion. Of course there must be some matter in the nebula – it has been observed evaporating off filaments through its polarisation behaviour – but in this model, it has a very minor role in the dynamics.

By contrast, the contact discontinuity of an electromagnetic shell expands at a relativistic speed, and the reflected wave has a much smaller amplitude than the incident wave. Consequently, there is no necessity to destroy a lot of magnetic flux. Stated another way, there *need* be little resistance in the circuit. The effective load consists of the performance of work on the expanding blast wave. This is where most of the power that is generated by the central magnetic rotator ends up. (The distinction between this inertial load and a dissipative load is quite unimportant for the behavior of the black hole magnetosphere.)

The simple electromagnetic solution will only remain valid until the end of the outward-propagating, electromagnetic pulse catches up with the blast wave. This occurs at some radius $R_{\rm free} \sim \Gamma(R_{\rm free})^2 ct_{\rm source}$. Thereafter, the

[m] The contact discontinuity should be formally Kruskal-Schwarzschild stable.

[n] If we were to prevent this dissipation, in a thought experiment, through having the current flow along rigid, perfect conductors, then the reflected electromagnetic waves would react back on the source and shut down the power supply.

surrounding blast wave which, by now, has acquired almost all of the energy in the explosion, will expand freely with $\Gamma \propto R^{-3/2}$.

This elementary model of a relativistic electromagnetic shell can be easily generalized to accommodate different assumptions about the variation of the flux supply with time and latitude and the density variation in the external medium.

3.5 Gamma Ray Bursts

As well as bring out some formal points, the electromagnetic solution just described provides a possible model for GRBs [1811]. Suppose that a magnetic rotator spins off toroidal magnetic field as it slows down and that this magnetic field fills an anisotropic, relativistically expanding shell in a uniform medium. Suppose, further, that the flux distribution near the light cylinder is consistent with the current being concentrated along the axis and in the equatorial plane, as described above. The current density is most intense on the axis and, although there is no requirement that flux be destroyed as long as the expansion is relativistic, in practice the magnetic pinch will become hydromagnetically unstable to sausage and kink modes (in the comoving frame) after expansion beyond a radius where the stabilising, poloidal field becomes insignificant [2]. These global instabilities, which should have a longitudinal wavelength comparable in size to the width of the current distribution, may sustain an electromagnetic turbulence spectrum which should ultimately be responsible for particle acceleration and the excitation of transverse gyrational motion [17]. [o] The reason why particles are accelerated is that, when the power cascades down to short enough wavelengths, there are too few charged particles to carry the electrical current and field-parallel electric fields will develop. This is the microscopic source of the dissipation. These particles will emitγ-rays, primarily through the synchrotron process, though also by inverse Compton scattering. γ-rays, in excess of threshold will create fresh pairs. All of this will take place in a frame moving outward with the electromagnetic velocity and the emission will be strongly beamed outward. The global pinch instabilities also provide a plausible explanation for the large $\sim 1 - 10$ ms fluctuations in the observed γ-ray flux that are observed [p].

As well as create electrical resistance, these instabilities The afterglow is formed after the blast wave becomes free of its electromagnetic driver. Now, in most afterglow models, including those involving jets, it is supposed that the

[o] This turbulence may have already been seen in the measured fluctuation power spectrum [Chang].

[p] Attributing this variation to internal shocks tied to the source, as in the fluid model, implies that the γ-rays originate at much smaller radii than expected on the electromagnetic model.

expansion velocity does not vary with latitude. However, an electromagnetically-driven blast wave necessarily creates an anisotropic explosion and this has important consequences for observations of the afterglow, especially in the ultrarelativistic phase of expansion. If we continue to use our simple model, we find that the afterglow expansion varies most rapidly and remains relativistic for longest closest to the symmetry axis. As $L_\Omega \propto \csc^2\theta$ the energy contained in each octave of θ is roughly constant This means that the most intense bursts and afterglows in a flux-limited sample will be seen pole-on and should exhibit achromatic breaks, which might be mistaken for jets. The inferred explosion energy will be roughly independent of θ and characteristic of the total energy. When the expansion becomes non-relativistic, the remnant will have a prolate shape which might be measurable. [q]

This electromagnetic model provides a solution to the puzzle of how to launch a blast wave that extends over an angular scale $>> \Gamma^{-1}$ and where the individual parts are out of causal contact. In the electromagnetic model, the energy is transferred to the blast wave by a magnetic shell that pushes (unevenly) on the surrounding gas all the way out to $R_{\rm free}$. It also supplies an origin for the magnetic flux in the blast wave, for which the alternative origin in the bounding shock front seems very hard to explain. In the present model, magnetic field can simply be mixed into the blast wave (and the shock-accelerated relativistic electrons) at the contact discontinuity through instabilities, much like what seems to happen in regular supernova remnants.

3.6 Some Numbers

Let us give some illustrative orders of magnitude for a model of a long duration GRB. The electromagnetic energy flux near the pole is $L_\Omega \sim 10^{50}$ erg s^{-1} sterad^{-1} and lasts for a time $t_{\rm source} \sim 100$ s. The associated EMF in the electrical circuit $\sim 10^{22}$ V [r] and the current is $\sim 10^{20}$ A. [s] The external density is uniform $n \sim 1$ cm^{-3}. The blast wave is driven by the electromagnetic pulse with Lorentz factor $\Gamma \propto R^{-1/2}$ until $R \sim R_{\rm free} \sim 10^{17}$ cm, $\Gamma \sim \Gamma_{\rm free} \sim 100$. Thereafter there is a freely expanding blast wave with $\Gamma \propto R^{-3/2}$ until the expansion becomes non-relativistic when $R \sim R_{\rm NR} \sim 3 \times 10^{18}$ cm.

Most of the GRB emission (around ~ 1 MeV) is produced when $R \lesssim R_{\rm free}$ as synchrotron emission by ~ 100 GeV electrons in a co-moving magnetic field

[q] It is tempting to associate some of the barrel-shaped supernova remnants observed in our Galaxy with the remnants of electromagnetic explosions.

[r] A potential difference this large, made available along the contact discontinuity, provides one of the few astrophysical options for accounting for UHE cosmic rays.

[s] For comparison the values are $\sim 3\times10^{14}$ A, $\sim 3\times10^{16}$ V for the Crab Nebula and $\sim 10^{18}$ A, $\sim 10^{20}$ V for Cygnus A.

of strength $B \gtrsim 30$ G. The comoving cooling time of these electrons is ~ 3 s, a fraction $\lesssim 10^{-4}$ of the expansion timescale and so if the ~ 100 GeV pair energy density is maintained at a significant fraction of the equipartition energy density, then the magnetic energy can be efficiently transformed into γ-rays. The opacity to pair production for a γ-ray of energy E is $\sim 0.1(E/1\mathrm{GeV})$. The Thomson optical depth depends upon the details of the particle acceleration but is plausibly much smaller than unity so that the observed γ-rays can freely escape without erasing the variability.

3.7 AGN Jets

Having briefly discussed electromagnetic models of plerions and GRBs, it remains to re-consider extragalactic jets in this context [Fletcher]. Although electromagnetic/hydromagnetic models of the energy release have been quite widely discussed for a long while, it has generally been supposed that this electromagnetic energy is transformed into a particle-dominated flow at a safe distance from the black hole and that the observed emission is from a high β plasma. How plausible is it that the entire radio source is an electromagnetic structure, that the observed emission trace out unstable currents rather than strong shock fronts? [t] One of the merits of this hypothesis is that it may provide a dynamical rationalisation of the doctrine of equipartition. Pinched currents, on all scales, may continue to become unstable until their stresses are balanced by pressure. Another merit is that it provides a natural explanation for the helical structures that are often seen in VLBI maps. A third advantage is that currents can account for distributed particle acceleration in well-resolved jets, as spectral studies suggest may be required. It is natural ask if the extended lobes associated with the powerful FRII sources like Cygnus A are filled with unstable though fundamentally toroidal and force- free magnetic field. , then this could present a quite strong signal in the Faraday rotation maps which has not really been seen. Similarly the polarisation structure of FR2 radio sources does look like a shear flow and there are some features, in radio maps, like the knots in M87, which are naturally interpreted as strong shocks.

4 Summary and Prospects

In these impressionistic sketches, I have outlined two principles, which although not completely new, are now not commonly discussed – that most accreting

[t] Of course, under force-free conditions, the currents also trace out the magnetic field in the frame in which the electric field vanishes.

black holes (excluding those where the mass supply is within an order of magnitude or so of the Eddington rate) swallow only a tiny fraction of the gas supplied at the Bondi radius and that ultrarelativistic, high energy phenomena are fundamentally electromagnetic not gas dynamical. The link between these two principles may be that carrying off the binding energy of the accreting gas in a fluid outflow, rather than radiation, is a necessary condition for extracting energy electromagnetically from a spinning black hole. At least this is the story in the case of AGN jets. This condition is also satisfied by pulsars. GRBs, which may be either accreting black holes or neutron stars, would be similar to either. An extension of this line of argument is that AGN jets may be essentially electromagnetic all the way to their hot spots.

There are many possible discriminatory observations. For the most interesting and immediate case of GRBs, it is not predicted that there will be an accompanying neutrino signal. By contrast, a gravitational wave signal is expected in some, though not all, models and would be strongly diagnostic if detected. It is unlikely, though not impossible, that GRBs will be associated with Type II supernovae. More immediate but less specific diagnostics include relating the duration and character of the GRB with the inferred observation angle of the burst – higher inclination should be associated with longer and less intense bursts. The spectrum and polarisation of the afterglow emission might also contain some clues, though the lack of a usable theory of particle acceleration [u] and magnetic field amplification at ultrarelativistic shocks makes this a bit problematic. A more detailed discussion of the GRB emission, than presented here should account for the MeV breaks observed in γ-ray spectra.

From a more theoretical perspective, there is much to be learned about the properties of force-free electromagnetic fields and especially their stablity. The possible relationship of the GRB fluctuation power spectrum to an underlying turbulence spectrum is especially tantalising. Undoubtedly, numerical simulations will be crucial as the problem is essentially three dimensional. Force-free electromagnetism is easier to study than relativistic MHD and may well be a very good approximation in many of these sources.

Turning to plerions, the most direct, observational challenge is to see if there really is a strong, dissipating shock as expected with a fluid wind or a flow of electromagnetic energy towards the axis and the equatorial plane as predicted by the electromagnetic model and as appears to be exhibited by the Crab Nebula.

[u] Recent, promising progress on understanding relativistic shocks predicts a power law distribution function with a logarithmic slope of 2.2[9], provided that the scattering is essentially normal to the shock front. What is not yet clear is whether these scattering conditions are present.

Finally, for AGN, we would like to detect and understand the mass outflows predicted above for low mass supply rates. (We already know that rapidly accreting stellar and massive holes drive dense, high speed outflows.) In addition, we need to see if the very well-observed radio jets and their lobes can be re-interpreted in terms of an unstable Z-pinch. A good place to start is through mapping the magnetic field using polarisation measurements and the internal mass density through internal depolarisation data. In addition, detailed imaging spectra from radio to X-ray energies can be used to determine where the particles are being accelerated – at shock surfaces or in volumes containing strong, unstable currents – and how the propagate away from these acceleration sites at different energies.

The conjunction of recent, impressive discoveries, upcoming observational facilities and powerful computing capability makes this a propitious time to be studying current high energy emission from black holes.

Acknowledgments

I thank the organisers of this workshop for their gracious and generous hospitality, my colleagues Mitch Begelman and Max Lyutikov for their collaboration on the above and several of the conference participants for discussion which helped to clarify some of the arguments I have presented. Support under NASA grant 5-2837 is gratefully acknowledged.

References

1. Balbus, S. A. & Hawley, J. F. RMP **70**, 1 (1998)
2. Begelman, M. C. ApJ **493**, 291 (1998)
3. Begelman, M. C. ApJ , in press (2002)
4. Blandford, R. D. & McKee, C. F. Phys. Fluids **19**, 1130 (1976)
5. Blandford, R. D. Prog. Theor. Phys. Supp. **143**, 182 (2001)
6. Blandford, R. D. *Lighthouses of the Universe* ed. R. Sunyaev *et al* Berlin:Springer-Verlag (2002)
7. Blandford, R. D. & Begelman, M. C. MNRAS **303**, L1 (1999)
8. Frank, J. H., King, A. R. & Raine, D. J. *Accretion Power in Astrophysics* Cambridge: Cambridge University Press (1992)
9. Kirk, J. G., Guthmann, A. W., Gallant, Y. A. & Achterberg, A. ApJ **542**, 235 (2000)
10. Krolik, J. H. *Active Galactic Nuclei* Princeton: Princeton University Press (1999)
11. Lyutikov, M. & Blackman, E. G. MNRAS **321**, 177 (2001)
12. Mészáros, P. ARAA , in press (2002)

13. Narayan, R., Mahadevan, R. & Quataert, E. *The Theory of Black Hole Accretion Discs* ed. M. A.. Abramowicz, G. Bjornsson & J. E. Pringle Cambridge: Cambridge Univeristy Press p148 (1998)
14. Shakura, N. I. & Sunyaev, R. A. Astron. Astrophys **254**, 22 (1973)
15. Stone, J. M., Pringle, J. E. & Begelman, M. C. MNRAS **310**, 1002 (1999)
16. Stone, J. M. & Pringle, J. E. MNRAS **322**, 461 (2001)
17. Thompson, C. & Blaes, O. M. *Phys. Rev.* D **57**, 3219 (1998)
18. Usov, V. V. MNRAS **267**, 1035 (1994)
19. Weisskopf, M. *et al*ApJ **536**, 81 (2000)
20. Wilms, J., Reynolds, C. S., Begelman, M. C., Reeves, R., Molendi, S., Staubert, R. & Kendziorra, E. MNRAS **328**, L27 (2001)

PAIR PRODUCTION CASCADE IN BLACK-HOLE MAGNETOSPHERE

Kouichi HIROTANI
Code 661, Laboratory for High-Energy Astrophysics
NASA/Goddard Space Flight Center, MD 20771, USA
E-mail: hirotani@milkyway.gsfc.nasa.gov

Isao OKAMOTO
National Astronomical Observatory,
Osawa 2-21-1, Mitaka, Tokyo 181-8588, Japan
E-mail: okamoto@th.nao.ac.jp

We investigate a stationary pair production cascade in the magnetosphere of a rotating black hole. The charge depletion due to global flows of charged particles, causes a large electric field along the magnetic field lines. Migratory electrons and/or positrons are accelerated by this field to radiate gamma-rays via inverse-Compton scatterings. Some of such gamma-rays collide with the X-rays and materialize as pairs in the accelerator. The replenished charges partially screen the electric field, which is self-consistently solved, together with the distribution functions of particles and gamma-rays. By solving these Vlasov equations, we demonstrate that sufficient amount of plasmas can be supplied for the Blandford-Znajek process to work effectively in the magnetosphere of a super-massive black hole residing in an active galactic nucleus.

1 Introduction

The model of a stationary, plasma-filled magnetosphere around a rotating black hole and/or its accretion disk has been evoked, to explain the extraction of energy in the vicinity of a black hole and the formation of relativistic jets (Blandford 1976, 1977; Lovelace 1976). ¿From recent observations of γ-ray emissions from AGNs, the Blandford-Znajek process is suggested to be a viable mechanism for energizing the relativistic jets (see e.g., Montigony et al. 1995 for recent review of observations).

In the γ-ray emitting region, jets must be dominated by the kinetic energy flux of $e^{\pm}$ plasma over the electromagnetic flux originating from the central engine. A pair-cascade model of γ-ray jets has been proposed by Blandford (1993), Blandford and Levinson (1995; BL95), and Levinson and Blandford (1996). Their bulk jet dynamics for converting the Poynting flux to the kinetic energy is as follows: Once the bulk motion of the jets attains the Lorentz factor Γ larger than the threshold energy above which the opacity to pair production on the background photons exceeds unity, copious pair production would ensue, leading to a sharp increase of the inertia and radiative drag acting

upon the $e^{\pm}$ outflow. This is a phenomenon which is expected to occur above the annihilation radius $r_{\rm ann}$ for the $e^{\pm}$ plasma (BL95).

On the other hand, in the region well below $r_{\rm ann}$ where rapid annihilation is dominating, the jet energy may well be regarded predominantly electromagnetic. Levinson (1996) examined the structure of the "inner jet", where the Poynting flux is being converted to the kinetic flux. By numerically solving the equation of motion for the bulk jet coupled to the appropriate kinetic equations governing the evolution of $e^{\pm}$ pair and γ-ray number densities, he showed that for steep spectra, such as the standard spectrum invoked by BL95, the acceleration scenario predicts that the majority of the power extracted from the central engine will be emitted as soft X-rays rather than being converted into pairs and high-energy γ-rays as required by the observations. The spectrum incident upon the inner jet must be sufficiently flat.

The source of the electromagnetic power eventually resulting in the γ-ray jet may be the spin of the central black hole (Blandford 1993). It is however the electric current that sustain the electromagnetic power in the field-dominated region of the magnetosphere, and it is charged particles, even though inertia-free, that carry the electric current. In other words, energy and angular momentum are transported out of the horizon by the electric current which can freely flow into and out of the horizon. However, even though under the assumption of massless, it must be real charged particles that constitute the electric current, and because no particle can classically escape beyond the horizon, electrons must flow into the horizon for the electric current to flow out, and positrons must flow into there for the electric current to flow in. We thus have to contrive a process of plasma supply deep within the magnetosphere somewhat different from the process of converting the Poynting flux into pairs and γ-rays in the jet much above the horizon.

We assume a stationary axisymmetric degenerate force-free magnetosphere around a rotating black hole. The absolute space around the hole with the mass M and the angular momentum J is described in a Boyer-Lindquist spatial coordinate system with the two scalar functions α and ω (Macdonald & Thorne 1982; Thorne et al. 1986):

$$ds^2 = \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2 + \varpi^2 d\phi,$$

where

$$\rho^2 = r^2 + a^2 \cos^2\theta, \quad \Delta = r^2 - \frac{2GMr}{c^2} + a^2,$$

$$\Sigma^2 = (r^2 + a^2)^2 - a^2 \Delta \sin^2\theta, \quad \varpi = \frac{\Sigma}{\rho} \sin\theta,$$

and

$$\alpha = \frac{\rho\Delta^{1/2}}{\Sigma}, \quad \omega = \frac{2aGMr}{c\Sigma^2}; \tag{1}$$

a is the spin parameter defined by $a \equiv J/Mc$. The horizon radius is denoted with $r_{\rm H} = GM/c^2 + \sqrt{(GM/c^2)^2 - a^2}$. Here α is the lapse function or the redshift factor, and ω is the angular velocity of so-called zero angular momentum observers (ZAMOs), which vanishes at infinity ($\alpha \to 1$) and coincides with the uniform rotation of the hole, $\omega_{\rm H} \equiv c^3 a/(2GMr_{\rm H})$ at the horizon ($\alpha = 0$). The general relativistic effects upon the force-free magnetosphere appear through α and ω.

Denoting the magnetic flux function with Ψ, one can define the poloidal magnetic field in terms of Ψ

$$\boldsymbol{B}_{\rm P} = -\frac{\boldsymbol{e}_\phi \times \boldsymbol{\nabla}\Psi}{2\pi\varpi}, \tag{2}$$

where $\boldsymbol{e}_\phi$ is the unit toroidal vector. In the axisymmetric system the electric field is purely poloidal. Assuming the frozen-in condition, we obtain

$$\boldsymbol{E}_{\rm P} = -\frac{\boldsymbol{v}_{\rm F}}{c} \times \boldsymbol{B}_{\rm P} = -\frac{\Omega_{\rm F} - \omega}{2\pi\alpha c}\boldsymbol{\nabla}\Psi \tag{3}$$

where $\boldsymbol{v}_{\rm F} = [(\Omega_{\rm F} - \omega)\varpi/\alpha]\boldsymbol{e}_\phi$ is the rotational velocity of field lines measured by ZAMOs, and $\Omega_{\rm F}(\Psi)$ is the angular velocity of field lines.

The charge density in the degenerate force-free magnetosphere is given by the Goldreich-Julian charge density, i.e.

$$\rho_{\rm GJ} \equiv \frac{1}{4\pi}\boldsymbol{\nabla}\cdot\boldsymbol{E}_{\rm P} = -\frac{1}{4\pi}\boldsymbol{\nabla}\cdot\left(\frac{\Omega_{\rm F} - \omega}{2\pi\alpha c}\boldsymbol{\nabla}\Psi\right) \tag{4}$$

everywhere except within the thin gap of which structure is our main concern. Since $\boldsymbol{v}_{\rm F}$ and hence $\boldsymbol{E}_{\rm P}$ changes sign at the surface where ω equals $\Omega_{\rm F}$, one can anticipate that $\rho_{\rm GJ}$ also changes sign in the vicinity of this surface. Some analyses reveal that $\rho_{\rm GJ}$ is negative far from the hole and positive near the hole along field lines near the rotation axis.

In pulsar magnetospheres, it is known that the null-charge surface, where $\rho_{\rm GJ}$ changes its sign, is the surface around which a vacuum accelerator can arise, provided that thre is no particle injection across either of the boundaries (Hirotani & Shibata 2001a,b; 2002). In this paper, we assume no particle injection; therefore, an accelerator can be expected to be located at null surface also in a black-hole magnetosphere.

A typical example of the distribution of the null surface is depicted in Fig. 1. As the figure indicates, the null surfaces (solid lines) nearly coincide with

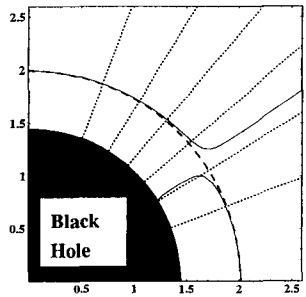


Figure 1: Distribution (side view) of the 'null surface' (solid lines), where ρ_{GJ} vanishes around a rapidly rotating hole ($a = 0.9$ and $\Omega_F = 0.5\omega_H$). The magnetic field lines (dotted lines) are assumed to be radial near the horizon. In this figure, we adopt $\Psi \propto \sin^2\theta$. The null surface almost coincides with the surface of $\omega = \Omega_F$ (dashed line) except in the middle latitudes.

the surface of $\omega = \Omega_F$ (dashed line). The magnetic field lines (dotted lines) are assumed to be radial in this figure, because the force-free transfield equation makes the field line structure be asymptotically radial for $\alpha \to 0$ (Macdonald & Thorne 1982; Okamoto 1992). For illustration purpose, Ψ is assumed to be proportional to $\sin^2\theta$.

In spite of the core significance in feeding relativistic γ-ray jet from the central engine, the real process of supplying charged particles in the magnetosphere has not been substantiated for a realistic model of radiation field in AGNs until Beskin et al. (1992). For the typical parameters of the central radiation in AGN, they estimated the width of the region of plasma production and the particle energies in it.

The purpose of this paper is to further extend the analysis of Beskin et al. (1992) and to clarify more quantitatively the micro process of interaction of radiation – pair creation under the existence of parallel electric field.

In the next section, we formulate basic equations describing a pair production cascade in the magnetosphere. We then solve them in §3 and demonstrate that sufficient amount of plasma is supplied by the cascade so that hole's rotational energy may be extracted effectively. In the final section, we sum up the results and discuss the differences from Beskin et al. (1992).

2 Pair Production Cascade Mechanism

In this section, we describe how a pair production cascade proceeds in a thin gap in the force-free magnetosphere. We first discuss physical processes in the potential gap leading to a cascade in § 2.1. We next formulate basic equations

in § 2.2, discuss non-gray analysis of the γ-ray distribution in § 2.3, and impose boundary conditions in § 2.4.

2.1 Saturated Lorentz Factor

Let us first consider the Poisson's equation that describes the acceleration field, $E_{\parallel} \equiv \boldsymbol{E}_{\rm P} \cdot \boldsymbol{B}_{\rm P}/|\boldsymbol{B}_{\rm P}|$, in the presence of the rotationally induced charge density, $\rho_{\rm GJ}$. If the transfield thickness of the potential gap is large compared with the longitudinal width, we can approximated the divergence of the acceleration field into one-dimension form and obtain

$$\frac{dE_{\parallel}}{dx} = 4\pi[e(n^{+} - n^{-}) - Ax], \tag{5}$$

where $A \approx \Omega_F B/(4\pi ce)$ is the expansion coefficient of $\rho_{\rm GJ}$ at the null surfaces ($x = 0$), which are indicated by the solid lines in fig. 1. For a spherical null surface, we have $A = \partial_r(\rho_{\rm GJ})$ at $x = 0$. If the real charge density $e(n^{+} - n^{-})$ deviates from $\rho_{\rm GJ} = Ax$, an acceleration field, $E_{\parallel}$, is induced along the magnetic field lines. Migratory particles are accelerated by $E_{\parallel}$ and upscatter soft photons (emitted by the accretion disk, for instance). The resultant γ-ray photons collide with the soft photons, to materialize as pairs, which are accelerated by $E_{\parallel}$ to maintain a pair-production cascade (Blandford & Znajek 1977).

In the potential gap, the $e^{\pm}$'s lose their perpendicular momentum rapidly by upscattering ambient UV photons. However, their longitudinal motion will be maintained by the electrostatic acceleration due to $E_{\parallel}$ field. Therefore, the motion of a single e^{+} or e^{-} can be approximated one-dimensionally and saturates at the Lorentz factor

$$\Gamma(x) = \sqrt{1 + eE_{\parallel}(x)/\sigma_{\rm T}U_{\rm b}} \qquad \approx 5{\times}10^{3} \left(\frac{E_{\parallel}(x)}{10^{3}{\rm V/m}}\right)^{1/2} \left(\frac{U_{\rm b}}{10^{6}{\rm ergs/cm^{3}}}\right)^{-1/2}, \tag{6}$$

where $U_{\rm b}$ refers to the energy density of back ground soft photons.

2.2 Particle Continuity Equation

Let us first consider the continuity equations of $e^{\pm}$'s with spatial number densities $n^{\pm}(x)$. Without loss of any generality, we can assume $E_{\parallel} > 0$; in this case, e^{+}'s (or e^{-}'s) migrate outwards (or inwards). Since particle velocities are virtually $\pm c$, we obtain

$$\pm c\frac{dn^{\pm}}{dx} = \int_{0}^{\infty} \eta_{\rm p}(\epsilon_\gamma) \left[F^{+}(x,\epsilon_\gamma) + F^{-}(x,\epsilon_\gamma)\right] d\epsilon_\gamma, \tag{7}$$

as the continuity equaiton. Here, the angle-averaged pair production rate, $\eta_{\rm p}$, is defined in e.g., Berestetskii et al. (1989). $F^{\pm}(x,\epsilon_\gamma)$ are the distribution functions of the γ-rays propagating in $\pm\, x$-directions, respectively, at a position x and in the dimension-less energy interval $\epsilon_\gamma \sim \epsilon_\gamma + d\epsilon_\gamma$; $dN_{\rm s}/d\epsilon_{\rm s}$ refers to the number density of background soft photons in the dimension-less energy interval $\epsilon_{\rm s} \sim \epsilon_{\rm s} + d\epsilon_{\rm s}$. It must be noted that the γ-ray photons, which are produced by the inverse Compton scatterings, are highly beamed in the same direction of $e^{\pm}$'s one dimensional motion. That is, their distribution functions can be fully described in terms of F^{+} and F^{-}.

We can easily see that the current, which is carried by $e^{\pm}$'s along a given field line is conserved along x. ¿From one combination of equations (7) we have

$$\frac{d}{dx}\left[n^{+}(x) + n^{-}(x)\right] = 0, \tag{8}$$

which yields

$$n^{+}(x) + n^{-}(x) = j_0/e \tag{9}$$

where the current density j_0 is constant along a field line. In order that the energy and angular momentum may be extracted effectively from a rotating supermassive black hole, j_0 must take the value of

$$j_0 \sim 10^{-13}\left(\frac{a}{M}\right)\left(\frac{M}{10^8 M_\odot}\right)\left(\frac{B}{10^4{\rm G}}\right)\frac{\rm abampere}{\rm cm^2}. \tag{10}$$

in order of magnitude (e.g. Thorne et al. 1986).

2.3 Gamma-ray Boltzmann Equation

We next derive the Boltzmann equations for the γ-ray photons. As we noted in the previous subsection, the γ-ray photons are scattered only in the $\pm x$-direction. Thus the γ-ray distribution functions, $F^{\pm}$, obey

$$\pm c\frac{d}{dx}F^{\pm}(x,\epsilon_\gamma) = \eta_{\rm c} n^{\pm}\sqrt{1-1/\Gamma^2} - \eta_{\rm p}F^{\pm}, \tag{11}$$

where $\eta_{\rm c}(\epsilon_\gamma,\Gamma)$ is the Compton redistribution function [$\rm s^{-1}$] defined by

$$\eta_{\rm c} \equiv c\int_{\epsilon_{\rm min}}^{\epsilon_{\rm max}} d\epsilon_{\rm s}\frac{dN_{\rm s}}{d\epsilon_{\rm s}}\sigma_{\rm KN}\delta(\epsilon_\gamma - \Gamma^2\epsilon_{\rm s}); \tag{12}$$

$\sigma_{\rm KN}$ is the Klein-Nishina cross section (e.g., Rybicki& Lightman 1979).

Let us assume a power-law spectrum for the soft photon field such that

$$\frac{dN_{\rm s}}{d\epsilon_{\rm s}} = C(\alpha)\epsilon_{\rm s}^{-\alpha}, \tag{13}$$

where $C(\alpha)$ is a decreasing function of α and is defined by

$$C(\alpha) \equiv \frac{2-\alpha}{\epsilon_{\rm max}^{2-\alpha} - \epsilon_{\rm min}^{2-\alpha}} \frac{U_{\rm b}}{m_{\rm e}c^2}; \tag{14}$$

$\epsilon_{\rm max}$ and $\epsilon_{\rm min}$ are the cutoff energy of the above spectrum. In what follows, we adopt $\epsilon_{\rm max} = 30{\rm keV}/m_{\rm e}c^2 = 5.87\times10^{-2}$ and $\epsilon_{\rm min} = 10{\rm eV}/m_{\rm e}c^2 = 1.95\times10^{-5}$. The energy density of the soft photon field can be estimated as

$$U_{\rm b} \sim \frac{L_{\rm b}/c}{4\pi(5r_{\rm H})^2} = 1.47\times10^6 \left(\frac{M}{10^8 M_\odot}\right)^{-1} \left(\frac{L_{\rm b}}{L_{\rm Edd}}\right) {\rm ergs/cm^3}, \tag{15}$$

where $L_{\rm b}/L_{\rm Edd}$ is the luminosity of the background radiation field normalized by the Eddington luminosity $L_{\rm Edd} = 1.25\times10^{46}(M/10^8 M_\odot)$ ergs/s.

Let β_i and β_{i-1} be the upper and lower limit of the i-th dimension-less energy bin. Then taking β_i to be sufficiently close to β_{i-1}, one can approximate integrals in the right hand side of equation (7) with the summation of the following integrals over each bin

$$\int_{\beta_{i-1}}^{\beta_i} \eta_{\rm p}(\epsilon_\gamma) F^{\pm}(x,\epsilon_\gamma) d\epsilon_\gamma \approx \eta_{{\rm p},i} f_i^{\pm}(x), \tag{16}$$

where

$$\eta_{{\rm p},i} \equiv \eta_{\rm p}\left(\frac{\beta_{i-1}+\beta_i}{2}\right), \tag{17}$$

$$f_i^{\pm}(x) \equiv \int_{\beta_{i-1}}^{\beta_i} F^{\pm}(x,\epsilon_\gamma) d\epsilon_\gamma. \tag{18}$$

Then, instead of equation (7), we use

$$\pm c\frac{dn^{\pm}}{dx} = \sum_{i=1}^{m} \eta_{{\rm p},i} \left[f_i^{+}(x) + f_i^{-}(x)\right], \tag{19}$$

where m refers to the number of energy bins. Likewise, equations (11) become

$$\pm c\frac{d}{dx} f_i^{\pm}(x) = \eta_{{\rm c},i}(\Gamma) n^{\pm}(x) - \eta_{{\rm p},i} f_i^{\pm}(x), \qquad (i = 1,2,\cdots,m) \tag{20}$$

where

$$\eta_{{\rm c},i}(\Gamma) \equiv \int_{\beta_{i-1}}^{\beta_i} \eta_{\rm c}(\epsilon_\gamma,\Gamma) d\epsilon_\gamma. \tag{21}$$

To investigate the pair production mechanism, it is sufficient to consider γ-ray photons with energies (dimension-less) satisfying

$$\epsilon_\gamma > \frac{2}{1-\mu}\frac{1}{\epsilon_s}, \tag{22}$$

of which minimum value is $1/\epsilon_{max}$. Photons below this threshold energy $1/\epsilon_{max}$ never contribute to pair productions, and hence we may put the lower limit of the lowest energy bin, $\beta_0 m_e c^2$ to be $m_e c^2/\epsilon_{max} = 8.7$MeV (i.e., $\beta_0 = 1/\epsilon_{max}$). For the present purpose, it is actually enough to divide the range of ϵ_γ into 13 bins and set $\beta_1 = 1.5/\epsilon_{max}$, $\beta_2 = 2/\epsilon_{max}$, $\beta_3 = 2.5/\epsilon_{max}$, $\beta_4 = 3/\epsilon_{max}$, $\beta_5 = 4/\epsilon_{max}$, $\beta_6 = 5/\epsilon_{max}$, $\beta_7 = 10/\epsilon_{max}$, $\beta_8 = 20/\epsilon_{max}$, $\beta_9 = 30/\epsilon_{max}$, $\beta_{10} = 40/\epsilon_{max}$, $\beta_{11} = 50/\epsilon_{max}$, $\beta_{12} = 60/\epsilon_{max}$, and $\beta_{13} = 70/\epsilon_{max}$.

2.4 Boundary Conditions

To solve the set of ordinary differential equations (5), (19), and (20), we assume that $E_\parallel$, Γ, n^+, $f_i^\pm$'s have the following symmetric properties:

$$E_\parallel(x) = E_\parallel(-x), \tag{23}$$

and

$$n^+(x) = n^-(-x), \qquad F^+(x,\epsilon_\gamma) = F^-(-x,\epsilon_\gamma). \tag{24}$$

The solutions obtained under these symmetries will not miss the essential features of the pair production cascade in the black hole magnetosphere. In view of these symmetry properties, it is sufficient to solve the differential equations only in the range $0 \le x \le H$.

We impose the folowing boundary conditions:

$$2n_+ = j_0/e, \quad f_i^+ = f_i^- \quad \text{at } x = 0, \tag{25}$$

and

$$E_\parallel = 0, \quad n_+ = j_0/e, \quad n_- = 0, \quad f_i^- = 0 \quad \text{at } x = H, \tag{26}$$

where $i = 1, 2, \cdots, m$. We thus obtain total $(2m+4)$ boundary conditions for the $(2m+3)$ unknown functions $E_\parallel(x)$, $n^\pm(x)$, and $f_i^\pm(x)$ $(i = 1, 2, \cdots, m)$. In general, the excessive boundary condition should be compensated by treating H be free. For a small j_0, it follows from equation (5) that $dE_\parallel/dx = -4\pi AH$. However, as j_0 increases, $|dE_\parallel/dx|$ decreases to vanish at a certain value of j_0. Above this critical current density, j_{cr}, the magnetosphere is over filled by charged particles and the gap cannot be closed (see fig. 2 in Hirotani &

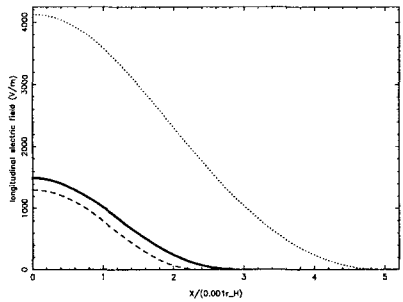


Figure 2: Examples of longitudinal electric field $E_{\parallel}(x)$. The abscissa is normalized by $10^{-3}r_{\rm H}$. The solid line represents the solution corresponding to $\alpha = 2.0$. The dashed and dotted lines to $\alpha = 1.5$ and 2.5, respectively.

Okamoto 1998). to investigate an active magnetosphere, we consider the case when j_0 is close to $j_{\rm cr}$. In this case, we obtain

$$dE_{\parallel}/dx \approx 0 \quad \text{at } x = H. \tag{27}$$

Under these $(2m + 5)$ boundary conditions, we can solve H and $j_{\rm cr}$, as well as $E_{\parallel}$, $n^{\pm}$, and $f_i^{\pm}$.

3 Structure of the Potential Gap

In the preceding section, we formulated the basic equations and described the procedure to solve them under suitable boundary conditions. Adopting plausible values of external parameters α and $U_{\rm b}$, we are now able to solve these equations and investigate the structure of the gap.

3.1 Electric Field Distribution

To grasp the rough feature, we first show some examples of the solutions of $E_{\parallel}(x)$ in figure 2. For a fixed value of $U_{\rm b} = 10^6\,{\rm ergs/cm^3}$, we depicted the three cases of $\alpha = 1.5$, 2.0, and 2.5 denoted by the dashed, solid, and dotted lines, respectively. This figure indicates that $E_{\parallel}$ is maximum at the symmetry point, $x = 0$, decreases monotonically with increasing x, and finally vanishes at the boundary, $x = H$. The corresponding Lorentz factor, which is solved from equation (6), becomes several hundred.

As figure 2 indicates, the typical half width of the potential gap is $H \sim 0.005r_{\rm H}$ for $U_{\rm b} = 10^6\,{\rm ergs/cm^3}$. However, H has a strong dependence on $U_{\rm b}$. This is because the pair production mean free path $l_{\rm p}$, which essentially

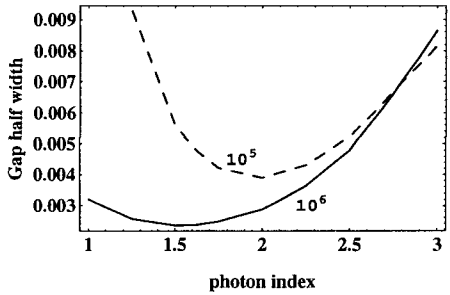


Figure 3: Gap half width H vs. photon index α. The ordinate is normalized by the hole's radius, $r_{\rm H} = 3 \times 10^{13}$cm. The solid line describes $H(\alpha)$ for $U_{\rm b} = 10^6\,{\rm ergs/cm^3}$, while the dashed one for $U_{\rm b} = 10^5\,{\rm ergs/cm^3}$.

describes H together with the number of γ-rays (N_γ) produced by a single e^+ or e^- via inverse Compton scatterings, strongly depends on $U_{\rm b}$. That is, the chance of a γ-ray photon to collide with background soft photons is proportional to $U_{\rm b}$; as a result, $l_{\rm p}$ increases with decreasing $U_{\rm b}$.

The results of H vs. α and $U_{\rm b}$ are summarized in figure 3. The solid line indicates H vs. α for $U_{\rm b} = 10^6{\rm ergs/cm^3}$, while the dashed one for $U_{\rm b} = 10^5{\rm ergs/cm^3}$. This figure indicates that H decreases with α for small α (i.e., hard spectrum), but it increases for large α (i.e., soft spectrum). The key point in understanding this behavior is that H be described by $l_{\rm p}/N_\gamma$, pair production mean free path divided by the number of γ-rays produced by a single e^+ or e^-.

Let us evaluate $l_{\rm p}$ and N_γ at $x = H$. As $f_i^+(x)$ and $\eta_{,i}$ have already been solved, we can compute $l_{\rm p}$ from

$$\frac{1}{l_{\rm p}} = \frac{\sum_{i=1}^{13} \eta_{{\rm p},i} f_i^+(H)}{\sum_{i=1}^{13} f_i^+(H)}. \tag{28}$$

For $U_{\rm b} = 10^6\,{\rm ergs/cm^3}$, $\log_{10} l_{\rm p}/r_{\rm H}$ vs. α at $x = 0$ is depicted by the dash-dotted line in fig. 4. As expected, $l_{\rm p}$ increases monotonically with α, because the number density of target X-ray photons decrease when the spectrum becomes soft.

Having observed the variations of $l_{\rm p}$ and N_γ, we can then go on to consider $l_{\rm p}/N_\gamma$. Figure 4 tells us that $\log_{10}[(l_{\rm p}/r_{\rm H})/N_\gamma]$, which is plotted by the dash-dotted line, becomes minimum for $\alpha \sim 2$. Moreover, $l_{\rm p}/N_\gamma$ describes very well

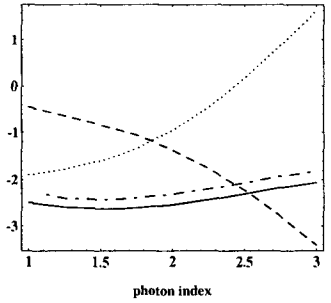


Figure 4:
The curve of $l_{\rm p}(\alpha)/r_{\rm H}$ at $x = 0$ for $U_{\rm b} = 10^6\,{\rm ergs/cm^3}$. The solid line represents $\log_{10}(H/r_{\rm H})$, while the dashed, dotted, and dash-dotted lines represent $\log_{10} N_\gamma^{-1}$, $\log_{10} l_{\rm p}/r_{\rm H}$, and $\log_{10}(l_{\rm p}/r_{\rm H})/N_\gamma$, respectively.

the width of the gap H, of which logarithmic values are plotted by the solid line.

We have seen in this subsection how H depends smoothly on $U_{\rm b}$ and α. Once that is understood, we are in a better position to evaluate the maximum allowed current density $j_{\rm cr}$. Remining that $j_0 = j_{\rm cr}$ gives $dE_\parallel/dx = 0$ at the boundary, we obtain

$$j_{\rm cr} = AH. \tag{29}$$

As a result, the pair production cascade provides sufficient current density for the effective energy and angular momentum extraction from a rotating suppermassive black hole. The explicit values of $j_{\rm cr}$ are depicted in figure 5 as a function of α and $U_{\rm b}$.

4 Summary and Discussion

To sum up, we have shown that a 'null surface', where B_z vanishes, can be formed close to a rotating black hole, by the space-time dragging effect. By solving the Vlasov equations, we explicitly demonstrated that a stationary pair-production cascade is maintained close to the null surface. For a typical X-ray field strength in an AGN, the strength of the acceleration fiels becomes a few thousand V $\rm m^{-1}$, which limits the particles' Lorentz factors below $10^{2.5\sim4}$. An accelerated e^+ or e^- scatters background UV and X-ray photons to produce $10^{1\sim3}$ γ-ray photons of which energies exceed the threshold of pair production. The cascade supplies sufficient amount of pair plasma necessary in maintaining the electric current in the magnetosphere and thus ensures efficient extraction

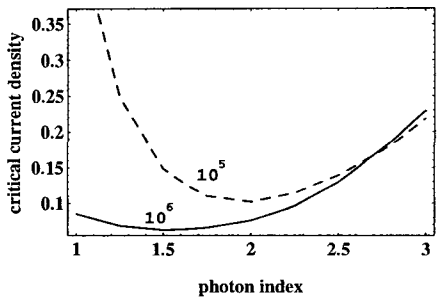


Figure 5: Critical current density vs. α. The solid line indicates $j_{\rm cr}(\alpha)$ for $U_{\rm b} = 10^6\,{\rm ergs/cm^3}$, while the dashed line for $U_{\rm b} = 10^5\,{\rm ergs/cm^3}$. The ordinate is normalized by $10^{-13}{\rm abampere/cm^2}$

of the hole's rotational energy in the form of electromagnetic energy. The width of the potential gap ($2H$) in which the cascade proceeds, is much less than the hole's radius, $r_{\rm H}$. Therefore, the expansion of $\rho_{\rm GJ}$ at $x = 0$ in equation (5) is self-consistently justified.

It will be necessary to make clear the differences between this work (see also Hirotani & Okamoto 1998) and the previous work of Beskin et al. (1992). They implicitly assumed that all of the γ-ray photons created by the inverse Compton scatterings contribute to the cascade, and explicitly put $l_{\rm p} \sim l_{\rm c} \sim H$. However, as seen in §3, $l_{\rm p}$ is much larger than $l_{\rm c}$ and H. This means that very small fractions of γ-ray photons are enough to maintain the cascade, and most of them escape freely from the gap. This situation is similar to the pair production cascade which is expected to occur in pulsar magnetospheres (e.g., Daugherty & Harding 1982), but might be dissimilar in that the γ-ray photons from the gap will be unobservable, because as should be so, the γ-ray flux going out of the gap is negligibly small compared with the outgoing Poynting flux from the hole. Owing to such a difference, figure 3 differs qualitatively from figure 6 in Beskin et al. (1992).

In addition to the above difference, we explicitly solved the γ-ray distribution functions by considering their spectrum, whereas in Beskin et al. (1992) the treatment is not explicitly mentioned. It is in fact important to take the γ-ray spectrum into account, because $l_{\rm p}$, which essentially controls the gap width.

Let us briefly discuss the consistency that the existence of the plasma source gap should not significantly affect the global magnetospheric structure. For one thing, as we have seen, the gap width, H, is much smaller than the global length scales, such as $r_{\rm H}$. What is more, as figure 6 indicates, the electric

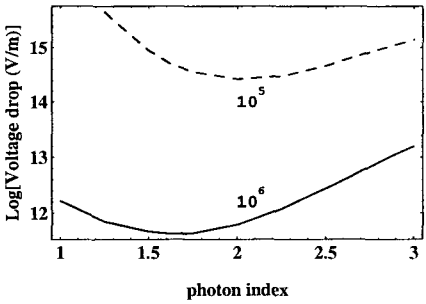


Figure 6: Total voltage drop in the gap as a function of α. The solid line corresponds to the case of $U_{\rm b} = 10^6\,{\rm ergs/cm^3}$, while the dashed line to that of $U_{\rm b} = 10^5\,{\rm ergs/cm^3}$.

potential drop in the gap is sufficiently small compared with the total EMF, $\sim 10^{19}\,{\rm volts}\,(a/M)(M/10^8 M_\odot)(B/10^4{\rm G})$. Even for the less luminous case of $U_{\rm b} = 10^5\,{\rm ergs/cm^3}$, in which current density j_0 allows almost most effective energy extraction from a rotating supermassive black hole, the potential drop in the gap is as small as 0.1% of the total EMF.

It might finally be interesting to investigate whether or not the same mechanism works around a stellar mass black hole, from which the rotational energy may be extracted, resulting in a galactic γ-ray jet (Levinson & Blandford 1996). Around a supermassive black hole, which is the concern of this paper, the relaxation length $l_{\rm drag}$ at which $e^\pm$'s lose their perpendicular momenta owing to scatterings with background UV photons, is much smaller than $l_{\rm p}$, $l_{\rm c}$ and H. Therefore, particle motion and γ-ray behavior can be treated one-dimensionally; in addition, electron/positron spectra are approximated by mono-energetic ones. However, as Beskin et al. (1992) pointed out, around a stellar mass black hole, neither the drag due to background photons nor the synchrotron loss cannot make the motion of $e^\pm$'s be one-dimensional, because their mean free paths are much larger than $l_{\rm c}$ or H. This issue will be our next target.

Acknowledgments

One of the authors (K.H.) wishes to thank Drs. S. Shibata and M. Takahasi for valuable comments and criticisms.

References

1. Beskin, V. S., Istomin, Ya. N., & Par'ev, V. I. 1992, Sov. Astron. 36(6), 642
2. Berestetskii, V. B., Lifshitz, E. M. & Pitaevskii, L. P., 1989, Quantum Electrodynamics 3rd, Japanese translation, Marruzen, Tokyo
3. Blandford, R. D. 1976, MNRAS, 176, 465
4. Blandford, R. D., 1989, in Theory of Accretion Disks, ed. P. Meyer, W. Duschl, J. Frank & E. Meyer-Hofmeister, Dordrecht: Klewer, 35 Friedlander, N. Gehrel, & D. J. Macomb, New York: AIP, 553
5. Blandford, R. D., 1993, in Proc. Compton Symp. 1992, ed. M.
6. Blandford, R. D., 1994, in Cosmical Magnetism, ed. D. Lynden-Bell, Dordrecht: Kluwer, 171
7. Blandford, R. D., & Levinson, A. 1995, ApJ, 441, 79
8. Blandford, R. D., & Znajek, R. L. 1976, MNRAS, 179, 433
9. Daugherty, J. K., & Harding, A. K. 1982, ApJ, 252, 337
10. Hirotani, K., Okamoto, I., 1998, ApJ 497, 563
11. Hirotani, K., Shibata, S., 2001, MNRAS 325, 1228
12. Hirotani, K., Shibata, S., 2001, ApJ 558, 216
13. Hirotani, K., Shibata, S., 2002, ApJ 564, 369
14. Levinson, A. & Blandford, R. D., 1996, ApJL, 456, L29.
15. Lovelace, R. V. E., 1976, Nature, 262, 649
16. Macdonald, D. A. & Thorne, K. S. 1982, MNRAS, 198, 345
17. Montigny, C. von, 1995, ApJ, 440, 525
18. Okamoto, I. 1992, MNRAS, 254, 192
19. Rybicki, G. B. & Lightman, A. P., 1979, Radiative processes in astrophysics, John Wiley & Sons, New York
20. Thorne, K. S., Price, R. H., & Macdonald, D. A. eds., 1986, Black Holes: The Membrane Paradigm, Yale University Press, New Haven & London

THE TRANSFIELD EQUATION OF THE AXISYMMETRIC, NONSTATIONARY MAGNETOSPHERE OF A BLACK HOLE

SEOK JAE PARK
Korea Astronomy Observatory
61-1 Hwaam-Dong, Yuseong-Gu, Daejeon 305-348, Korea
E-mail: sjpark@kao.re.kr

An axisymmetric, stationary, general-relativistic, electrodynamic engine model of an active galactic nucleus was formulated by Macdonald and Thorne, and consisted of a supermassive black hole surrounded by a plasma magnetosphere and a magnetized accretion disk. Based on this initial formulation, a nonstationary version of their model is constructed in this paper, and the associated 'Transfield Equation' is derived for a magnetosphere in the force-free limit. The result is consistent with previous work, and, as expected, is more general in many aspects.

1 Introduction

It is widely believed that the central engines in active galactic nuclei are related to supermassive black holes. An axisymmetric, stationary, general-relativistic, electrodynamic engine model was formulated by Macdonald & Thorne (1982[1], hereafter MT), and consisted of a supermassive black hole surrounded by a plasma magnetosphere and a magnetized accretion disk. MT assumed the magnetic field line velocity to be purely toroidal, as this is actually demanded by the axisymmetry and stationarity conditions.

Based on the MT model, we investigated the axisymmetric, *nonstationary* case for the central engine (Park & Vishniac 1989a[2], hereafter PV; Park & Vishniac 1989b[3]). The axisymmetric and nonstationary conditions are represented by (PV, eq. [3.1])

$$\mathbf{m} \cdot \nabla f = 0, \quad \mathcal{L}_{\mathbf{m}} \mathbf{f} = \mathbf{0}, \tag{1}$$

and

$$\frac{\partial f}{\partial t} \equiv \dot{f} \neq 0, \quad \frac{\partial \mathbf{f}}{\partial t} \equiv \dot{\mathbf{f}} \neq \mathbf{0}, \tag{2}$$

respectively. Here, $\mathbf{m}$ is a Killing vector of the axisymmetry, $\mathcal{L}$ means the Lie derivative, and f and $\mathbf{f}$ are any scalar and vector, respectively.

In PV, we derived all the fundamental equations in a fully time-dependent manner, and investigated the electrodynamics of the black hole. The main point in PV was to add the secular effects of mass accretion to the original axisymmetric, stationary model. Thus, in PV, the magnetic field line velocity

$\mathbf{v}$ has not only a toroidal component $\mathbf{v}^{\mathrm{T}}$, but also a poloidal component $\mathbf{v}^{\mathrm{P}}$. This poloidal velocity component $\mathbf{v}^{\mathrm{P}}$ was, however, assumed to be confined to the radial direction of a cylindrical polar coordinate system.

In this paper, we analyze a force-free black hole magnetosphere, and we do *not* assume $\mathbf{v}^{\mathrm{P}}$ to be confined to the radial direction in cylindrical polar coordinates. In doing this, we are then able to obtain the axisymmetric, *nonstationary* 'Transfield Equation', which, to our knowledge, has not yet been published elsewhere. In the stationary limit, this result is consistent with previous work; and it is more general in many aspects, as expected.

Throughout this paper, we define our units such that $c = G \equiv 1$, and we assume the central black hole to possess a Kerr metric with total mass M, angular momentum J, and angular momentum per unit mass a ($\equiv J/M$).

2 Reformulated Force-Free Electrodynamics

First, we summarize the basic equations of the axisymmetric, *nonstationary*, general-relativistic electrodynamics as already set out in PV. We use spherical polar coordinates (r, θ, φ) to describe the spacetime of the black hole. Throughout this paper, $\mathbf{m}$ has the same magnitude as ϖ, which is the separation between the black hole's symmetry axis and a given Fiducial Observer (hereafter FIDO; Thorne et al. 1986[4]),

$$\varpi \equiv \frac{\Sigma}{\rho} \sin\theta, \tag{3}$$

where

$$\rho^2 \equiv r^2 + a^2 \cos^2\theta, \tag{4}$$

$$\Sigma^2 \equiv \left(r^2 + a^2\right)^2 - a^2 \Delta \sin^2\theta, \tag{5}$$

and

$$\Delta \equiv r^2 + a^2 - 2Mr. \tag{6}$$

Let ∂A be an $\mathbf{m}$-loop, A be any surface bounded by ∂A, but not intersecting the black hole's event horizon, and let $d\mathbf{s}$ be the normal vector of an infinitesimal area on A. Then, we may define the total electric current passing downward through A as $I(t, \mathbf{r})$; the total magnetic flux passing upward through A as $\Psi(t, \mathbf{r})$; and the total electric flux passing upward through A as $\Phi(t, \mathbf{r})$; these definitions may be written as (PV, eq. [3.2])

$$I(t, \mathbf{r}) \equiv -\int_A \alpha \mathbf{j} \cdot d\mathbf{s}, \tag{7}$$

$$\Psi(t,\mathbf{r}) \equiv \int_A \mathbf{B}\cdot d\mathbf{s}, \tag{8}$$

and

$$\Phi(t,\mathbf{r}) \equiv \int_A \mathbf{E}\cdot d\mathbf{s}, \tag{9}$$

where α is the lapse function of the FIDO,

$$\alpha = \frac{\rho}{\Sigma}\sqrt{\Delta}. \tag{10}$$

In terms of these quantities, the electromagnetic fields, as observed by the FIDO, are given by (cf. PV, eqs. [3.3], [3.5] and [3.6])

$$\mathbf{E}^{\mathrm{T}} = -\frac{2}{\alpha\varpi^2}\left(\frac{\dot{\Psi}}{4\pi}\right)\mathbf{m}, \tag{11}$$

$$\mathbf{B}^{\mathrm{T}} = -\frac{2}{\alpha\varpi^2}\left(I - \frac{\dot{\Phi}}{4\pi}\right)\mathbf{m}, \tag{12}$$

$$\mathbf{E}^{\mathrm{P}} = \frac{\nabla\Phi\times\mathbf{m}}{2\pi\varpi^2}, \tag{13}$$

$$\mathbf{B}^{\mathrm{P}} = \frac{\nabla\Psi\times\mathbf{m}}{2\pi\varpi^2}, \tag{14}$$

and

$$\alpha\mathbf{j}^{\mathrm{P}} = -\frac{\nabla I\times\mathbf{m}}{2\pi\varpi^2}, \tag{15}$$

where T, P denote the toroidal and poloidal components, respectively.

If we fix the observer as the FIDO at a given point near a Kerr black hole, the Maxwell's equations which satisfy the axisymmetric, *nonstationary* conditions (1) and (2) are given by (PV, eq. [2.9])

$$\nabla\cdot\mathbf{E} = 4\pi\varrho_{\mathrm{e}}, \tag{16}$$

$$\nabla\cdot\mathbf{B} = 0, \tag{17}$$

$$\nabla\times(\alpha\mathbf{E}) = -\dot{\mathbf{B}} + (\mathbf{B}\cdot\nabla\omega)\mathbf{m}, \tag{18}$$

and

$$\nabla \times (\alpha \mathbf{B}) = \dot{\mathbf{E}} - (\mathbf{E} \cdot \nabla \omega)\mathbf{m} + 4\pi\alpha \mathbf{j}, \tag{19}$$

where ω is the FIDO's angular velocity

$$\omega = \frac{2aMr}{\Sigma^2}. \tag{20}$$

Following the arguments previously given in MT and PV, we employ both the 'highly-conducting' condition

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \mathbf{0}, \tag{21}$$

and the 'force-free' condition

$$\varrho_e \mathbf{E} + \mathbf{j} \times \mathbf{B} = \mathbf{0}. \tag{22}$$

In equation (21), $\mathbf{v}$ is, in the first instance, the plasma bulk flow velocity. Under the ideal magnetohydrodynamic conditions assumed here, this should be identical to the 'frozen-in' magnetic field line velocity, which we choose to measure relative to the FIDO. Using condition (21), we can set (PV, eq. [4.2])

$$\mathbf{v} \equiv \mathbf{v}^{\mathrm{T}} + \mathbf{v}^{\mathrm{P}} = -\frac{\omega - \Omega^{\mathrm{F}}}{\alpha}\mathbf{m} + \mathbf{v}^{\mathrm{P}}, \tag{23}$$

where Ω^{F} is the angular velocity of the magnetic field line around the black hole's rotation axis. Equation (23) means that the FIDO observes not only the toroidal component $\mathbf{v}^{\mathrm{T}}$ of $\mathbf{v}$, but also the poloidal component $\mathbf{v}^{\mathrm{P}}$, which arises from the accretion of plasma along with its 'frozen-in' magnetic flux. Note that we do *not* impose any constraints on $\mathbf{v}^{\mathrm{P}}$; this is an important difference between this paper and PV (in which $\mathbf{v}^{\mathrm{P}}$ was confined to move only parallel to the equatorial plane).

The electric field vector components are, therefore, given by

$$\mathbf{E}^{\mathrm{T}} = -\mathbf{v}^{\mathrm{P}} \times \mathbf{B}^{\mathrm{P}} = \frac{2}{\alpha\varpi^2}\left(\frac{\alpha \mathbf{v}^{\mathrm{P}} \cdot \nabla\Psi}{4\pi}\right)\mathbf{m}, \tag{24}$$

and

$$\mathbf{E}^{\mathrm{P}} = -\mathbf{v}^{\mathrm{T}} \times \mathbf{B}^{\mathrm{P}} - \mathbf{v}^{\mathrm{P}} \times \mathbf{B}^{\mathrm{T}} = \frac{\omega - \Omega^{\mathrm{F}}}{2\pi\alpha}\nabla\Psi + \frac{2}{\alpha\varpi^2}\left(I - \frac{\dot{\Phi}}{4\pi}\right)\mathbf{v}^{\mathrm{P}} \times \mathbf{m}. \tag{25}$$

Taking the vector cross products $\mathbf{v}^{\mathrm{P}}\times$ and $\mathbf{v}^{\mathrm{T}}\times$ of equation (25), we obtain the magnetic field vector components

$$\mathbf{B}^{\mathrm{T}} = -\frac{2}{\alpha\varpi^2(\mathbf{v}^{\mathrm{P}})^2}\left\{\left(\frac{\alpha\mathbf{v}^{\mathrm{P}}\cdot\nabla\Phi}{4\pi}\right)\mathbf{m} + \frac{\omega-\Omega^{\mathrm{F}}}{4\pi}\varpi^2\mathbf{v}^{\mathrm{P}}\times\nabla\Psi\right\}, \tag{26}$$

and

$$\mathbf{B}^{\mathrm{P}} = -\frac{\alpha}{2\pi(\omega-\Omega^{\mathrm{F}})\varpi^2}\nabla\Phi + \frac{2}{(\omega-\Omega^{\mathrm{F}})\varpi^2}\left(I-\frac{\dot{\Phi}}{4\pi}\right)\mathbf{v}^{\mathrm{P}}, \tag{27}$$

respectively.

Setting equation (24) for $\mathbf{E}^{\mathrm{T}}$ equal to equation (11), we obtain the relation

$$\frac{\dot{\Psi}}{4\pi} = -\frac{\alpha\mathbf{v}^{\mathrm{P}}\cdot\nabla\Psi}{4\pi}, \tag{28}$$

which reveals the fact that it is the poloidal velocity component $\mathbf{v}^{\mathrm{P}}$ which determines the nonstationarity; if $\mathbf{v}^{\mathrm{P}} = \mathbf{0}$, then $\dot{\Psi} = 0$ and we revert to the stationary model (as in MT). We may repeat the same procedure for the other 3 field components: equating (25) for $\mathbf{E}^{\mathrm{P}}$ to (13), or (26) for $\mathbf{B}^{\mathrm{T}}$ to (12), or (27) for $\mathbf{B}^{\mathrm{P}}$ to (14), we emerge with another useful equation (for the conduction plus displacement current through area A)

$$I - \frac{\dot{\Phi}}{4\pi} = \frac{\alpha}{4\pi(\mathbf{v}^{\mathrm{P}})^2}\mathbf{v}^{\mathrm{P}}\cdot\left(\nabla\Phi + \mathbf{v}^{\mathrm{T}}\times\nabla\Psi\right), \tag{29}$$

whose right-hand side should always remain finite, including in the stationary case (i.e. $\mathbf{v}^{\mathrm{P}} = \mathbf{0}$). For later use, we define a new, time-dependent, vector quantity

$$\dot{\mathbf{C}} \equiv \frac{2}{\alpha\varpi^2}\left(I-\frac{\dot{\Phi}}{4\pi}\right)\mathbf{v}^{\mathrm{P}} = \frac{1}{2\pi\varpi^2}\left(\nabla\Phi + \mathbf{v}^{\mathrm{T}}\times\nabla\Psi\right), \tag{30}$$

which, it should be noted, necessarily vanishes (i.e. $\dot{\mathbf{C}} = \mathbf{0}$) in the stationary case. In terms of $\dot{\mathbf{C}}$, we may write the poloidal field vector components in their general, time-dependent, forms as

$$\mathbf{E}^{\mathrm{P}} = \frac{\omega-\Omega^{\mathrm{F}}}{2\pi\alpha}\nabla\Psi + \dot{\mathbf{C}}\times\mathbf{m}, \tag{31}$$

and

$$\mathbf{B}^{\mathrm{P}} = -\frac{\alpha}{\omega - \Omega^{\mathrm{F}}}\left(\frac{1}{2\pi\varpi^2}\nabla\Phi - \dot{\mathbf{C}}\right). \tag{32}$$

3 The Transfield Equation

From equations (11), (16) and (31), we immediately get the charge density

$$4\pi\varrho_{\mathrm{e}} = \nabla\cdot\mathbf{E}^{\mathrm{P}} = \nabla\cdot\left(\frac{\omega - \Omega^{\mathrm{F}}}{2\pi\alpha}\nabla\Psi + \dot{\mathbf{C}}\times\mathbf{m}\right). \tag{33}$$

The current vector $\mathbf{j}$ is given by equation (19) as

$$4\pi\mathbf{j} = \frac{1}{\alpha}\nabla\times(\alpha\mathbf{B}) + \frac{1}{\alpha}(\mathbf{E}\cdot\nabla\omega)\mathbf{m} - \frac{1}{\alpha}\dot{\mathbf{E}}. \tag{34}$$

Substituting equations (14) and (31) into equation (34), we get

$$\begin{aligned} 4\pi j^{\mathrm{T}} = & -\frac{\varpi}{2\pi\alpha}\nabla\cdot\left(\frac{\alpha}{\varpi^2}\nabla\Psi\right) + \frac{\omega - \Omega^{\mathrm{F}}}{2\pi\alpha^2}\varpi\nabla\omega\cdot\nabla\Psi \\ & + \frac{\varpi}{\alpha}\dot{\mathbf{C}}\times\mathbf{m}\cdot\nabla\omega - \frac{1}{\alpha\varpi}\dot{\mathbf{E}}\cdot\mathbf{m}. \end{aligned} \tag{35}$$

Substituting equation (21) into equation (22) enables us to cast the current vector as

$$\mathbf{j} = \varrho_{\mathrm{e}}\mathbf{v} + \xi\mathbf{B}, \tag{36}$$

where ξ is a scalar function. The current vector components may then be written as

$$\mathbf{j}^{\mathrm{P}} = \varrho_{\mathrm{e}}\mathbf{v}^{\mathrm{P}} + \xi\mathbf{B}^{\mathrm{P}}, \tag{37}$$

and

$$\mathbf{j}^{\mathrm{T}} = \varrho_{\mathrm{e}}\mathbf{v}^{\mathrm{T}} + \xi\mathbf{B}^{\mathrm{T}}. \tag{38}$$

If we substitute equations (23), (33), and (35) into (38), we finally arrive at the *nonstationary* 'Transfield Equation' (otherwise known as the 'Stream Equation'), which, in our current formulation, may be written as:

$$\nabla \cdot \left[\frac{\alpha}{\varpi^2} \left\{ 1 - \frac{(\omega - \Omega^{\mathrm{F}})^2}{\alpha^2} \varpi^2 \right\} \nabla \Psi \right] - \left(\frac{\omega - \Omega^{\mathrm{F}}}{\alpha} \nabla \Psi + 2\pi \dot{\mathbf{C}} \times \mathbf{m} \right) \cdot \nabla \Omega^{\mathrm{F}}$$

$$-2\pi \nabla \cdot \{ (\omega - \Omega^{\mathrm{F}}) \dot{\mathbf{C}} \times \mathbf{m} \} + \frac{2\pi}{\varpi^2} \dot{\mathbf{E}} \cdot \mathbf{m} - \frac{16\pi^2 \xi}{\varpi^2} \left(I - \frac{\dot{\Phi}}{4\pi} \right) = 0, \tag{39}$$

This more general form of the Transfield Equation was not derived in PV, and, to our knowlege, has never before been published elsewhere. It is unfortunate, however, that equation (39) seems to be too complicated to solve directly via analytic methods; if this is true, then it would best be attacked via numerical computation.

We are now in a position to verify that our equations reduce to those of the stationary case, when all the nonstationary assumptions are removed. It is readily seen that all the nonstationary equations revert to their stationary counterparts, as in MT, as soon as all the time-derivative terms are set to zero. Equation (37) then becomes

$$\mathbf{j}^{\mathrm{P}} = \xi \mathbf{B}^{\mathrm{P}}, \tag{40}$$

in the stationary case, and, from equations (14), (15), and (40), we find

$$\xi = -\frac{1}{\alpha} \frac{dI}{d\Psi}, \tag{41}$$

and thus equation (36) becomes (MT, eq. [6.3])

$$\mathbf{j} = \varrho_{\mathrm{e}} \mathbf{v} - \frac{1}{\alpha} \frac{dI}{d\Psi} \mathbf{B}. \tag{42}$$

If we set all the time-derivative terms in equation (39) to zero, and substitute equation (41), we finally obtain

$$\nabla \cdot \left[\frac{\alpha}{\varpi^2} \left\{ 1 - \frac{(\omega - \Omega^{\mathrm{F}})^2}{\alpha^2} \varpi^2 \right\} \nabla \Psi \right] - \frac{\omega - \Omega^{\mathrm{F}}}{\alpha} \nabla \Omega^{\mathrm{F}} \cdot \nabla \Psi + \frac{16\pi^2}{\alpha \varpi^2} I \frac{dI}{d\Psi} = 0, \tag{43}$$

which is the well-known, *stationary* 'Stream Equation' (MT, eq. [6.4]). This equation (43) was originally derived by Blandford & Znajek (1977)[5], and numerically solved first by Macdonald (1984)[6].

References

1. Macdonald, D.A., & Thorne, K.S., *MNRAS* **198**, 345 (1982) (MT)
2. Park, S.J., & Vishniac, E.T., *Ap. J.* **337**, 78 (1989a) (PV)
3. Park, S.J., & Vishniac, E.T., *Ap. J.* **347**, 684 (1989b)
4. Thorne, K.S., Price, R.H., & Macdonald, D.A., eds., "Black Holes: The Membrane Paradigm" (Yale Univ. Press, New Haven CT, USA, 1986).
5. Blandford, R.D., & Znajek, R.L., *MNRAS* **179**, 433 (1977)
6. Macdonald, D.A., *MNRAS* **211**, 313 (1984)

GAMMA-RAY BURSTS AND GRAVITATIONAL RADIATION FROM BLACK HOLE-TORUS SYSTEMS

M.H.P.M. van PUTTEN
MIT 2-378, Cambridge, MA 02139, USA
E-mail: mvp@schauder.mit.edu

Cosmological gamma-ray bursts (GRBs) are probably powered by systems harboring a rotating black hole. This may result from hypernovae or black hole-neutron star coalescence. We identify short/long bursts with hyper- and suspended-accretion states around slowly/rapidly rotating black holes [van Putten & Ostriker, ApJ, 552, L31, 2001]. Black holes may be activated into producing outflows by a surrounding torus magnetosphere, in the form of baryon poor jets as input to the observed GRB/afterglow emissions. Here, we attribute these outflows to a differentially rotating gap in an open flux-tube along the axis of rotation of the black hole [van Putten, Phys. Rep., 345, 1 (2001)]. A high incidence of the black hole luminosity into the surrounding matter is expected by equivalence in poloidal topology to pulsar magnetospheres. For long bursts, this predicts a large fraction of the black hole spin-energy emitted in gravitational waves by a quadrupole moment in the surrounding torus [van Putten & Levinson, ApJ, 555, L41, 2001]. This suggests that long GRBs may be the most powerful LIGO/VIRGO burst sources of gravitational waves in the Universe with an expected duration of 10-15s on a horizontal branch of 1-2kHz in the $\dot{f}(f)$−diagram [van Putten, Phys. Rev. Lett., 87, 091101, 2001]. Gravitational wave-emissions from GRBs, therefore, promise calorimetric evidence for Kerr black holes.

1 Introduction

Cosmological gamma-ray bursts (GRBs) are the most enigmatic events in the Universe (Fig. 1). Their emissions are characteristically nonthermal in the few hundred keV range with a bi-modal distribution in durations, of short bursts around 0.3s and long bursts around 30s.[13] Redshift determinations from long bursts indicate a cosmological origin, probably associated with the formation of young massive stars.[22,4] GRBs, therefore, are probably most frequent within a redshift $z = 1 - 2$. The recent proposal[5] that GRB relics could be found in some of the galactic soft X-ray transients (SXTs), notably so GRO J1655-40 [12] and V4641 Sgr,[18] further suggests an association with black holes of about $3 - 14M_{\odot}$.

The inner engine producing the GRBs should be energetic and compact. Angular momentum forms a canonical energy reservoir, and GRB inner engines are probably no exception. This is consistent with breaking of spherical symmetry.[36] Hypernovae, then, follow up on this idea by postulating massive young stars in binaries as their progenitors. High angular momentum is also present in coalescing black hole-neutron star systems.

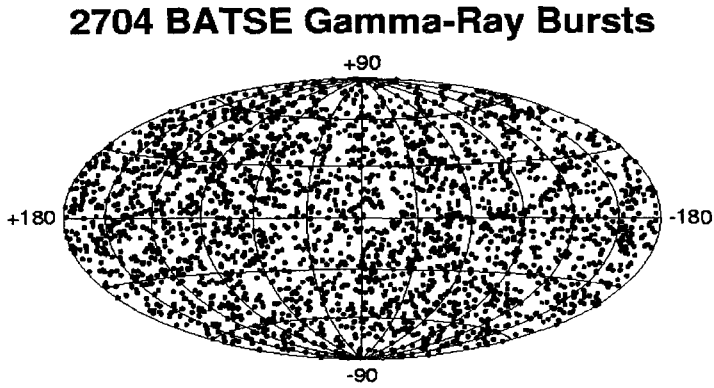


Figure 1: Shown are the locations of 2704 GRBs from the GRBs from the BATSE Catalogue over a nine-year period. The projection is in galactic coordinates. (Courtesy of NASA Marshall Space Flight Center, Space Sciences Laboratory.) The isotropic distribution and a $< V/V_{max} >$ distinctly less than the Euclidean value of $1/2$ (Schmidt 1999) are indicative for their cosmological origin. For long bursts, this is further evidenced by redshifts of order unity from GRB afterglow emissions and a probably association with star-forming regions (Paczynski 1998; Bloom et al. 2000). Hence, long bursts are expected to track closely the star-formation rate, which peaks at about $z = 1 - 2$. Remnants of GRBs are potentially found in some of the Soft X-ray transients, with notable candidates GRO J1655-40 (Israelian et al, 1999) and V4641Sgr (Orosz et al., 2001)

Here, I discuss some physical aspects of black hole-torus systems and a few observational predictions. Black holes may be active in accord with the Rayleigh criterion, in contact with a torus magnetosphere supported by baryonic matter. This is expected to be manifest in non-thermal emissions, in outflows along an open flux-tube on the axis of rotation of the black hole and various emissions from a surrounding torus:

- At supercritical magnetic field-strengths, outflows may spontaneously be created by frame dragging-induced potential differences, as follows from a perturbative calculation in Hawking's approach.[29] In competition with equilibration by charge separation, pair-creation is expected to be confined to a gap along the open flux-tube. Asymptotic boundary conditions on the horizon and infinity, linked by current continuity, define the net dissipation in the gap, and hence a powerful outflow whenever the black hole spins more rapidly than twice the angular velocity of the torus.[31]

- The torus is expected to be luminous in several channels, notably so in gravitational radiation, neutrino emissions and Poynting flux winds. Powered by the spin-energy of the black hole [30], the fluence in gravitational waves may reach about 1% of the black hole mass. The coupling to the rotational energy of the black hole operates by equivalence in poloidal topology to pulsar magnetospheres.[28,6] Detection by LIGO/VIRGO of these emissions may result in calorimetric evidence of Kerr black holes [33], whenever the fluence is determined to be in excess of the rotational energy of a rapidly rotating neutron star.

These prospects suggest several observational predictions.

Increasing evidence towards clustering in the true energy in GRB emissions [9,19,24] indicates a standard opening angle of about $\theta_H \simeq 35^o$ of the open magnetic flux-tube on the horizon of the black hole [32]. This predicts a cut-off $\theta_j \leq 35^o$ in the observed opening angles θ_j on the celestial sphere in any large statistical sample. Furthermore, with outflows created in long and short bursts alike, HETE-II may detect afterglows also to short GRBs [30] (see also [16]). The latter is expected to differ mostly in net fluence, as their inner engine operates for a significantly shorter time than in long GRBs.

Gravitational radiation appears to be a major channel in the emissions from the torus in long bursts, along with emissions in neutrinos and winds coming off the torus. Calculations in the suspended-accretion state indicate a net luminosity in gravitational waves of about one-third the net luminosity of the black hole.[33] This amounts to a fluence E_{gw} about 1% of the mass-energy of the central black hole. The frequency f_{gw} in these emissions at twice the Keplerian

frequency of the torus, as it develops a quadrupole moment in its mass distribution, is expected to be about 1-2kHz for a black hole mass of about $10M_{\odot}$. This range overlaps with the design bandwidth of 0.1-1.5kHz of the upcoming Laser Interferometric Gravitational Wave Observatories LIGO/VIRGO.

This raises an unanticipated prospect: calorimetric evidence for Kerr black holes from the emission in gravitational waves from the torus.[34] Indeed, consider the product $\alpha = 2\pi E_{gw} f_{gw}$, which expresses a measure for the ratio of rotational energy to the linear size of the inner engine. It appears that α from black hole-torus systems may reach values in excess of those attainable by rapidly rotating neutron stars. A LIGO/VIRGO detection of a large α, therefore, individually or as an average over a sample of detections, would be evidence for the Kerr relationship $E_{rot} \sim M/3$ between the rotational energy E_{rot} and the mass M of a rapidly rotating black hole.

The proposed association of gamma-ray bursts to black hole-torus systems will be reviewed in §2, and prospects for GRBs as potential LIGO/VIRGO sources is outlines in §3. We close with a comment on the potential for calorimetric evidence of Kerr black holes in GRBs.

2 A theory of gamma-ray bursts from black hole-torus systems

A black hole-torus system is of compact dimension, consistent with the short time-variability in the GRB light-curves and the proposed GRB-SXT association. The mass in the surrounding torus or disk will be limited, in both the hypernovae and binary black hole-neutron star coalescence scenario. This introduces relatively short time-scales of accretion, leaving a central Kerr black hole as the major energy reservoir. This poses two questions: what accounts for the duration in long GRBs and how can the rotational energy of the black hole create baryon poor jets?

2.1 Formation of black hole-disk or torus systems

A black hole-torus system may form from binary black hole-neutron star coalescence. Here, the neutron star gradually approaches the black hole by angular momentum loss in gravitational radiation. The neutron star will then be subject to tidal interactions, which may lead to break-up outside the inner most stable circular orbit (ISCO) when the central black hole is sufficiently small in mass. For this to happen, canonical estimates provide a bound of $3.7M_{\odot}$ on non-rotating black holes and a bound of $28M_{\odot}$ on rapidly rotating black holes. This indicates a substantially wider window of mass for the rotating case. It follows that a torus is more likely to form around a Kerr black hole.[21]

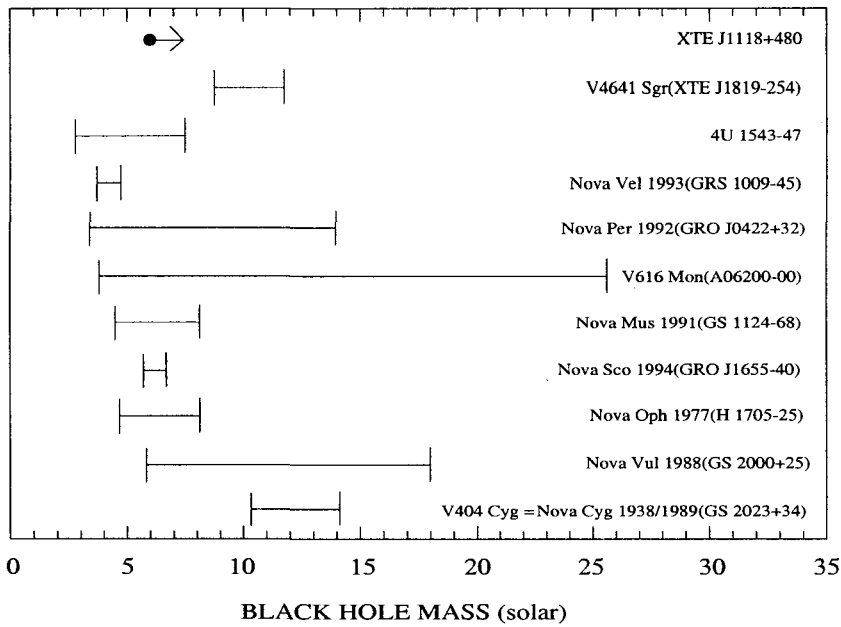


Figure 2: Shown is the distribution of black hole masses in X-ray novae. The top four are XTE J118+408 (McClintock et al., 2001), V4641 Sgr (Orosz et al., 2001), 4U 1543-47 (Orosz et al., 1998) and Nova Vel 1993 (Filippenko et al., 1999); the lower seven are from Bailyin et al. (1998). This mass distribution manisfests a certain diversity in black hole masses of about $3 - 14M_{\odot}$. [Reprinted from van Putten, *Physics Reports*, 345 ©2001 Elsevier B.V.]

The collapsar, failed supernova or hypernova scenario envisions the collapse of the center of a young massive star with high angular momentum. The origin of the angular momentum is most likely orbital angular momentum from the progenitor binary system. While the details of orbital angular momentum transfer into the collapsing star are somewhat uncertain, collapse of a rapidly rotating object is expected to result in a compact core surrounded by matter stalled against an angular momentum barrier. If the core forms a black hole in prompt collapse then, furthermore, the black hole will have a *minimum* mass, sufficient to account for the angular momentum J_H in view of the Kerr constraint $J_H^2 \leq M^2$ (in geometrical units, with M denoting the Schwarzschild radius Gm/c^2, where G is Newton's constant, m the mass of the black hole and c the velocity of light) [30]. For example, a Lane-Emden relationship with polytropic index $n = 3$ for the progenitor star gives $M \geq 10M_{\odot}$, consistent with the observed range of $3 - 14M_{\odot}$ in SXTs shown in Fig. 2.

In both scenarios, a magnetized neutron star or young massive star - represented by a magnetic moment density – will result in a disk or torus endowed with a net poloidal flux. It may be appreciated that the mass in the disk or torus a finite in a much more stricter sense than in analogous configurations believed to exist in active galactic nuclei. With magnetic regulated accretion, accretion of $0.1M_{\odot}$ becomes fairly rapid on a time-scale of a second or less onto onto a $10M_{\odot}$ mass black hole. Depleting the surroundings of baryonic matter inevitably prevents any further activity of the black hole. This suggests that additional physical processes should account for the relatively long duration in long bursts. In a recent proposal, the magnetic moment density of the surrounding torus is believed to permit a suspended-accretion state for the duration of spin-down of the central black hole. A bi-modal distribution of durations then occurs when the ratio of black hole-to-disk or torus mass is large.[30]

2.2 The lowest energy state of the black hole

The black hole will be surrounded by a torus magnetosphere, supported by the accretion disk. The black hole will adjust to a lowest energy state by developing an equilibrium magnetic moment[31]

$$\mu_H^e \simeq aBJ_H, \tag{1}$$

where $a = J_H/M$ denotes the specific angular momentum of a black hole with mass M and angular momentum J_H and B denotes the (average) poloidal magnetic field. This results from a minimum in the energy $\mathcal{E}(q) = (1/2)Cq^2 - \mu_H B$, where $C \simeq 1/r_H$ denotes the capacitance of the black hole, q the charge on the horizon, and Carter's identity[7] $\mu_H = qJ_H/M$ ("no fourth hair"). The minimum of $\mathcal{E}$ at $q \simeq BJ_H$ recovers Wald's result.[35] Similar results are found in a largely force-free magnetosphere.[14] This equilibrium magnetic moment preserves an essentially maximal and uniform horizon flux. This serves to preserve a strong coupling to the magnetosphere and, hence, to the inner face of the surrounding torus. It also permits the black hole to support an open flux-tube to infinity along its axis of rotation, particularly so in a suspended accretion state. Frame-dragging will act on this flux-tube, to to produce a differentially rotating gap for the creation of baryon poor outflows.

2.3 Hyper- and suspended-accretion states for a ring

GRBs show a bi-modal distribution in durations, as show in Fig. 3. We attribute this to different modes of angular momentum losses in black hole plus disk or torus systems.

As magnetic fields can be very efficient in mediating angular momentum transport, magnetic regulated accretion times tend to be short. Indeed, the accretion of a magnetized ring is illustrative, showing evolution towards a finite-time singularity[30]

$$\varpi = (1 - t/t_f)^p \tag{2}$$

of its radius $\varpi = R(t)/R(0)$, where $R(0)$ denotes the initial radius and t_f the final time of collapse. Here, $p = 1/2$ and 2 for a split monopole geometry (SMG) and toroidal field geometry (TFG), respectively. The final time in both field geometries is governed by the ratio of kinetic-to-magnetic energy $\delta E_k/\delta E_B$ - a free parameter, at present not well constrained from first principles or current numerical simulations. For a fiducial value of 10^2, SMG applied to the initial evolution ring and TFG applied to the final evolution of the ring indicates accretion times t_f of a few seconds or less – consistent with the time-scale of short bursts. This analysis does not take into account any action by the black hole back back onto the ring, i.e.: the results apply to the accretion onto slowly rotating or Schwarzschild black holes.

What then, may account for the duration of long bursts? A ring surrounding a rapidly rotating black hole will support, by a magnetic moment density, poloidal magnetic field-lines connected the ring to the horizon. These field-lines permit energy and angular momentum transport from the spin of the black hole into the ring. This process operates by equivalence in poloidal topology to pulsar magnetospheres, wherein the ring and the horizon of the black hole are, respectively, equivalent to a pulsar and infinity. The angular velocity of the equivalent pulsar is the relative angular velocity between the ring and the black hole. When the black hole spins more rapidly than the ring, the ring receives angular momentum *like a pulsar being spun-up when infinity wraps around it.*

The ring – an element of the disk or torus – will assume a suspended-accretion state arises when radiative losses in energy and angular momentum to infinity are replenished by gain from the central black hole. For a balance on magnetic torques alone, we find a critical angular velocity Ω_H^* of the black hole: [30]

$$\Omega_H^* = \Omega_T[1 + (f_w/f_H)^2 |\ln(\theta_0/2)|^{-1}] \tag{3}$$

for fractions f_w and f_H of magnetic flux which connect to infinity (in an outgoing Poynting flux-wind) and the black hole (in an ingoing Poynting flux-wind), respectively, where θ_0 denotes the minimum poloidal angle in TFG.

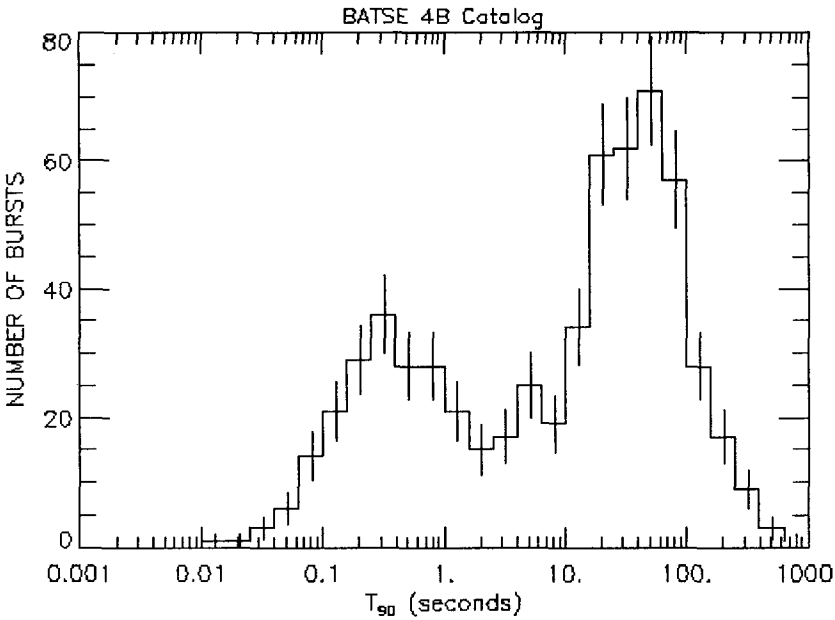


Figure 3: Shown is the bi-modal distribution of GRB durations of the 4B Catague, set by a T90 duration parameter, on lightcurves integrated over all 4 channels (E >20keV). (Courtesy of NASA Marshall Space Flight Center, Space Sciences Laboratory.) The population of long bursts is probably associated with young massive stars and, hence, with a redshift of $z = 1 - 2$. This indicates a redshift corrected mean value of the intrinsic duration of about 10-15s. In van Putten & Ostriker (2001), shorts bursts are identified with magnetic regulated hyperaccretion onto slowly rotating black holes, and long bursts with rapidly rotating black holes in a state of suspended accretion. Long bursts are potential LIGO/VIRGO sources of gravitational radiation by gravitational radiation from the torus, derived from the spin-energy of the black hole. This indicates a mean duration of 10-15s of gravitational wave-emissions commensurate with the redshift corrected GRB-event, for a cosmologically nearby sample within the detection sensitivity of LIGO/VIRGO.

This state lives as long as the black hole spins rapidly [30]

$$t_{long} = 88s \left(\frac{M}{10M_\odot}\right) \left(\frac{M/M_d}{100}\right) \left(\frac{E_k/E_B}{100}\right) g^2(\theta_0), \tag{4}$$

where $g(\theta_0)$ denotes a geometrical factor of order unity. We conclude that a bimodal distribution in duration occurs due to hyper- and suspended-accretion states whenever the ratio M/M_d is large.

2.4 Outflows from a differentially rotating gap

Frame-dragging introduces differential rotation along the axis of rotation of a Kerr black hole. Flux-tubes on a differentially rotating space-time background thus tend to develop potential differences by Faraday-induction. These potentials may drive large-scale currents. See [29] for a perturbative discussion about a Wald [35] field in Hawking's approach. At the same time, the magnetosphere tends to equilibrate locally by charge-separation. These two processes are generally in competition with one another, subject to current continuity linking asymptotic boundary conditions on the horizon and at infinity. In one proposal for the boundary conditions, discussed below, these considerations indicate that outflows may be created by a macroscopic gap in an open flux-tube along the axis of rotation of the black hole.[31,6]

An open flux-tube supported by the black hole is endowed with ingoing boundary conditions on the horizon of the black hole and outgoing boundary conditions at infinity. In field theory, these asymptotic boundary conditions assume conjugate radiative-radiative boundary conditions. In the continuum limit of a plasma which is asymptotically in charge-separated equilibrium, these become slip-slip boundary conditions: the angular velocities of the flux-surfaces on the horizon may differ from that of the black hole, and the angular velocity at infinity may be non-zero. In contrast, a flux-surfaces supported by baryonic matter are fixed to its angular velocity, namely that of the disk or torus (a no-slip boundary condition). This will hold to within a fair approximation over an appreciable scale relative to the system size. Recall that this well-known corotation law is based on the singular limit of perfect conductivity; deviations of order unity will arise over distance scales of order $1/\epsilon$, upon deviations from the corotation charge-density to order ϵ.

Equilibration towards a force-free state introduces an asymptotic null-condition on the current carried by the flow going into the black hole: $j^2 \to 0$ upon approaching the horizon, where j^b denotes the four-current.[25,31] This expresses the condition that the current becomes asymptotically convective: $j^r = \pm \alpha j^t$, where α denotes the redshift factor on-axis of the Kerr black

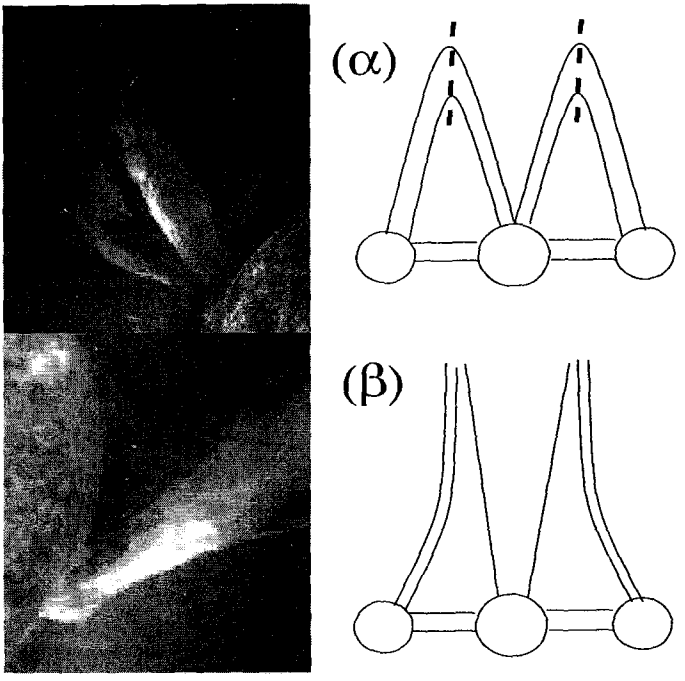


Figure 4: Images of closed (*top left*) and open (*bottom left*) topology of flux-tubes in the solar atmosphere from the Solar Heliospheric Observatory (Courtesy of the SOHO/EIT consortium, ESA-NASA). The footpoints of the tubes are rooted in the surface of the Sun with no-slip boundary conditions. The open tube represents a violent coronal mass ejection. The present proposal builds on potentially similar structures in a torus magnetosphere around a black hole, wherein the first corresponds to field-lines connecting the torus (with no-slip boundary conditions) to the black hole (with slip boundary conditions) as schemetically indicated in (α); the second corresponds to field-lines extending from the black hole to infinity (with slip boundary conditions on the horizon) and from the torus to infinity (with no-slip boundary conditions on its surface) as schemetically indicated in (β). Since open flux-tubes (β) form from closed loops (α), these fluxes are the same in magnitude and opposite in sign for the black hole and the torus. This co-axial flux-structure permits current closure at infinity.

hole. In approaching the horizon, drift-currents are suppressed by a divergent Lorentz factor. Here, we shall consider the proposal that the boundary condition at infinity is similar in an ultrarelativistic outflow. It would be of interest to study this proposal self-consistently with the micro-physics in the gap.

Frame-dragging appears explicitly in the expression for the electric charge-density ρ in the equilibrium charge-separated limit. Indeed, the equilibrium charge-density is associated with a time-like coordinate which is orthogonal to the azimuthal Killing vector. Hence, this charge-density corresponds to the density-at-infinity as seen by zero angular momentum observers (ZAMOs); the "true" angular velocity of a flux-surface is that relative to a local ZAMO with angular velocity $-\beta$. Consequently, we have the expression $\rho = -(\Omega+\beta)B/2\pi$ as a modified Goldreich-Julian density.[10,3,11,31] The asymptotical condition $j^2 = 0$ on the horizon and infinity now expresses electric current mediated by convection of this modified Goldreich-Julian density. Integrating over an effective area corresponding to a given flux surface A_ϕ =const., we find *current sources* $I_- = \Omega_- A_\phi$ at infinity and $I_+ = (\Omega_H - \Omega_+)A_\phi$ on the horizon. Here, Ω_- and Ω_+ denote the Boyer-Lindquist angular velocities of the two asymptotically equilibrated sections attached to infinity and the horizon, respectively. Current continuity enforces the condition $\Omega_- = \Omega_H - \Omega_+$.

Global current closure may obtain over the surrounding torus. Here, we appeal to a potential similarity to solar flares, as observed by the Transient Region Corona Experiment (TRACE) and the Solar Heliospheric Observatory (SOHO). While magnetic field-lines form closed loops when supported by compactly supported current sources, these loops are occasionally unstable, and flare as open prominences, as shown in the left column of Fig. 4. This suggests that open flux-tubes might form from loops in the torus magnetosphere connecting the black hole and the torus, sketched in the right column of Fig. 4. If so, this gives rise to an inner open flux-tube supported by the magnetic moment of the black hole and outer open flux-tube supported by the magnetic moment of the inner face of the torus. Creating open flux-tubes in the fashion is accompanied by an algebraic constraint: the inner and outer flux-tubes carry a magnetic flux which is equal in magnitude and opposite in sign. Applying the same asymptotic boundary condition $j^2 = 0$ to the outflow from the torus – and this is expected to be an approximation to within order unity – we obtain global current closure in the form of $I_- = I_+ = I_T = \Omega_T A_\phi$, where Ω_T denotes the angular velocity of the torus. The result is a differentially rotating gap between forementioned two equilibrium sections, one attached to infinity and the other attached to the horizon, with a Faraday-induced potential drop $\Delta V = (\Omega_+ - \Omega_-)A_\phi$. The power dissipated in this gap becomes

$$P = \Omega_T(\Omega_H - 2\Omega_T)A_\phi^2. \tag{5}$$

Thus, a gap forms with macroscopic dissipation whenever the black hole spins more rapidly than *twice* the angular velocity of the torus.[31]

Of some interest is the formation of the gap ($\Omega_H > 2\Omega_T$) while being in a state of hyper-accretion ($\Omega_H < \Omega_H^*$). Attributing the power released in the gap to the input to the observed GRB and afterglow emissions, suggests the possibility for afterglow emissions to short bursts. Unless the environment to short bursts is dramatically different from long bursts, HETE-II should see afterglows to short bursts as well.[30]

2.5 Clustering and spread in GRB emissions

Recent analysis of achromatic breaks in a sub-sample of GRB lightcurves indicates that these emissions are beamed, and that their true fluence E_{grb} is standard (with a dynamic range of about one decade) at few times 10^{50}ergs [9,19]. At the same time, the opening angle displays a rather wide dynamic range, between a few degrees and a few tens of degrees.

In [32], we consider a geometrically standard inner region in the vicinity of the black hole, when the torus is thick relative to the size of the black hole. In this event, the opening angle of the open flux-tube on the horizon commensurate with the true emissions in GRBs is about 35^o. Collimation of this flux-tube down to an opening angle θ_j on the celestial sphere may derive from winds coming off the torus, possibly so along forementioned outer flux-tube. This will account for a true output in outflow of about $E_{grb}(\theta_H/\theta_j)\epsilon^{-1}$, where $\epsilon \sim 0.15$ denotes the efficiency of kinetic energy to gamma-rays. Though substantial, this output remains well below the energy deposit into the torus by the black hole. This predicts a bound

$$\theta_j \leq 35^o \tag{6}$$

in any large sample of observations. We attribute variations in the opening angle θ_j to a diversity in torus parameters.

3 GRBs: the tip of the iceberg?

For long bursts, the equivalence in poloidal topology to pulsar magnetospheres indicates a high incidence of the black hole luminosity into the torus. Only a small fraction of less than one percent of the black hole output is associated with the true output in GRB-afterglow emissions.

3.1 Gravitational radiation from a torus around a black hole

The torus processes the input from the black hole by emission in various channels. This can be estimated in a suspended accretion state, including gravitational radiation, neutrino emissions and Poynting flux-winds. Gravitational radiation will be emitted as the torus develops non-axisymmetries, which features several aspects which suggest considering long GRBs as potential sources for LIGO/VIRGO. Namely, the torus is strongly coupled to the spin-energy of the black hole; lumpiness in the torus will produce gravitational radiation at twice the Keplerian angular frequency, i.e., in the range of 1-2kHz; the emission in gravitational radiation should dominate over emissions in radio waves; the true rate of GRBs should be frequent as inferred from their beaming factor of a few hundred. These gravitational wave emissions from the torus are powered by the spin energy of the black hole. This sets it apart from such emissions in neutron star-neutron star mergers or by fragmentation in collapse towards supernaovae.

Non-axisymmetries in the torus are expected from dynamical and, potentially, radiative instabilities. Notably so, a geometrically thick torus, consistent with the recent indication that long GRBs may be standard, is generally subject to the Papaloizou-Pringle instability[20]. If the torus reaches a mass on the order of that of the central black hole, it will be unstable to self-gravity. Of interest is further the possibility of a Chandrasekhar-Friedman-Schutz instability, or radiative instabilities since lumps of matter radiate preferentially on inner orbits. It would be of interest to study these radiative instabilities in further detail. The resulting gravitational wave-emissions may be quasi-periodic (QPO). This may be reminiscent of the observed QPOs in accretion disks in X-ray binaries, some of which have been attributed to general relativistic effects in orbital motions[27].

A detailed calculation in the suspended-accretion state gives the estimate $L_{gw} \simeq L_H/3$ for the luminosity in gravitational waves as a fraction of the black hole luminosity L_H.[33] This gives a fluence in gravitational waves

$$E_{gw} \simeq 1\% M \tag{7}$$

in terms of the mass M of a rapidly spinning Kerr black hole. This suggests to consider black hole-torus systems as potential LIGO/VIRGO sources of gravitational waves. The abundance of GRBs on a cosmological scale suggests to consider an interesting contribution to the stochastic background in gravitational radiation.[1,8]

3.2 Calorimetric evidence for Kerr black hole

The existence of black hole remains circumstantial, in particular for the population of stellar mass black holes. Obtaining evidence based on first principles is a challenging task which occasionally drives new observational strategies. Perhaps, then, LIGO/VIRGO may contribute by calorimetry on gravitational wave emissions. Successful detection of a burst in gravitational waves of a duration commensurate with the redshift corrected duration of 10-15s of long bursts would provide evidence for a compact and high-angular momentum inner engine. Additionally, a fluence in gravitational waves in excess of the rotational energy of a rapidly spinning neutron star would support the presence of a central Kerr black hole. Kerr black holes have the unique property of storing up to about a third of their mass in rotational energy. This has no baryonic counter part.

This may be pursued by the observable combination

$$\alpha = 2\pi E_{gw} f_{gw}. \tag{8}$$

Here, α is dimensionless upon use of geometrical units, with energy expressed in terms of the corresponding Schwarzschild radius and frequency in 1/cm. It may be noted that α is a compactness parameter reminiscent of the dimensionless Kerr parameter a/M. Indeed, for large specific angular momenta a, the ratio a/M becomes effectively a measure for the stored rotational energy to the linear size of the black hole. The estimates above show that a black hole-torus system may produce $\alpha = 0.01 - 0.015$, while a neutron star satisfies $\alpha < 0.007$. This indicates an opportunity to detect α in excess of that permitted by a neutron star, to serve as calorimetric evidence of a Kerr black hole. Notice that E_{gw} requires a distance estimate to the source. In practice, this may require statistical analysis on a sample of detections.

Acknowledgment. The author thanks the Korea Institute for Advanced Study for their hospitality and for hosting a very stimulating meeting. He also thanks G.E. Brown, C.W. Lee and A. Levinson for continuing conversations. This work is partially supported by NASA Grant No. 5-7012 and an MIT C.E. Reed Award.

References

1. Auriemma, G., 2001, J. Physics G., to appear; astro-ph/0108314
2. Bailyn, C.D., et al., 1998, ApJ, 499, 367
3. Beskin, V. S. 1997, Usp. Fiz. Nauk, 167, 689, Eng. Trans. in Physics - Uspekhi, 40, 659
4. Bloom, J.S., Kulkarni, S.R.,& Djorgovski, S.G., 2000, astro-ph/0010176

5. Brown, G.E., Lee, C.-H., Wijers R.A.M.J., Lee, H.K., Israelian G. & Bethe H.A., 2000, NewA, 5, 191
6. Brown, G.E., Bethe, H.A., & Lee, C.-H., 2001, Selected Papers: Formation and Evolution of Black Holes in the Galaxy (World Scientific), in preparation
7. Carter B., 1968, Phys. Rev., 174, 1559
8. Coward, D., et al., 2001, in preparation
9. Frail, D.A. et al., astro-ph/0102282
10. Goldreich, P., & Julian, W. 1969, ApJ, 157, 869
11. Hirotani, K., 1998, ApJ, 497, 563
12. Israelian, G., Rebolo, R., Basri, G., Casares, J., & Martín, E.L., 1999, Nature, 401, 142
13. Kouveliotou, C. *et al.*, 1993, ApJ, 413, L101
14. Lee, H.K., Lee, C.H., & van Putten, M.H.P.M., 2001, MNRAS, 324, 731
15. McClintock, J.E., et al., 2001, ApJ, 551, L147
16. Panaitescu, A., Kumar, P., & Narayan, R., et al., 2001, astro-ph/0108132
17. Orosz, et al., 1998, APJ, 499, 375
18. Orosz, J.A., Kuulkers, E., McClintock, J.E., Garcia, N.R., Jain, R.K., Bailyn, C.D., & Remillard, R.A., 2001, ApJ, 555, 489
19. Panaitescu, A., & Kumar, P., 2001, astro-ph/0109124
20. Papaloizou, J.C.B., & Pringle, J.E., 1995, MNRAS, 208, L12
21. Paczyński, B.P. 1991, Acta Astron, 41m, 257
22. Paczyński, B.P. 1998, ApJ, 494, L45
23. Filippenko, A.V., et al., 1999, Pub. ASP 111 (792), 969
24. Piran, T., et al., 2001, ApJL, to appear; astro-ph/0108033
25. Punsly, B., & Coroniti, F.V., 1989, Phys. Rev. D., 40, 3834; ibid. 1990, ApJ, 350, 518
26. Schmidt, M., 1999, Astron. Astroph. Suppl. Ser. 138, 409
27. Stella, L., in Proc. X-ray Astron. 1999: Stellar Endpoints, AGN and the Diffuse Background, edited by G. Malaguti, G. Palumbo, and N. White (Singapore, Gordon and Brach, New York, 2000)
28. van Putten, M.H.P.M., Science, 284, 115 (1999).
29. van Putten, M.H.P.M., 2000, Phys. Rev. Lett., 84(17), 3752
30. van Putten, M.H.P.M., & Ostriker, E.C., 2001, ApJ, 552, L31
31. van Putten, M.H.P.M., 2001, Phys. Rep., 345, 1
32. van Putten, M.H.P.M., & Levinson, A., 2001, ApJ, 555, L41
33. van Putten, M.H.P.M., 2001, Phys. Rev. Lett., 87, 091101
34. van Putten, M.H.P.M., submitted
35. Wald R.M., 1974, Phys. Rev. D., 10, 1680
36. Woosley, S.E. 1993, ApJ, 405, 273

This page is intentionally left blank

PART 4

Supernova and Gamma Ray Bursts

This page is intentionally left blank

PROGENITOR PROPERTIES AS KEY FOR THE DIVERSITY OF TYPE Ia SUPERNOVAE

P. Höflich
Department of Astronomy, University of Texas, Austin, TX 78712, USA

I. Dominguez
Dept. de Física Teórica y del Cosmos, 18071 Granada, Spain

Christopher L. Gerardy & Robert A. Fesen
Wilder Laboratory, Dartmouth College, Hanover, NH 03755, USA

A. Howell, L. Wang
Laurence Berkeley Lab, 1 Cyclotron Rd., Berkeley, CA 94720, USA

We discuss the influence of the progenitor on the observable properties of Type Ia Supernovae (SN). The study is based on spherical and 3-D calculations for the explosions, light curves, and flux and polarization spectra. The current status of models and open questions are addressed. In some detail, we discuss the a) the influence of the main sequence mass and metalicity, b) the nature of subluminous SNe Ia at the example of SN1999by and, c) the thermonuclear runaway in the exploding WD. In combination with recent observations, our study suggests that the progenitor properties are key to understand the diversity of SNe Ia. The implications are discussed for the use of SNe Ia to constrain the cosmological equation of state.

1 Introduction and Overview

The last decade has witnessed an explosive growth of high-quality data for supernovae both from the space and ground observatories with spectacular results, and new perspectives for the use of SNe Ia as cosmological yard sticks and for constraining the physics of supernovae. One of the most important new developments in observational supernova research was to establish the long-suspected correlation between the peak brightness of SNe Ia and their rate of decline (Fig. 1) by means of modern CCD photometry [51]. SNe Ia have provided new estimates for the value of the Hubble constant (H_0) based on a purely empirical procedure [11,54], and on a comparison of detailed theoretical models with observations [16,43]. The values obtained are in good agreement with one another. More recently, the routine successful detection of supernovae at large redshifts [50,49,56], has provided strong evidence for a positive cosmological constant ($\Omega_\Lambda = 0.7$), which prompted the quest for the nature of the the 'dark' energy, i.e. cosmological equation of state [48,1,44]. The results on future projects to measure the cosmological equation of state depend on the empirical

$M(\Delta M_{15})$ which is calibrated locally, and the success depends on the ability to improve its accuracy by factors of 3 to 10. This leaves systematic effects as the main source of concern , in particular, due to the evolution with redshift [5,61,10]. A good understanding of the physics of SNe Ia seems to be critical.

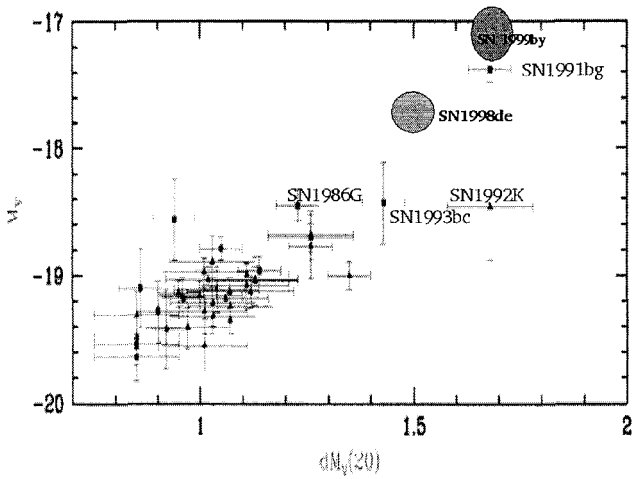


Figure 1: Observed light curve maximum brightness - decline over 20 days for 42 SN of the CTIO sample (Hamuy et al. 1996, Höflich et al. 1996). The strongly subluminous SNe 1998de and 1999by are shown including their error circles.

There is general agreement that SNe Ia result from some process of combustion of a degenerate white dwarf [23]. Within this general picture, three classes of models have been considered: (1) An explosion of a CO-WD, with mass close to the Chandrasekhar mass, which accretes mass through Roche-lobe overflow from an evolved companion star [65]. The explosion is then triggered by compressional heating near the WD center. (2) An explosion of a rotating configuration formed from the merging of two low-mass WDs, caused by the loss of angular momentum due to gravitational radiation [63,24,45]. (3) Explosion of a low mass CO-WD triggered by the detonation of a helium layer [42,66,67]. Only the first two models appear to be viable. The third, the sub-Chandrasekhar WD model, has been ruled out on the basis of predicted light curves and spectra [18,43].

Because DD-models allow to reproduce the observations, we use this scenario to test the influence of the underlying stellar population on the explosion. Delayed detonation (DD) models [25,68,69] have been found to reproduce the optical and infrared light curves and spectra of 'typical' SNe Ia reasonably well [14,15,16,9,43,64,32]. This model assumes that burning starts as subsonic defla-

gration and then turns to a supersonic, detonative mode of burning. Due to the one-dimensional nature of the model, the speed of the subsonic deflagration and the moment of the transition to a detonation are free parameters. The moment of deflagration-to-detonation transition (DDT) is conveniently parameterized by introducing the transition density, $\rho_{\rm tr}$, at which DDT happens. The amount of ^{56}Ni, M_{56Ni}, depends primarily on ρ_{tr} [14,15,60], and to a much lesser extent on the assumed value of the deflagration speed, initial central density of the WD, and initial chemical composition (ratio of carbon to oxygen). Models with smaller transition density give less nickel and hence both lower peak luminosity and lower temperatures[14,15,60]. In DDs, almost the entire WD is burned, i.e. the total production of nuclear energy is almost constant. This and the dominance of ρ_{tr} for the ^{56}Ni production are the basis of why, to first approximation, SNe Ia appear to be a one-parameter family. The observed $M(\Delta M_{15})$ can be well understood as a opacity effect[17], namely, as a consequence of the rapidly dropping opacity at low temperatures[13,26]. Less Ni means lower temperature and, consequently, reduced mean opacities because the emissivity is shifted from the UV towards longer wavelengths with less line blocking. A more rapidly decreasing photosphere causes a faster release of the stored energy and, as a consequence, steeper declining LCs with decreasing brightness. The DD models thus give a natural and physically well-motivated origin of the $M(\Delta M_{15})$ relation of SNe Ia within the paradigm of thermonuclear combustion of Chandrasekhar-mass CO white dwarfs. Nonetheless, variations of the other parameters lead to some deviation from the $M(\Delta M_{15})$. E.g. a change of the central density results in an increased binding energy of the WD and a higher fraction of electron capture close to the center which reduce the ^{56}Ni production[17]. Empirically, the $M(\Delta M_{15})$ has been well established with a rather small statistical error σ (0.12^m [55], 0.16^m [58], 0.14^m [52], 0.17^m [47] but this scatter needs significant improvements by a factor of ≈ 5 to determine the nature of the dark energy [1] This may imply a correlation between free model parameters, namely the properties of the burning front, the main sequence mass of the progenitor M_{MS}, and the central density of the WD at the time of the explosion, and it requires rather small variations in the explosion energy, i.e. the amount of burning for both normal bright and subluminous SNe Ia should be similar. Note that detailed analyzes of observed spectra and light curves indicate that mergers and deflagration models such as W7 may contribute to the supernovae population[16,12].

One of the uncertainties is the description of the nuclear burning fronts. The propagation of a detonation front is well understood but the description of the deflagration front and the deflagration to detonation transition (DDT) pose problems. On a microscopic scale, a deflagration propagates due to heat con-

duction by electrons. Though the laminar flame speed in SNe Ia is well known, the front has been found to be Rayleigh-Taylor (RT) unstable, increasing the effective speed of the burning front [41]. More recently, significant progress has been made toward a better understanding of the physics of flames. Starting from static WDs, hydrodynamic calculations of the deflagration fronts have been performed in 2-D [39,53,33] and 3-D [34,27,29]. It has been demonstrated that R-T instabilities govern the morphology of the burning front in the regime of linear instabilities, i.e. as long as perturbations remain small. During the first second after the runaway, the increase of the flame surface due to RT remains small and the effective burning speed is close to the laminar speed ($\approx 50 km/sec$) if the ignition occurs close to the center. Khokhlov [29] also shows that the effective burning speed is very sensitive to the energy release by the fuel, i.e. the local C/O ratio. Therefore, the actual flame propagation will depend on the detailed chemical structure of the progenitor. Despite advances, the mechanism is not well understood which leads to a DDT or, alternatively, to a fast deflagration in the non-linear regime of instabilities. Possible candidates for the mechanism are, among others, the Zeldovich mechanism, i.e. mixing of burned and unburned material [28], crossing shock waves produced in the highly turbulent medium, or shear flows of rising bubbles at low densities [36]. Currently, none of the proposed mechanisms have been shown to work in the environment of SNe Ia. However, as a common factor, all these mechanism will depends on on the physical conditions prior to the DDT.

For the identification of the most common scenario, the main problem is related to the insensitivity of the WD structure to the progenitor star and system. However, new, high-quality data and advances in the models for supernovae opened up new opportunities to constrain the physics of supernovae. For the first time, a direct connection with the progenitors seems to be within reach. In particular, there is mounting evidence for a connection between the properties of the progenitor, and the physics of the explosion. In the following sections, some of this evidence shall be reviewed.

2 Influence of the progenitor star

Detailed stellar evolution calculations [59] have been performed to quantify the influence of the main sequence mass M_{MS} and the metalicity Z of the progenitor on the structure of the exploding WD which are thought to be the progenitors of SNe Ia. Both the typical M_{MS} and Z can be expected to change as we go back in time. We consider the entire range of potential progenitors with 1.5 to 7 $M_{\odot}$ and metalicities between Z=0.02 to 1×10^{-10}. Our study is based on the delayed detonation scenario with specific parameters which give

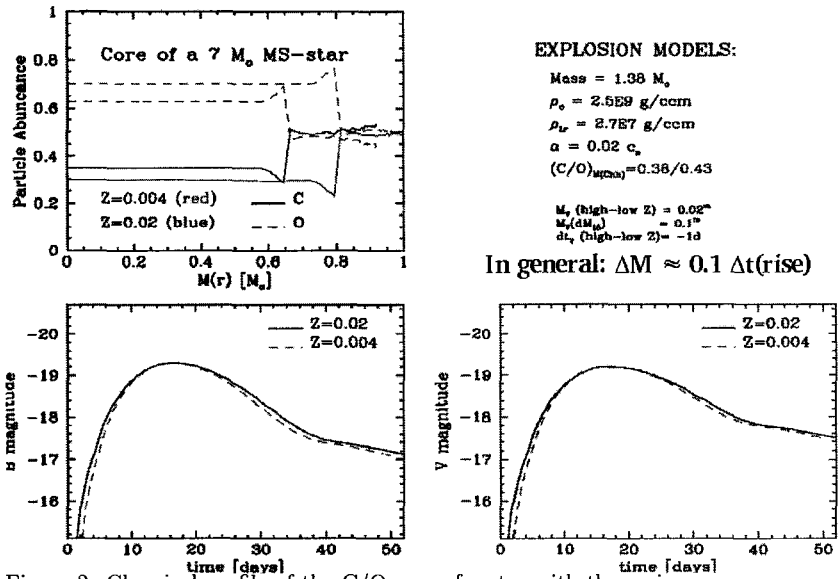


Figure 2: Chemical profile of the C/O core of a star with the main sequence mass of $7M_{\odot}$ after central Helium burning for various metalicities (upper panel). The region of reduced C abundance is produced during the central helium burning and depends on both M_{MS} and the metalicity. Eventually, this C/O core becomes the accreting WD in typical SNe Ia scenarios. In the lower panels, we compare the resulting light curves in B (left) and V (right) of a delayed detonation model (from Höflich et al. 2000).

a good account of typical light curves and spectra for a normal bright SNe Ia. Based on the structures for the WD, detailed model calculations have been performed for the hydrodynamical explosion, nucleosynthesis & LCs (Fig. 2).

All parameters but the progenitors have been fixed to produce LCs and spectra typical for 'normal' SNe Ia. In this work, rise times to maximum light are between 17.7 to 19.4 days, $M_V = -19.25^m$ to -19.11^m, and $B - V = +0.02^m$ to -0.07^m. The main sequence mass has been identified as the decisive factor to change the energetics of the explosion and, consequently, dominates the variations in the rise-time and decline of LCs (Fig. 3). M_{MS} has little effect on the color index B-V (Fig. 4). For similar decline rates ΔM_{15}, the flux at maximum brightness relative to the flux on the radioactive tail decreases systematically with M_{MS} by about 0.2^m. This change goes along with a reduction of the photospheric expansion velocity v_{ph} by about 2000 km/sec. A change in the central density of the exploding WD has similar effects but produces the opposite dependency between the brightness to tail ratio and v_{ph} and, therefore, can be separated. The metalicity alters the

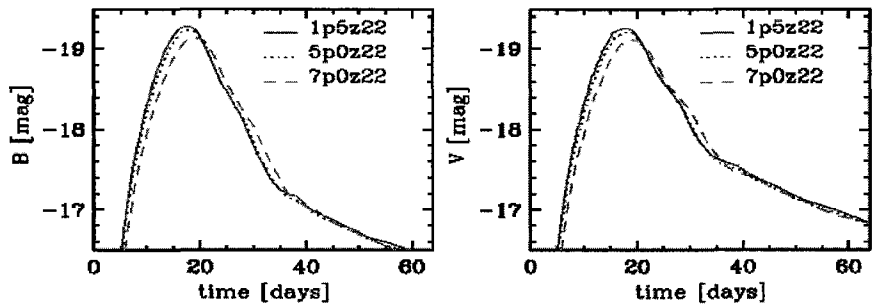


Figure 3: Influence of the main sequence mass of the progenitor on the LCs. The region of low C due to the central helium burning for the models with 1.5, 5. and 7 $M_{\odot}$ contains about 0.27, 0.46 and 0.70 $M_{\odot}$, respectively (from Dominguez, Höflich & Straniero 2001).

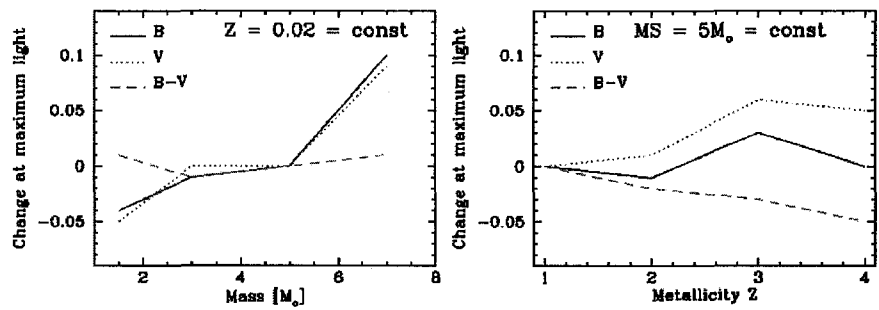


Figure 4: Influence of M_{MS} (left) and metalicity (right) on B (—), V (····) and B-V (- - -) at maximum light. All quantities are given relative to the reference model with a main sequence mass of $5M_{\odot}$ and solar metalicity. In the right panel, the numbers 1,2,3,4 refer to Z of 0.02, 0.001, 0.0001 and 10^{-10}, respectively. (from Dominguez et al. 2001).

isotopic composition of the outer layers of the ejecta. Selective line blanketing at short wavelengths decreases with Z and changes systematically the intrinsic color index B-V by up to -0.06^m, and it alters the fluxes in the U band and the UV. The change in B-V is critical if extinction corrections are applied. The offset in the calibration of $M(\Delta M_{15})$ is not monotonic in Z and, in general, remains $\leq 0.07^m$. The main correlations between progenitor and observables are given in Fig. 5).

We use our results and recent observations to constrain the progenitors, and to discuss evolutionary effects of SNe Ia with redshift. The narrow spread in the fiducial rise-time to decline relation [57] in local SNe Ia restricts the range of main sequence masses to a factor of 2. The upper limit of 1 day for the difference between the local and distance sample support the need for a positive

Feature	Effect on LC & spectra	Order of effect
Initial metallicity Z	a) little effect on B, V, R, I, ... b) strong influence on U and UV c) Strong, individual lines (e.g. 1mu FeII)	 factor 3 in Z -> 0.2 to 0.5m PopI -> strong line/ PopII-> no l.
C/O ration depends on MS mass of prog. and Z	a) Change of the rise-time/decline rel. b) Expansion velocities (Doppler shift of lines) c) Peak/Tail ratio PT	$\Delta M = 0.1\,\Delta t$[days]) Δv(Si)[km/sec]= +20,000 Δ M PT = C Δ M with C $\simeq$ +1
Change of central density rho of the initial WD -> region with neutron capture increases	a) Similar brightness at maximum for same 56 Ni production but faster, earlier rise and slower decline with increasing rho b) Peak to tail ratio changes c) Si velocity at maximum light and asymtotic Si velocity is higher	A change of rho(c) from 1.5E9 to 2.5E9 g/ccm changes width of LC by 2 days = 0.2 mag PT = C Δ M with C $\simeq$ -1 Δ(Si)[km/sec] = -20,000 ΔM
Merger/PDDs vs. classical DD	a) Slower rise and decline compared to DD b) Spectra: significant of C/O	change by about 2 to 4 days C/O down to 13,-14000 km/sec upper limit of Mg, etc.

Figure 5: Summary of observational effects due to a change in $\rho_c, M_{MS} \& Z$.

cosmological constant. The size of evolutionary effects are small ($\Delta M \approx 0.2^m$) but are absolutely critical for the reconstruction of the cosmological equation of state. For more details, see [7].

3 The subluminous SN 1999by

The nature of subluminous SNe Ia is important for the use of SNe Ia as distance indicators, and to understand the nature of SNe Ia. In particular, the calibration of the brightness decline relation depends critically on subluminous SNe Ia because they significantly increase the required baseline observations (Fig.1. From this perspective, it may turn out to be critical to determine whether all SNe Ia form a homogeneous class of objects or not, or whether a significant amount of material would remain unburned. in which we must expect a large scatter around the brightness decline relation because the explosion energy may vary strongly for SNe Ia (see above). Up to SN1999by, theoretical interpretations of subluminous SNeIa include all three types of explosion mechanism: the centrally triggered detonation of a sub-Chandrasekhar mass WD, deflagration and delayed detonations of massive WD and two merging WDs [68,15,43,38]. The possibility has also been raised that SN1991bg-like SNe Ia should not be classified with other SNe Ia at all as they may arise from different progenitors.

On April 30, SN 1999by was independently discovered at about 15^m by [2], and [46]. The distance estimates to the host galaxy, NGC 2841 [46] implies

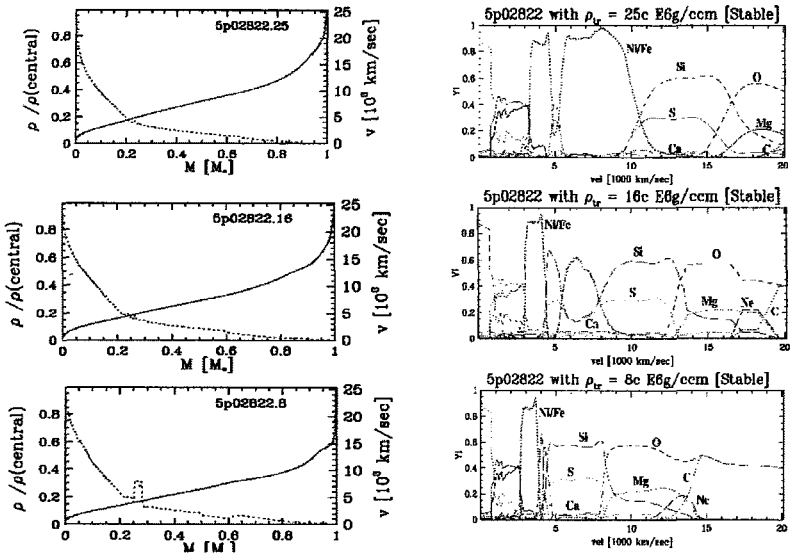


Figure 6: Density (blue, dotted) and velocity (red, solid) as a function of the mass (left panels), and abundances of stable isotopes (right panels). as a function of the expansion velocity for delayed detonation models with $\rho_{tr} = 8$, 16 and 25 $x10^6 g/cm^3$ (top to bottom). In addition, ^{56}Ni is given. All models are based on the same M_{Ch}-mass progenitor with a Main Sequence Mass of 3 $M_\odot$, solar metalicity and a central density of $2 \times 10^9 g/cm^3$ at the time of the explosion. In all cases, the entire WD is burned and, thus, all models have similar explosion energies.

magnitudes of SN 1999by are $M_B = -16.68^m$ and $M_V = -17.12^m$. Thus, SN 1999by is underluminous by roughly 2.5 magnitudes as compared to a typical Type Ia supernova but similar to SN1991bg[8,30].

This is the first subluminous SN Ia (and arguably the first SN Ia) for which IR spectra with good time coverage are available. These observations allowed us to determine the chemical structure of the SN envelope.

Based on a delayed detonation model, a self-consistent set of hydrodynamic explosions, light curves, and synthetic spectra have been calculated (Figs. 6, 7). This analysis has only two free parameters: the initial structure of the progenitor and the description of the nuclear burning front. The light curves and spectral evolution follow directly from the explosion model without any further tuning, thus providing a tight link between the model physics and the predicted observables.

By varying a single parameter, the transition density at which detonation occurs, a set of models has been constructed which spans the observed brightness variation of Type Ia supernovae. The absolute maximum brightness de-

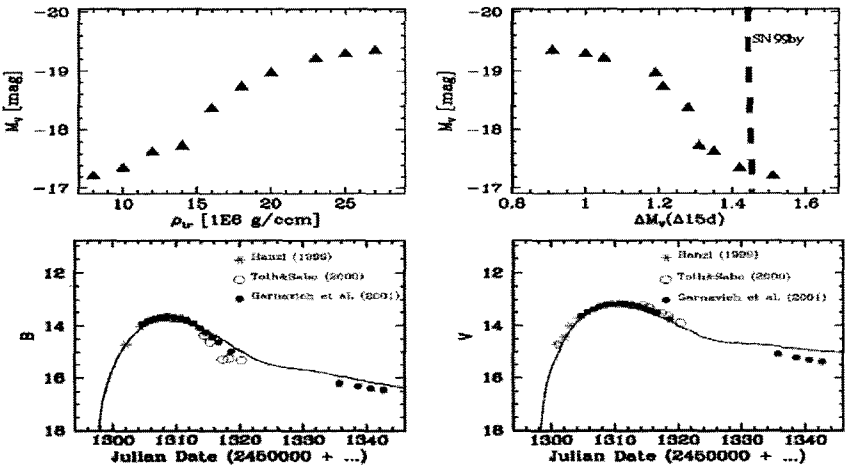


Figure 7: Maximum brightness M_V as a function of the ^{56}Ni mass (upper left), ρ_{tr} (lower left) and $M_V(\Delta M_{\Delta t=15d})$ (upper right) for delayed detonation models with ρ_{tr} of 8, 10, 12, 14, 16, 18, 20, 23, 25 and 27 $\times 10^6 g/cm^3$ from left to right. The red bar (right) gives the brightness to decline ratio as observed for SN1999by. In the lower panels, the comparison between theoretical and observer ed B and V LCs are given, derive a distance of 11 ± 2.5 Mpc, consistent with independent estimates (from Höflich et al. 2001).

pends primarily on the ^{56}Ni production which, for DD-models, depends mainly on the transition density ρ_{tr}. The brightness-decline relation $M_V(\Delta M_{\Delta t=15d})$ observed in SNe Ia is also reproduced in these models (see above).

Within M_{Ch} WD models, a certain amount of burning during the deflagration phase is needed to pre-expand the WD and avoid burning the entire star to ^{56}Ni. This implies a lower limit for the ^{56}Ni mass of about $0.1 M_\odot$ $\pm 25\%$ and, consequently, implies a minimum brightness for SNe Ia within this scenario. The best mode for SN1999by ($\rho_{tr} = 8 \times 10^7 g/cm^3$, selected by matching the predicted and observed brightness decline ratio) is close to this minimum Ni yield (Fig. 7. Without any further modification, this subluminous model has been used to analyze the IR-spectra from May 6, May 10, May 16 and May 24, 1999, which correspond to -4 d, 0 d, +7 d and +14 d after maximum light. The photosphere ($\tau_{Thomson} = 1$) recedes from 14000, to 10500, 6500 and 4000 km s^{-1}. The observed and theoretical spectra agree reasonably well with respect to the Doppler shift of lines and all strong features could be identified (Fig. 8). Before maximum light, the spectra are dominated by products of about explosive carbon burning (O, Mg), and Si. Spectra taken at +7 d and +14 d after maximum are dominated by products of incomplete Si burning. At about 2 weeks after maximum, the iron-group elements begin to show up.

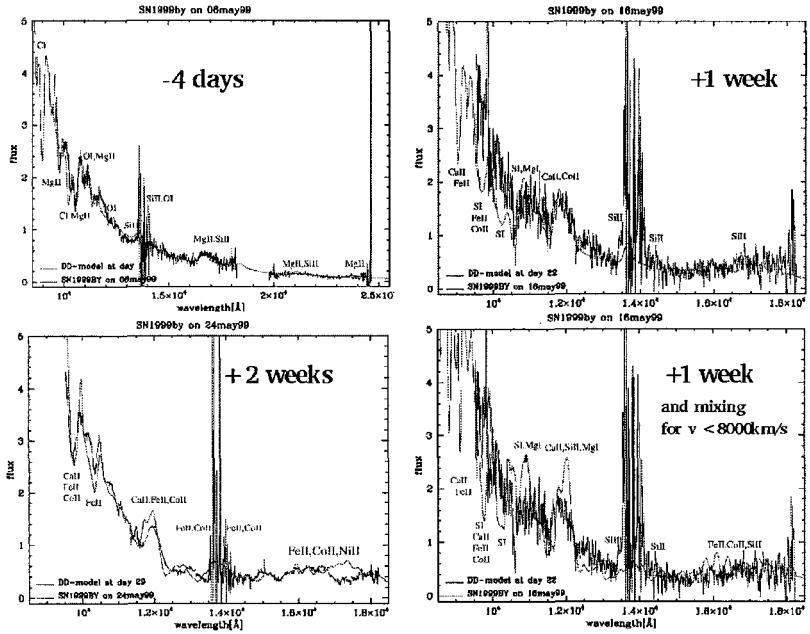


Figure 8: Comparison of the observed NIR spectrum of SN1999by on May 6 (upper left), May 16 (upper right), and May 24, 1999 (lower left). At those times, the I Thomson scattering photosphere is located at v = 13,000, 7000 and 4000 km/s, respectively. For SN1999by, the spectra are formed in layers of explosive carbon and incomplete silicon burning up to about 2 weeks after maximum light. This is in strict contrast to normal bright SNe Ia where the photosphere enters the layers of complete Si burning already at about maximum light. In the lower right, we show a comparison of the observed and theoretical spectrum if we impose mixing of the inner 0.7 $M_\odot$ (lower right) as it can be expected based on detailed 3-D models for the deflagration front (Khokhov, 2001). Obviously, strong mixing of the inner layers can be ruled out (see text).

The long duration of the phases dominated by layers of explosive carbon burning and incomplete Si burning implies massive layers of these burning stages that are comparable with our model results. (≈ 0.45 and $0.65 M_\odot$ for explosive carbon and incomplete silicon burning, respectively.) This, together with the ^{56}Ni mass, argues that SN 1999by was the explosion of a WD at or near the Chandrasekhar mass.

Finally, we note that the observed IR spectra are at odds with recent 3-D calculations [29] which predict that large scale chemical inhomogeneities filling 50 – 70% (in mass) of the WD will be formed during the deflagration phase. When the effect of such inhomogeneous mixing is tested by mixing the

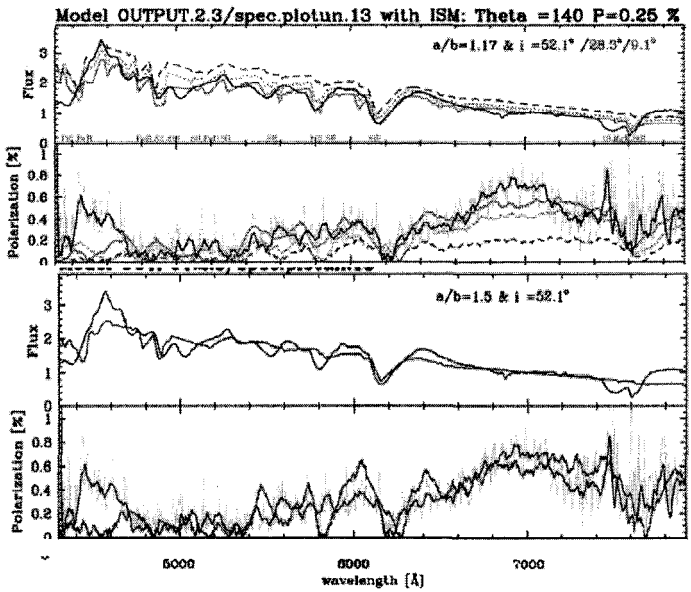


Figure 9: Flux spectra and polarization spectra at day 15 after the explosion for the subluminous delayed-detonation model in comparison with the SN1999by at about maximum light in B. The interstellar component of P has been determined to $P = 0.25\%$ with a polarization angle of 140^o. The observed flux and the smoothed polarization spectra are the solid black lines. The light blue lines are the original data for P at a resolution of 12.5 Å. In the observations, the polarization angle is constant indicating rotational symmetry of the envelope. The structure of the spherical model for SN1999by has been mapped into oblate ellipsoids with axis ratios A/B of 1.17 and 1.5. The model seen nearly equator-on ($i = 9.1°$) does the best job of reproducing the general features of the polarization spectrum:seen from the equatorial direction. depolarization from 4900 Å to 5500 Å, moderate polarization from 5500 Å to 6100 Å, a depolarization in the Si 6150 Å feature, and a rising polarization to the red (from Howell, Höflich, Wang & Wheeler 2001).

inner layers of our SN 1999by model, significant differences appear between the model spectra and the observed data (Fig. 8). This suggests that no significant large-scale mixing took place in SN 1999by. The lack of observed mixing and the asphericity seen in SN 1999by may be important clues into the nature of subluminous SNe Ia, and may be related to the reason for a low DDT transition density (and hence the low luminosity). Alternatively, we may have an extended smoldering phase of the WD prior to the explosion which skips the deflagration phase altogether. In addition to the studies of optical light curves and spectra mentioned above, detailed polarization spectra of SN 1999by have been obtained and analyzed[22] (Fig. 9. Whereas 'normal'

SNe Ia tend to show little or no polarization[62] this supernova was significantly polarized, up to 0.7%, indicating an overall asphericity of $\approx$ 20%. These result suggests that there may be a connection between the observed asphericity and the subluminosity in SNe Ia and supports the suggestion that rapid rotation of the progenitor. Pure deflagration scenarios or mergers are unlikely because those models predict a significant amount of unburned material at the outer layers, and helium-triggered explosions can be ruled out. In either case, the pre-conditioning of the WD prior to the explosion seems to be a key for understanding SNe Ia.

4 The thermonuclear runaway

From the analysis of LCs and spectra, and the study of flame fronts in SNe Ia, there are strong indications for the importance of the initial structure of the WD prior to the nuclear runaway. In this section we present the evolution of the WD just prior to the thermonuclear runaway based on multidimensional calculations. In particular, we want to address the following questions: 1) Do mixing processes change the chemical structure of the WD prior to the explosion? 2) Does the thermonuclear runaway occur in multiple spots? 3) Does the thermonuclear runaway happen in a static white dwarf?

We have studied the final hours of a Chandrasekhar-mass WD prior to the thermonuclear runaway to investigate the pre-conditioning of exploding WDs, namely chemical mixing and the ignition process.

The initial model has been constructed from results of stellar evolution for a star with 3 solar masses with solar metalicity, followed by a subsequent accretion phase close to the onset of the thermonuclear runaway[7]. The WD has a mass of $\approx 1.37 M_{\odot}$. Its chemical structure is characterized by a central region of $0.348 M_{\odot}$ with a low C-concentration ($\approx$ 24%) surrounded by a mantel with $C/O \approx 1$ originating from the He-shell burning and the phase of accretion. A few hours before runaway, the thermal structure of the progenitor shows a rather flat temperature profile, and a steep entropy profile because the rapidly dropping density.

Prior to the runaway, the central regions undergo mild C-burning. The resulting energy release drives convective motion in the inner region of low C-concentration and, gradually, increases the entropy of the core up to the point of ignition (Fig. 10). Due to the convection, the entropy of the core is almost constant. Within the resolution of our models, the carbon-concentration gradient at the boundary between the core and the mantle prevents direct mixing, e.g. due to overshooting convective elements. However, the increasing entropy of the core results in a negative entropy gradient at the core boundary which

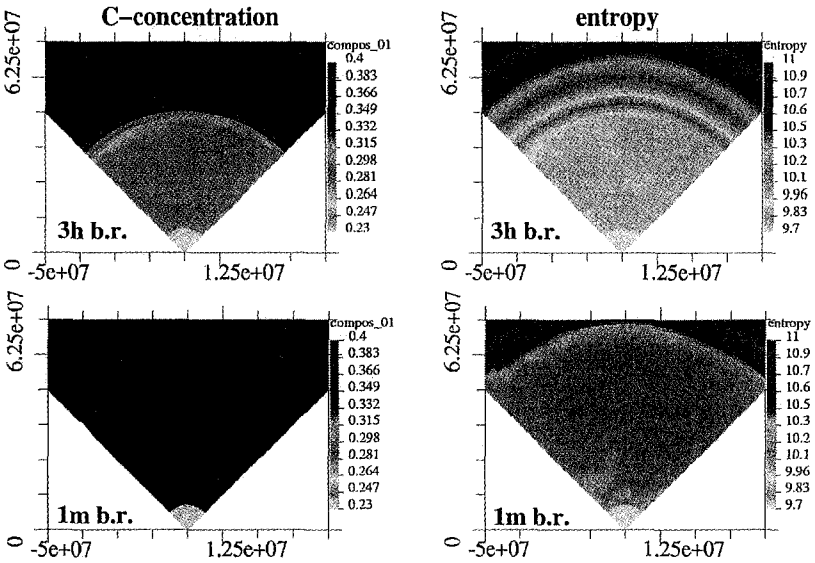


Figure 10: Evolution of the chemical profile in a M_{Ch} mass WD with a radius of $\approx 1600km$. Carbon concentration (left) and entropy (right) at three hours (upper) and one minute (lower) before runaway in the inner 700km. Runaway occurs when the mean entropy in the turbulent core rises to about 10.1. Note that the final size of the mixed region corresponds to the distance at with the same entropy can be found in the initial model (from Höflich & Stein 2001).

compensates for the carbon-concentration gradient. This increases the region with constant entropy and produces mixing of C-rich region into the core with typical fluctuation of about 1%. We find that the central C-abundance increases from 24 to about 37 %. The initial WD is not homogenized, but the jump in the carbon abundance is reduced by a factor of ≈ 2. Note, however that the amount of burning in a non-rotating WD cannot account for the pre-expansion needed for SNe 1999by.

At the time of the explosion, a pattern of large scale, convective elements has been established with sizes of typically 100 km and convective velocities between ≈ 40 to ≈ 120 km/sec (Fig. 11). Differential velocities between adjoining eddies are larger by a factor of 2 which is well in excess of the laminar deflagration speed. Thus, the change of the morphology of the burning front of SNe Ia is determined by the pre-conditioning of the WD during the early

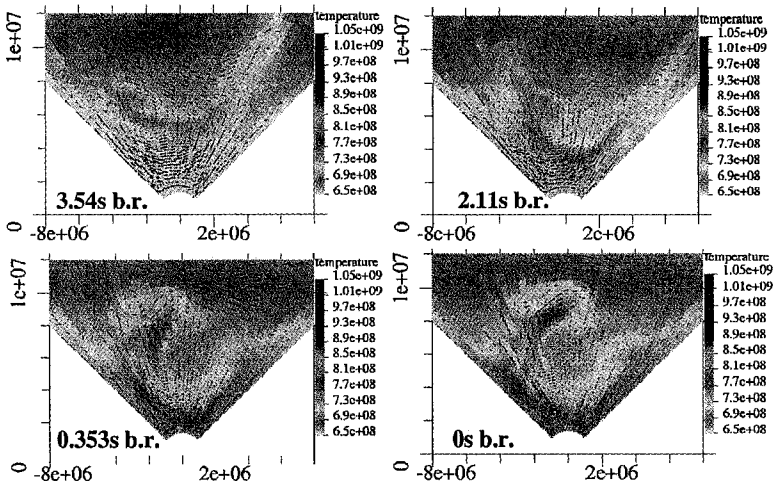


Figure 11: Final evolution of the temperature structure up to the runaway. In the lower left plot, the runaway occurs in the second red zone from the left right at the inner boundary (T=1.74^9 K). In addition, the velocity field is given. Black, red and green vectors correspond to velocity ranges of 0 to 50, 50 to 100, and 100 to 150 km/s, respectively. The runaway occurs close to the center (see text). The velocity field is formed by rising bubbles, and to first order, the convective velocity field is determined by the typical final/terminal velocity. Note that the The asymptotic velocity of a bubble should be larger than its its mean velocity. Thus, in general, the pre-existing velocity will dominate the motion of a bubble early on.

phase of the explosion for $\approx$ 0.5 to 1 sec [6,29]. Niemeyer et al. [40] found significantly shorter time scales for the growth of RT-instabilities in their study of strongly off-center explosions because he assumed that the ignition is significantly off-center. The effective surface of the front will be increased resulting in significantly higher burning speeds. Faster burning implies a larger region of low proton to nucleon ratio and, thus, a larger production of neutron rich isotopes in the central region. On the other hand, a reduction of the time scales for electron capture can be expected leading to the production of less neutron rich isotopes. Possible consequences for current estimates on the limits on the central densities of the WD should be noted [4].

The explosive nuclear burning front starts in one well defined region close to the center ($\approx 30km$). The size of the ignition region is determined by the grid resolution ($\approx 2km$). The explosive phase of burning is triggered by compressional heat when matter is brought inwards by convection. It starts close to the center because, there, the adiabatic heating combined with thermonuclear reactions are most effective for a given size of turbulent elements. We find no evidence for multiple spot or strong off-center ignition. We do not

expect it because the size of the eddies is comparable to the central distance of the ignition point, and the lack of any mechanism which would cause a synchronization within typical time scales for the runaway ($\leq 0.1sec$). Thus, the probability is fairly small for having a second ignition point during that time.

In the following, we want to put our basic results into context for our understanding and the quantitative modeling of SNe Ia. As mentioned in the introduction, the propagation of the deflagration front depends on the energy release and, consequently on the fuel (Khokhlov 2001). We find that the chemical profile in the WD will be strongly changed, but in a predictable way. We find that the initial velocity field must be expected to alter the flame propagation during the deflagration phase. Although the actual deflagration speed has little effect on the overall chemical structure of DD-models with the exception of the production of neutron rich isotopes close to the center, all proposed mechanisms for the DDT identify the pre-conditioning of the material during the deflagration phase as a key element (see introduction) which, in turn, is strongly effected by the initial WD. As mentioned above, the overall chemical structure of the initial WD is preserved, and the turbulent velocity field is limited to the inner, C-depleted core. Both the velocity field and the C-concentration influence the burning front. Therefore, the mass of the progenitor has a direct influence on the outcome because the core size depends mainly on the M_{MS} mass of the progenitor. The consequences are obvious with respect to the evolution of the SNe Ia with redshift and their use as a yardstick to measure cosmological parameters and the cosmological equation of state.

Acknowledgments

This research was supported in part by NASA Grant LSTA-98-022.

References

1. Albrecht, A., Weller, J. 2000, BAAS, 197, 6106
2. Arbour, R., Papenkova, M., Li, W. D.,
3. Bonanos, A., Garnavich, P., Schlegel, E., Jha, S., Challis, P., Kirshner, R., Hatano, K., Branch, D. 1999, BAAS 195, 3806
4. Brachwitz, F. 2001, PhD thesis, University of Basel/Switzerland, supervised by F.K. Thielemann
5. Branch, D., Romanishin, W., Baron, E. 1996, ApJ 465, 73
6. Domínguez, I., Höflich, P. 2000, ApJ, 528, 854
7. Domínguez, I., Höflich, P., Straniero, O. 2001, ApJ, 557, 279
8. Filippenko, A. V. et al. 1992, AJ, 104, 1543

9. Fisher, A., Branch, D., Höflich, P., Khokhlov, A. 1998, ApJ, 494, 47
10. Hamuy M. et al. 2000, AJ 120, 1479
11. Hamuy, M. et al. 1996, AJ, 112, 2391
12. Hatano, K., Branch, D., Lentz, E. J., Baron, E., Filippenko, A. V., Garnavich, P. 2000, ApJ 543L, 49
13. Höflich, P., Khokhlov, A., Wheeler, J.C. 1993, A&A, 270, 223
14. Höflich, P. 1995, ApJ, 443, 89 [H95]
15. Höflich, P., Khokhlov, A., Wheeler, J.C. 1995, ApJ, 444, 211
16. Höflich, P., Khokhlov, A. 1996, ApJ, 457, 500 (HK96)
17. Höflich, P., Khokhlov, A., Wheeler, J. C., Phillips, M. M., Suntzeff, N. B., Hamuy, M. 1996a, ApJ, 472, L81
18. Höflich, P.; Dominik, C.; Khokhlov, A.; Müller, E.; Wheeler, J. C. 1996b, 17^{th} Texas Symposium on relativistic Astrophysics, Annals of the New York Academy of Science, 759, 348
19. Höflich, P., Wheeler, J. C., Thielemann, F. K. 1998, ApJ, 495, 617 (HWT98)
20. Höflich, P., Nomoto, K., Umeda, H., Wheeler, J. C. 2000, ApJ, 528, 590
21. Höflich, P., Stein, J. 2001, ApJ, submitted (astro-ph/0104226)
22. Howell, A., Höflich P., Wang L., Wheeler J. C. 2001, ApJ, 556, 302
23. Hoyle, P., Fowler, W.A. 1960, ApJ, 132, 565
24. Iben Jr, I. Tutukov, A.V. 1984, ApJs, 54, 335
25. Khokhlov, A. 1991, ApJ, 245, 114
26. Khokhlov, A., Müller, E., Höflich, P. 1993, A&A, 270, 223
27. Khokhlov, A. 1995, ApJ, 449, 695
28. Khokhlov A., Oran E.S., Wheeler J.C. 1997, ApJ, 478, 678
29. Khokhlov, A. 2001, ApJ, in press (astro-ph/0008463)
30. Leibundgut, B. et al. 1993, AJ, 105, 301
31. Lentz, E. J., Baron, E., Branch, D., Hauschildt, P., Nugent, P. E. 2000, ApJ, 530, L966
32. Lentz E. J., Baron E., Branch D., Hausschildt P. 2001, ApJ, 557, 2001
33. Lisewski, A. M.; Hillebrandt, W.; Woosley, S. E.; Niemeyer, J. C.; Kerstein, A. R. 2000, ApJ 537, 405L
34. Livne E. 1993, ApJ 406, 17L
35. Livne, E. 1998, ApJl, 527, 97
36. Livne E. 1998, talk in workshop on SNe Ia private communication with P. Höflich, Aspen 1998.
37. Livne, E. 1999, ApJl, 527, 97
38. Milne, P., The, L.-S., Leising, M.D. 1999, ApJs 124, 503
39. Niemeyer, J. C., Hillebrandt, W. 1995, ApJ, 452, 779

40. Niemeyer J. , Hillebrandt, W., Woosley S. E. 1996, ApJ 471, 903
41. Nomoto K., Sugimoto S., Neo S. 1976, ApSS, 39, L37
42. Nomoto, K. 1980, ApJ, 248, 798
43. Nugent P., Baron E., Hauschildt P., Branch., D. 1997, ApJ, 485, 812
44. Ostriker P., Steinhardt P.J. 2001, Scientific American, Jan. 2001, p. 47
45. Paczyński, B., 1985, in: Cataclysmic
46. Papenkova, M., Li, W. D., Filippenko, A. V. 1999, IAU-C 7156
47. Perlmutter, S., et al. 1999a, ApJ 517, 565
48. Perlmutter S., Turner M.S., White M. 1999, Physical Review Letters 83, 670
49. Perlmutter C. et al. 1997, ApJ 483, 565
50. Perlmutter C. et al. 1995, ApJ 440, L95
51. Phillips, M. M. et al. 1987, PASP 90, 592
52. Phillips, M.M., Lira P., Sunzeff N.B., Schommer R.A., Hamuy M., Maza J. 1999, AJ 118, 1766
53. Reinecke M., Hillebrandt W., Niemeyer J., Klein R., Gröbl A. 1999, AA 347, 724
54. Riess A.G., Press W.H., Kirshner R.P. 1995, ApJ 438, L17
55. Riess A.G., Press W.H., Kirshner R.P. 1996, ApJ 473, 88
56. Riess A.G., et al. 1998, AJ 116, 1009
57. Riess A.G., et al. 1999, AJ 118, 2675
58. Schmidt, B. P., et al. 1998, ApJ 507, 46
59. Straniero, O. Chieffi, A. and Limongi M. 1997, ApJ 490, 425
60. Umeda, H., Nomoto, K., Kobayashi, C., Hachisu, I., Kato, M. 1999, ApJ, 522, L43
61. Wang, L., Wheeler, J. C., Höflich, P. 1997, ApJ, 476, L27
62. Wang, L., Howell, A., Höflich, P., Wheeler J.C. 2001, ApJ 550, 1030
63. Webbink, R. F. 1984, ApJ, 277, 355
64. Wheeler, J. C., Höflich, P., Harkness, R. P., Spyromilio, J. 1998, ApJ, 496, 908
65. Whelan, J., Iben Jr., I. 1973, ApJ, 186, 1007
66. Woosley, S. E., Weaver, T. A., Taam, R. E. 1980, in: Type I Supernovae, ed. J. C. Wheeler, Austin, U. Texas, p.96
67. Woosley, S. E., Weaver, T. A. 1986, ARAA, 24, 205
68. Woosley, S. E., Weaver, T. A. 1994, ApJ, 423, 371
69. Yamaoka, H., Nomoto, K., Shigeyama, T., Thielemann, F. 1992, ApJ, 393, 55

GAMMA-RAY BURST LIGHT CURVES SEEN BY FOURIER ANLYSIS: COSMOLOGICAL TIME DILATION AND A USE AS A NEW DISTANCE ESTIMATOR

HEON-YOUNG CHANG
Korea Institute for Advanced Study
207-43 Cheongryangri-dong Dongdaemun-gu, Seoul 130-012, Korea
hyc@ns.kias.re.kr

SUK-JIN YOON
Center for Space Astrophysics and Department of Astronomy, Yonsei University
134 Shinchon-dong Seodaemun-gu, Seoul 120-749, Korea
sjyoon@csa.yonsei.ac.kr

CHUL-SUNG CHOI
Korea Astronomy Observatory
36-1 Hwaam-dong, Yusong-gu, Taejon 305-348, Korea
cschoi@kao.re.kr

We study the power density spectrum (PDS) of light curves of the observed gamma-ray bursts (GRBs) to search for a direct signature for cosmological time dilation in the PDS statistics with the GRBs whose redshifts z are known. We demonstrate that a time dilation effect is seen in GRB light curves. We also propose a redshift estimator for the long GRBs ($T_{90} > 20$) observed by the BATSE, based on the empirical relation between the redshift and the power-law index of PDSs of the observed GRBs. We attempt to determine the spatial distribution of the GRBs observed by the BATSE as a function of redshifts on the basis of the resulting redshift estimator. We find that the obtained spatial distribution of the observed GRBs seems consistent with that of the GRBs whose redshifts are reported, even though the estimated errors are marginal. The GRBs observed by BATSE seem distributed within $z \sim 5 - 6$. This result has implications on theoretical calculations of stellar formations at high redshifts and beaming geometry via a statistical study of the observed GRBs involving beaming-induced luminosity functions. Even though redshift estimates are subject to the stochastic nature of the observed PDSs and accuracy of estimates are limited by unknown properties of the GRBs, the encouraging conclusion of this study is that redshifts of the GRBs can be obtained with the GRB light curves whose redshifts otherwise remain unknown forever.

1 Introduction

Time dilation is a fundamental property of an expanding universe. Cosmological objects should not only be redshifted in energy but also extended in time because of the expansion of the universe. There have been observational experiments on time dilation with cosmological objects, such as high-redshift quasar or supernova observations. Since gamma-ray bursts (GRBs) are at cos-

mological distances the burst profiles should also be stretched in time due to cosmological time dilation by an amount proportional to the redshift, $1 + z$.

Without knowing their redshifts, different groups have investigated the correlation of the properties of GRB light curves and the burst brightness in order to look for a signature of time dilation (Norris et al. 1994; Mitrofanov et al. 1996; Lee & Petrosian 1997; Deng & Schaefer 1998; Stern et al. 1999). For instance, Norris et al. (1994) searched for time-dilation effects by dividing the bursts into groups based on their peak count rate and comparing some measure of burst duration with peak count rate. They have claimed that brighter bursts had shorter durations than dimmer ones and that the difference between the average durations of bright and dim bursts was consistent with a time-dilation factor of ~ 2. If bursts were standard candles, dimmer bursts would be time-dilated more than brighter bursts by a dilation factor $(1 + z_{\rm dim})/(1 + z_{\rm bright})$, where $z_{\rm dim}$ and $z_{\rm bright}$ are the redshifts of dim and bright bursts. However, one of the most serious limits of previous works is that inferences are all *indirect* and possibly misleading since the redshifts of most GRBs are unknown. Finding cosmological time dilation signature in light curves of GRBs is disputed. For instance, Mitrofanov et al. (1996) finds no time dilation in BATSE using the aligned peak test. Moreover, there are warnings and other possible explanations. Band (1994) has warned that an intrinsic burst luminosity function could easily produce similar effects. Questions have been raised as to whether or not the time stretching that is found is due to the intrinsic correlation between pulse width and burst brightness for bursts drawn from a volume-limited sample (Brainerd 1994). Yi & Mao (1994) also demonstrated that relativistic beaming in either Galactic halo or cosmological models can produce flux-duration relationships that might be consistent with the reported effects.

Here we present direct results on this topic which differ from those of previous works in two important ways. Firstly, we analyze the Fourier power spectra of a sample of GRB light curves to look for such an effect. Secondly, we use light curves of the GRBs whose redshift z is known so that we are able to infer the time dilation effect directly. Statistical significance is reduced because of a small size of GRB data sets. Nonetheless we have a direct measure of time dilation, since we use the GRB light curves for z-known samples. We also attempt to propose a new method as a redshift estimator of the observed GRBs, using the properties we demonstrate here. At present, there are too few redshift measurements with which to produce the global GRB formation rate. This fact is indeed hard to avoid unless observers set up networks of efficient telescopes in order for an immediate follow-up observations. Recently, there are pilot studies to overcome the problem mentioned above. Several authors (Stern

et al. 1999; Fenimore and Ramires-Ruiz 2000; Reichart et al. 2001) began to observe strong correlations between temporal properties of the observed GRBs and their brightness, which may have some implications that the measured spikiness can be used to obtain distances much like a Cepheid-like distance estimator. Currently, the luminosity estimator yields best-estimate luminosity distances that are accurate to a factor of 2 (see Reichart et al. 2001), while the rms errors of the redshifts estimated by our new method are around 1.

2 Power Density Spectrum of GRB Light Curves and Cosmological Time Dilation

The Fourier Transform analysis is widely used in studies of periodical phenomena and analyses of stochastic processes. It has been applied to studies of X-ray observation of AGNs and GRBs to investigate their central engines (e.g., Lawrence et al. 1985; Beloborodov et al. 1998). Contrary to the diverse and stochastic behavior in the time domain, long GRBs show a simple behavior in the frequency domain (e.g., Beloborodov et al. 1998). The power-law behavior is seen even in a single burst when it is bright and long. The power-law power density spectrum (PDS) provides a new tool for studies of GRBs themselves. Using the PDS analysis, Panaitescu et al. (1999) analyzed the temporal behavior of GRBs in the framework of a relativistic internal shock model. They set up their internal shock model and attempted to identify the most sensitive model parameters to the observed PDS. They concluded that the wind must be modulated such that collisions at large radii release more energy than those at small radii in order to reproduce consistent PDSs with the observation. However, it is also noted that the reported power-law behavior with the index of $-5/3$ and the exponential distribution of the observed PDS can be reproduced by adjusting the sampling interval in the time domain for a given decaying timescale of individual pulse in a specific form of GRB light curves (Chang and Yi 2000). Therefore, conclusions on the central engine of GRBs on the basis of the PDS analysis should be derived with due care.

Another valuable use of the PDS analysis can be realized bearing in mind that for a predetermined sampling interval the resulting power-law index is dependent upon the characteristic timescale of the observed light curve. For instance, consider a GRB light curve as a sum of exponential functions of time, $f(t)$, as considered in Chang and Yi (2000). Since the Fourier transform of $f(at)$ is given by $\frac{1}{|a|}F(\nu/a)$, where $F(\nu)$ is the Fourier transform of $f(t)$, ν being the cyclic frequency and a being a constant, fitting of the power law function to the PDS may result in a different power-law index as the constant a varies when the sampling interval is pre-determined. In other words, for the

Table 1: A list of the GRBs used in the analysis with the redshifts and T_{90}. The redshifts are quoted from a complied table in http : //www.aip.de/ $\sim$ jcg/grbgen.html. The duration T_{90} is taken from the updated BATSE 64 ms ASCII database.

GRB name	trigger number	redshift	T_{90}(sec)
GRB 000418	8079	1.118	2
GRB 991216	7906	1.02	15
GRB 990510	7560	1.619	68
GRB 990506	7549	1.3	130
GRB 990123	7343	1.60	63
GRB 980703	6891	0.966	411
GRB 980425	6707	0.0085	34
GRB 980329	6665	3.9	18
GRB 971214	6533	3.42	31
GRB 970508	6225	0.835	23

observed GRB light curves with the pre-determined sampling intervals, e.g., 64 ms, the cosmological time dilation stretches the light curve by an amount of $1+z$ and consequently results in changes in the obtained power-law index. This should be true because cosmological objects like GRBs should be strechted in time because of the expansion of the universe.

We have used light curves of GRBs from the updated BATSE 64 ms ASCII database[a]. From this archive we select the light curves of the GRBs whose redshifts are available. We list up the GRBs used in our analysis with BATSE trigger numbers and the reported redshifts in Table 1. We divide our sample into two subgroups so that we separate near and far GRBs. We calculate the Fourier transform of each light curve of GRBs and the corresponding PDS. Before averaging the calculated PDSs in each subgroup, we normalize GRB light curves by setting their peak fluxes to unity. We compare the slopes obtained by the linear fits as it is without time dilation correction with those after rescaling the individual GRB light curve to factor out $1 + z$. We have repeated the same process for the light curves of four different energy bands.

In resulting PDSs before time-dilation correction, power in lower frequencies is high for the far GRB subgroup, and for the near GRB subgroup power is concentrated in high frequencies. This is exactly what one may expect if light curves of GRBs are lengthened due to cosmological time dilation. We attempt power-law fits in the limited range, i.e., $-1.6 < \log f < 0$. The lower bound is determined in that the deviation from the power law begins due to the finite length of bursts. The upper bound is given such that the Poisson noise becomes dominant. Before the fitting the averaged PDS is smoothed on

[a] ftp : //cossc.gsfc.nasa.gov/pub/data/batse/

Table 2: Obtained slopes with the least square fits for the two subgroups of far and near GRBs. Fittings are repeated before and after correction of a time dilation effect by a factor of $1+z$.

	as-it-is		corrected	
channel	far	near	far	near
1	-1.706 ± 0.11	-1.543 ± 0.10	-1.564 ± 0.14	-1.508 ± 0.09
2	-1.805 ± 0.10	-1.522 ± 0.10	-1.585 ± 0.14	-1.465 ± 0.09
3	-1.881 ± 0.10	-1.510 ± 0.09	-1.600 ± 0.15	-1.416 ± 0.09
4	-1.371 ± 0.12	-1.098 ± 0.07	-1.398 ± 0.15	-1.113 ± 0.05

a scale of $\Delta \log f = 0.2$. The slopes and standard deviations obtained by the linear fit is summarized in Table 2. Exact values of the obtained slopes are subject to the range used in the fitting process. However, the trend is hardly affected. For all channels, the subgroup of the GRBs with higher redshifts result in exclusively steeper slopes compared with those with lower redshifts.

To see the effects of time dilation, we rescale the time interval of individual GRB light curve by a factor of $(1+z)^{-1}$, where z is the redshift of the individual GRB. This should remove the effect of time dilation, that is, the difference of slopes in the two subgroups resulting from cosmological time dilation. This manipulation has the effect of shifting the contributions of all GRBs to the range of higher frequencies. Resulting slopes of the fits summarized in Table 2. We note that the removal of the $(1+z)$ factor makes discrepancies of slopes in two subsamples reduced indeed, but differences still marginally remain.

3 Empirical Relation of Power-law Index and Redshift

In order to construct the empirical relation between the power-law index and the redshift of GRBs we have calculated the power-law index of the PDSs of 9 GRBs detected by the BATSE with known redshifts. We omit a short GRB, of which T_{90} is 2 seconds.

The PDSs can be described as a single power law with superimposed fluctuations which follow the exponential distribution, which may require the maximum likelihood method. However, by considering that we smooth the PDSs on the scale $\Delta \log \nu = 0.2$ before fitting, the least squares fit can be used since the error distribution may be modified to the normal distribution according to the *central limit theorem*. We use two different fitting routines corresponding to the normal error distribution and the exponential error distribution. What is shown in Figure 1 are empirical relations of the redshift and the power-law index obtained by the least squares fit and the maximum likelihood method. Firstly, we attempt to construct the empirical relation of the redshift and the

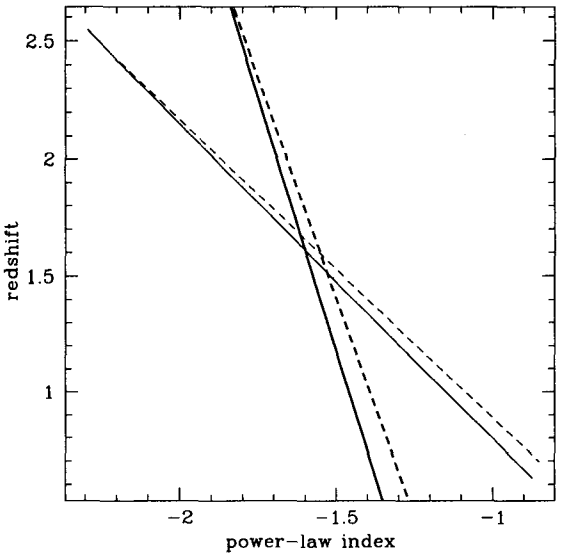


Figure 1: Relation of power-law index and redshift. The thin solid line results from fitting of individual PDSs by the least squares fit, the thin dashed line the maximum likelihood method. For comparison, the relation of redshift and power-law index with averaged PDSs is also shown. The thick solid line and the thick dashed line represent the least squares fit, the maximum likelihood method, respectively.

power-law index without averaging of the PDSs. That is, we construct the best fit using power-law indices of 9 individual data points. Results are shown in Figure 1, where the thin solid line results from the least squares fit, the thin dashed line the maximum likelihood method. A possible alternative way to extract the empirical relation from the noisy individual PDSs is to take the average PDSs over a sample of long GRBs. Because of the small number of data, we group the PDSs into 4 subgroups according to the reported redshifts, and average the PDSs before a fitting process : GRB 980329 + GRB 971214, GRB 990123 + GRB 990506 + GRB 990510, GRB 970508 + GRB 980703 + GRB 991216, GRB 980425. The thick solid line results from fitting of averaged PDSs by the least square fit, the thick dashed line the maximum likelihood method. We have attempted higher order polynomial fits but it did not end up with a monotonic relation as one should expect.

The error estimates of z, which are defined by a square root of the average of squared difference between the measured redshifts summarized in Table 1 and the expected redshifts by the fitting, are 1.09 and 1.11 for the empirical

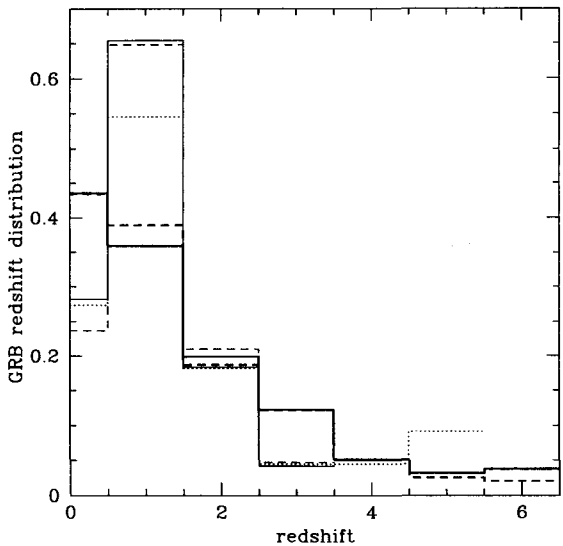


Figure 2: Normalized spatial distributions of 388 long GRBs observed by the BATSE are compared with that of the redshift-known GRBs, indicated by the same line types as in Figure 1 and the thin dotted line, respectively.

relation due to individual PDS fits, 1.72 and 1.56 due to the averaged PDS fits, for the least squares fit and the maximum likelihood fit, respectively.

4 Spatial Distribution of GRBs

We adopt light curves of the long GRBs from the updated BATSE 64 ms ASCII database as in processes above. We choose bursts with durations $T_{90} > 20$ s, where T_{90} is the time it takes to accumulate from 5 % to 90 % of the total fluence of a burst summed over all the four channels. Of those bursts, we further select bursts with the peak count rates satisfying $C_{\max}/C_{\min}$ for the 64 ms trigger timescale is greater than 1. Applying these criteria, we end up with 388 bursts.

In a similar way, we obtain power-law indices of PDSs of the selected GRBs and subsequently estimates of their redshifts. In Figure 2, the redshift distributions of the GRBs obtained by the relation we have in the previous section are shown. Different line types indicate same meanings as in Figure 1. Note that the thin dotted histogram represents the spatial distributions of the 22 GRBs whose redshifts are available at the web site, from where the

quoted redshifts in Table 1 are taken. It is interesting to note that the predicted redshift distributions of GRBs that are derived from fitting of individual power-law index without averaging PDSs appear to provide a better agreement with the redshift distribution of the GRBs with redshift-known.

The obtained GRB redshift distribution is quite suggestive. Firstly, it is seen in Figure 2 that the long GRBs observed by the BATSE, at least, are distributed well within $z \sim 5 - 6$. If one accepts an idea that the GRB formation rate should trace the SFR at high redshifts as at low redshifts, this is in apparent contrast to what is derived from some theoretical calculations of star formation. Theoretical calculations show that the birth rate of Pop III stars produces a peak in the SFR in the universe at redshifts $16 < z < 20$, while the birth rate of Pop II stars produces a much larger and broader peak at redshifts $2 < z < 10$ (Ostriker and Gnedin 1996). Secondly, according to the cumulative redshift distributions derived in terms of beaming-induced luminosity functions an extreme shape of a conic beam seems likely to be ruled out : although the observed $< V/V_{\rm max} >$ can be satisfied with the theoretical $< V/V_{\rm max} >$, a broad beam cannot explain the observed redshifts greater than $\sim 2 - 3$ (Kim et al. 2001) and a hollow beam expects too many GRBs farther than $z \sim 5 - 6$ (Chang and Yi 2001), if a SFR-motivated number density distribution of GRB sources is assumed.

5 Discussions and Conclusions

Claiming time dilation in light curves of GRBs with the anticorrelation of a timescale measure and a brightness measure has several difficulties. One difficulty is that this effect is correct only for standard candle sources with a standard duration, which we have evidence that it is not necessarily true. A broad luminosity function and/or an intrinsic spread in the durations could smear out the signature. Another possible difficulty with this anticorrelation is that it could be mimicked by intrinsic properties of the sources. An additional complication is that an intrinsic redshift of the time profiles from higher energy bands to lower energy bands may be present (Fenimore & Bloom 1995), which would bleach the cosmological signature.

We investigate the correlation between redshifts and timescale measures using available GRB data with known z. Unlike past indirect searches for cosmological time dilation, we use the GRBs whose z is known at the expense of statistical significance. We find that timescales tend to be shorter in bursts with small redshift, as expected from cosmological time-dilation effects, but we also find that there may be non-cosmological effects constituting to this correlation. Because of the small number of data, it is inconclusive that

these imperfect corrections require other explanations other than cosmological time dilation. The amount of observed stretching *may not* be the value expected from cosmological time dilation alone (Horack et al. 1996; Mészáros & Mészáros 1996). Challenging questions then are whether one may extract information on intrinsic properties of individual GRBs or whether one may distinguish a cosmological model by an analysis of the slope of the observed PDSs of GRBs.

Another important implication of our study should be pointed out. Beloborodov (1998) applied the Fourier transform technique to the analysis of light curves of long GRBs. With due care, the analysis of such kind may yield valuable information of the central engine of GRBs (Panaitescu et al. 1999; Chang & Yi 2000). However, the averaged power law index and the distribution of individual power should be corrected first in terms of a time dilation effect before making any physical points out of the results of the PDS analysis. We have followed similar procedures for the total sample as Beloborodov et al. (1998) did and obtained the slopes, which are close to the reported value $-5/3$ indeed. However, these slopes become flatter when time dilation correction is made before the analysis.

We demonstrate that this fact can be used as a redshift estimator, at least in principle. There are, however, problems and limits to determine redshifts accurately in both obtaining the relation and applying this relation to data. First of all, even though this method is in principle to work, it is not clear whether we may apply this method over the observed GRB light curves obtained by various satellite missions at the same time. It is partly because the PDS of each individual burst is composed of the power law and superimposed exponentially distributed fluctuations which make it difficult to recognize the power law in an individual burst, and partly because such the empirical relation is susceptible to observational conditions, such as, trigger timescale, detection sensitivity. Availability of more redshifts of GRBs may help to reproduce a better relation. With all the efforts in implementing a sophisticated algorithm to accommodate the diversity of the light curves, it is essential to understand a fundamental mechanism of GRBs to derive the intrinsic relation of the power-law index and the redshift. Secondly, We have implicitly assumed that all the long GRBs have a more or less same characteristic timescale and cosmological time dilation alone affects varying the characteristic timescale. We need to understand clearly what and how forms the flat part of PDS. Thirdly, the effect of redshift tends to flatten PDSs of GRBs on the contrary to the time dilation effect. It reflects a well-known fact that pulses in a single GRB are more narrow in a higher energy band (e.g., Norris et al. 1996). These effects combine and produce undesirable results in obtained power-law index. We have presumed

in this study that the time dilation effect on the power-law index is larger than that of the redshift as observed in Chang (2001) and ignored the effect of the redshift. However, it should be understood how the power-law index relates with the energy channels to improve the proposed method accommodating the redshift effect.

Acknowledgments

SJY is supported by the Creative Research Initiatives Program of the Korean Ministry of Science and Technology. This research has made use of data obtained through the High Energy Astrophysics Science Archive Research Center Online Service, provided by the NASA/Goddard Space Flight Center.

References

1. Band, D. 1994, ApJ, 432, L23
2. Beloborodov, A., Stern, B., & Svensson, R. 1998, ApJ, 508, L25
3. Brainerd, J. J. 1994, ApJ, 428, L1
4. Chang, H.-Y. 2001, ApJ, 557, L85
5. Chang, H.-Y. & Yi, I. 2000, ApJ, 542, L17
6. Chang, H.-Y. & Yi, I. 2001, ApJ, 554, 12
7. Deng, M., & Schaefer, B. E. 1998, ApJ, 502, L109
8. Fenimore, E. E., & Bloom, E. 1995, ApJ, 448, L101
9. Fenimore, E. E. & Ramirez-Ruiz, E. 2000, astro-ph/0004176
10. Horack, J. M., Mallozzi, R. S., & Koshut, T. M. 1996, ApJ, 466, 21
11. Kim, C., Chang, H.-Y., & Yi, I. 2001, ApJ, 548, 532
12. Lawrence, A., et al. 1985, MNRAS, 217, 685
13. Lee, T., & Petrosian, V. 1997, ApJ, 474, 37
14. Mészáros, A., & Mészáros, P. 1996, ApJ, 466, 29
15. Mitrofanov, I. G., et al. 1996, ApJ, 459, 570
16. Norris, J. P., et al. 1994, ApJ, 424, 540
17. Norris, J. P., et al. 1996, ApJ, 459, 393
18. Ostriker, J. P. & Gnedin, N. Y. 1996, ApJ, 472, L63
19. Panaitescu, A., Spada, M., & Mészáros, P. 1999, ApJ, 522, L105
20. Reichart, D. E., et al. 2001, ApJ, 552, 57
21. Stern, B., Poutanen, J., & Svensson, R. 1999, ApJ, 510, 312
22. Yi, I., & Mao, S. 1994, Phys. Rev. Lett., 72, 3750

HYPERNOVAE, BLACK HOLES, AND GALACTIC CHEMICAL EVOLUTION

Keiichi MAEDA, Ken'ichi NOMOTO, Hideyuki UMEDA, Takayoshi NAKAMURA
Department of Astronomy and Research Center for the Early Universe, School of Science, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, JAPAN

Paolo A. MAZZALI
Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34131 Trieste, Italy, and Research Center for the Early Universe, School of Science, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, JAPAN

We discuss the characteristics of nucleosynthesis in 'Hypernovae', i.e., supernovae with very large explosion energies ($\gtrsim 10^{52}$ ergs) both in 1D (spherical explosions) and 2D (aspherical explosions). The hypernova yields compared to those of ordinary core-collapse supernovae show the following characteristics: 1) Complete Si-burning takes place in more extended regions, so that the mass ratio between the complete and incomplete Si burning regions is generally larger in hypernovae than normal supernovae. 2) Si-burning takes place in lower density regions, so that the effects of α-rich freezeout is enhanced. Thus ^{44}Ca, ^{48}Ti, and ^{64}Zn are produced more abundantly than in normal supernovae. 3) Oxygen burning also takes place in more extended regions for the larger explosion energy. Then a larger amount of Si, S, Ar, and Ca ("Si") are synthesized, which makes the "Si"/O ratio larger. Aspherical explosions lead to a strong α-rich freezeout, thus enhancing the abundance ratios [^{44}Ca, ^{48}Ti, and ^{64}Zn / Fe] in the ejecta as well as the aspherical element distribution in the ejecta. We examine the nature of hypernovae by comparing the above characteristic with observations. We suggest that hypernova explosions are intrinsically aspherical events (which comes from the modeling of the late time spectra of SN1998bw), that such explosions are related with black hole formations (from the modeling of the black hole binary Nova Sco), and that hypernovae have made important contribution to the early Galactic (and Cosmic) chemical evolution.

1 Introduction

Type II and Ib/c supernovae (SNe II and SNe Ib/c) are catastrophic explosions which are induced when massive stars (8 to $\sim 100\,M_\odot$) reach the core-collapse stage finally at the end of their evolution. Until recently, we have usually considered supernovae with the kinetic energies of explosions $E_{51} = 1$ (where $E_{51} = E/10^{51}$ergs). These values have been estimated from supernovae 1987A, 1993J, and 1994 I, whose progenitors' masses have been modeled to be 13 - 20 $M_\odot$ (e.g., Nomoto et al. 1993, 1994; Blinnikov et al. 2000).

SN1998bw called into question if the above energy estimate applies to all core-collapse supernovae. The exceptionally bright Type Ic supernova (SN Ic)

SN 1998bw was discovered as the probable optical counterpart of the gamma-ray burst GRB980425 (Galama et al. 1998). The very broad spectral features and the light curve shape in the early phase have led to the conclusion that SN1998bw was a hyper-energetic explosion (kinetic energy $E_{51} \sim 40$) of a massive C+O star (Iwamoto et al. 1998; Woosley et al. 1999; Branch 2001; Nakamura et al. 2001a).

In this paper the term "hypernova" is used to refer to a SN explosion with $E_{51} \gtrsim 10$, regardless of the nature of the central engine (Nomoto et al. 2001a). Recently, other hypernova candidates have been recognized. SN1997ef and SN1998ey are also classified as SNe Ic, and show very broad spectral features similar to SN1998bw (Iwamoto et al. 2000). The spectra and the light curve of SN1997ef have been well simulated by the explosion of a $10M_\odot$ C+O star (which corresponds to $35M_\odot$ main sequence star) with $E_{51} = 10 \pm 2$ and $M(^{56}\mathrm{Ni}) \sim 0.15\ M_\odot$ (Iwamoto et al. 2000; Mazzali, Iwamoto, & Nomoto 2000). SN1997cy is classified as a SN IIn and unusually bright (Germany et al. 2000; Turatto et al. 2000). Its light curve has been simulated by a circumstellar interaction model which requires $E_{51} \sim 50$ (Turatto et al. 2000). The spectral similarity of SN1999E to SN1997cy (Cappellaro et al. 1999) would suggest that SN1999E is also a hypernova. Note that all of these estimates of E assume a spherically symmetric event. We review the characteristics of nucleosynthesis in hypernovae in section 2.

Though all of above modelings assume spherically symmetry, many suggestions are made on the explosion mechanism of very massive stars (e.g., $\geq 25M_\odot$) where asphericity plays a central role (e.g., Woosley 1993; MacFadyen & Woosley 1999; Pachezski 1998). we examine the effect of aspherical explosions on nucleosynthesis in hypernovae in section 3. We then investigate the degree of asphericity in the ejecta of SN 1998bw.

So far, all the modelings of hypernovae (e.g,, 1998bw, 1997ef, and 1997cy) imply that their progenitors' masses are $M \geq 25M_\odot$. Such very massive stars are likely to form black holes, while less massive stars form neutron stars (see, however, Wheeler et al. 2000). We consider this point in section 4, where we show that a black hole in a binary GRO J1655-40 (Nova Sco) is a relic of a supernova (most likely hypernova) explosion of very massive star ($\sim 40M_\odot$), which left a black hole as a compact remnant.

Finally contributions of hypernovae to chemical evolutions of galaxies are discussed in section 5, and summary is given in section 6.

2 Nucleosynthesis in Hypernova Explosions; Spherical Models

In core-collapse supernovae/hypernovae, stellar material undergoes shock heating and subsequent explosive nucleosynthesis. Iron-peak elements are produced in two distinct regions, which are characterized by the peak temperature, $T_{\rm peak}$, of the shocked material. For $T_{\rm peak} > 5 \times 10^9$K, material undergoes complete Si burning whose products include Co, Zn, V, and some Cr after radioactive decays. For 4×10^9K $< T_{\rm peak} < 5 \times 10^9$K, incomplete Si burning takes place and its after decay products include Cr and Mn (e.g., Hashimoto et al. 1989; Woosley, & Weaver 1995; Thielemann et al. 1996).

Figure 1 shows nucleosynthesis in hypernovae and normal supernovae for $E_{51} = 30$ and 1 (Nakamura et al. 2001b). We note the following characteristics of nucleosynthesis with very large explosion energies (Nomoto et al. 2001a,b).

1) Both complete and incomplete Si-burning regions shift outward in mass compared with normal supernovae, so that the mass ratio between the complete and incomplete Si-burning regions becomes larger. As a result, higher energy explosions tend to produce larger [(Zn, Co)/Fe], smaller [(Mn, Cr)/Fe], and larger [Fe/O]. The elements synthesized in this region such as ^{56}Ni, ^{59}Cu, ^{63}Zn, and ^{64}Ge (which decay into ^{56}Co, ^{59}Co, ^{63}Cu, and ^{64}Zn, respectively) are ejected more abundantly than in normal supernovae.

2) In the complete Si-burning region of hypernovae, elements produced by α-rich freezeout are enhanced because nucleosynthesis proceeds at lower densities (i.e., higher entropy) and thus a larger amount of ^{4}He is left. Hence, elements synthesized through capturing of α-particles, such as ^{44}Ti, ^{48}Cr, and ^{64}Ge (decaying into ^{44}Ca, ^{48}Ti, and ^{64}Zn, respectively) are more abundant.

3) Oxygen burning takes place in more extended, lower density regions for the larger explosion energy. Therefore, more O, C, Al are burned to produce a larger amount of burning products such as Si, S, and Ar. Therefore, hypernova nucleosynthesis is characterized by large abundance ratios of [Si/O], [S/O], [Ti/O], and [Ca/O].

3 Aspherical Hypernova Explosion and SN1998bw

Nakamura et al. (2001a) and Mazzali et al. (2001) have identified some signatures of asymmetric explosion in the late light curve and spectra of SN 1998bw. Maeda et al. (2002) have examined the effect of aspherical (jet-like) explosions on nucleosynthesis in hypernovae.

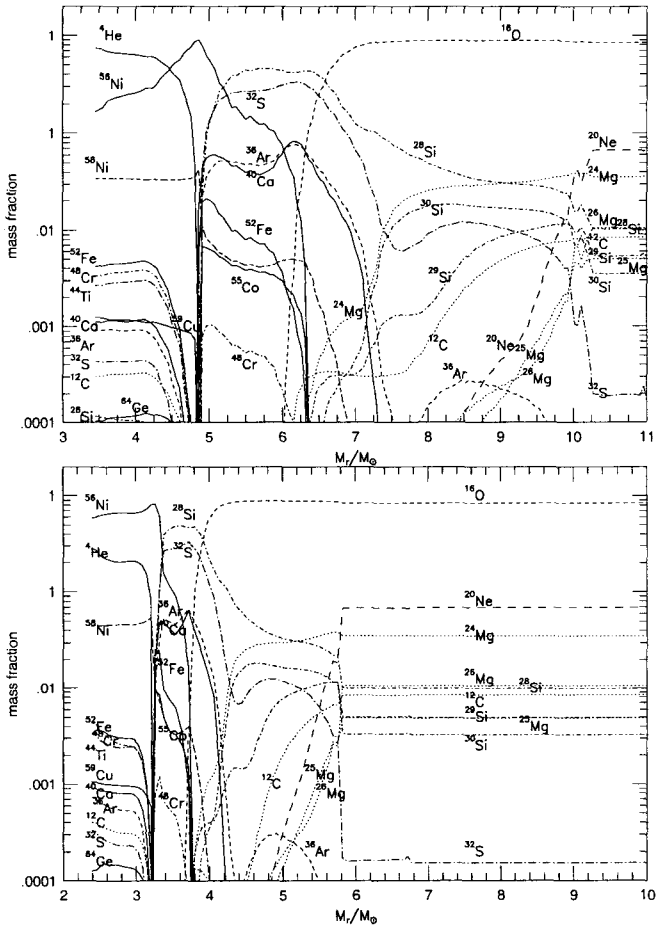


Figure 1: Nucleosynthesis in hypernovae and normal core-collapse supernovae for explosion energies of $E_{51} = 30$ (upper) and 1 (lower) ($E_{51} = E/10^{51}$ergs). The progenitor model is the $40M_{\odot}$ star model (Nakamura et al. 2001b).

3.1 Nucleosynthesis

Maeda et al. (2002) have constructed several asymmetric explosion models. The progenitor model is the 16 $M_\odot$ He core of the 40 $M_\odot$ star. The explosion energies E_{51} =1, 10, 30 are examined.

Figure 2 shows the isotopic composition of the ejecta of their model C (where $E_{51} = 10$). In the z-direction (i.e., the jet direction), where the ejecta carries more kinetic energy, the shock is stronger and post-shock temperatures are higher, so that explosive nucleosynthesis takes place in more extended, lower density regions compared with the r-direction. Therefore, larger amounts of α-rich freeze-out elements, such as ^{4}He and ^{56}Ni (which decays into ^{56}Fe via ^{56}Co) are produced in the z-direction than in the r-direction. Also, the expansion velocity of newly synthesized heavy elements is much higher in the z-direction.

In contrast, along the r-direction ^{56}Ni is produced only in the deepest layers, and elements ejected in this direction are mostly the products of hydrostatic nuclear burning (O), with some explosive oxygen-burning products (e.g., Si, S). The expansion velocities are much lower than in the z-direction.

Figure 3 (left panel) shows the 2D distribution of ^{56}Ni and ^{16}O in the homologous expansion phase. The distribution of newly synthesized elements (e.g., ^{56}Ni) is elongated in the z-direction, while that of ^{16}O is less aspherical. On the other hand, because the ejecta move more slowly in the r-direction, densities in this direction are higher than in the z-direction.

The main characteristics can be summarized as follows (see Maeda et al. (2002) and Nomoto et al. (2001b,c)).

(1) The complete Si-burning region is more extended for larger explosion energies. The aspherical explosion causes a region of higher entropy along the $z-$axis, which offers better conditions for the α-rich freezeout (Fig. 2). The high entropy inhibits the production of ^{56}Ni. Much ^{4}He is left after the freeze-out, so that the elements produced through ^{4}He capture are very abundant in the deepest region along the z-axis (Fig. 3). This results in the enhancement of the elements synthesized in the deepest region, such as ^{44}Ca (produced as ^{44}Ti), ^{48}Ti (as ^{48}Cr), and elements heavier than $A \sim 58$. Because of the enhancement of these elements and the simultaneous suppression of ^{56}Ni, the abundances of these elements relative to iron (e.g., [^{44}Ca, ^{48}Ti, ^{64}Zn /Fe]) are greatly enhanced. For more asymmetric explosion, the effect of α-rich freezeout is even larger.

(2) Incomplete Si-burning and O-burning regions are more extended for larger explosion energies (Nakamura et al. 2001b). This results in the enhancement of ^{28}Si, ^{32}S, ^{40}Ca, ^{54}Fe, and in the reduction of O. Asphericity has little

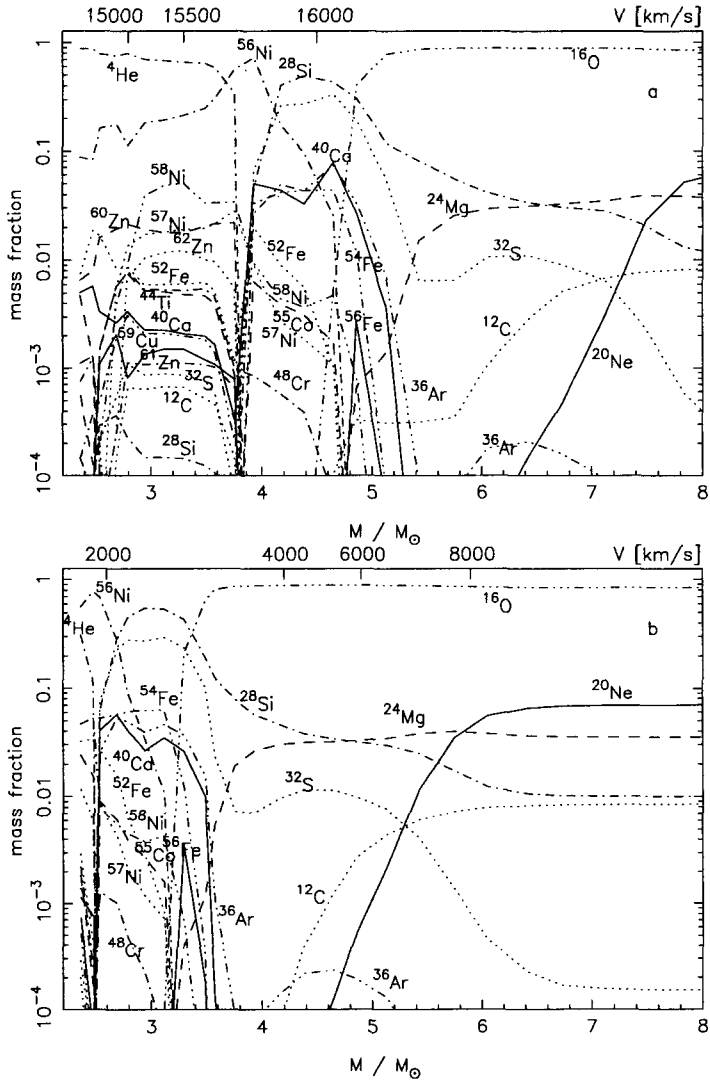


Figure 2: The isotopic composition of the ejecta of an aspherical explosion in the direction of the jet (z; upper panel) and perpendicular to the jet (r; lower panel). The ordinate indicates the initial spherical Lagrangian coordinate (M_r) of the test particles (lower scale), and the final expansion velocities (V) of those particles (upper scale).

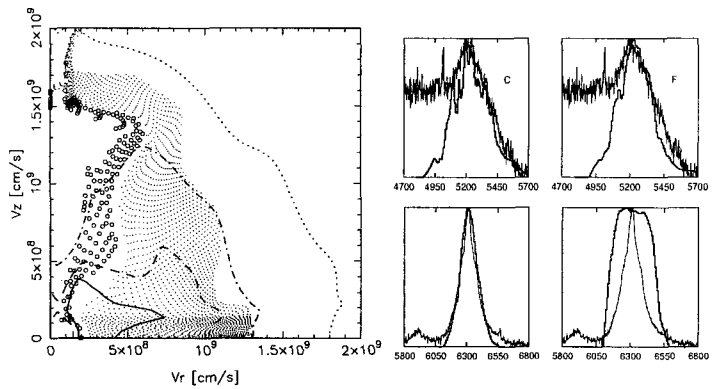


Figure 3: Left: The 2D distribution of ^{56}Ni (open circles) and ^{16}O (dots) in the homologous expansion phase. The lines are density contours at the level of 0.5 (solid), 0.3 (dashed), 0.1 (dash-dotted), and 0.01 (dotted) of the max density, respectively. Right: The profiles of the [Fe II] feature (upper panels) and of [O I] 6300, 6363 Å (lower panels) viewed at 15° from the jet direction (left panels; thick lines) and for a spherical model (right panels).

effect on the production of these elements.

The most pronounced effect of asphericity is that elements produced by the strong α-rich freezeout are greatly enhanced relative to iron (e.g., [Ti/Fe]). For other explosive burning products, the effect of a large explosion energy usually dominates over that of asphericity.

3.2 The Late Time Spectra of SN 1998bw

Mazzali et al. (2001) calculated synthetic nebular-phase spectra of SN 1998bw using a spherically symmetric NLTE nebular code based on the deposition of γ-rays from ^{56}Co decay in a nebula of uniform density and composition. They showed that the [O I] and the [Fe II] features can only be reproduced if different velocities are assumed for the two elements. A significant amount of slowly-moving O is necessary to explain the zero velocity peak of the [O I] line, and fast-moving Fe (faster then O) is needed to fit the broad-line Fe spectrum. This suggests that an aspherical distribution of Fe and O may explain the observations.

In order to verify the observable consequences of an axisymmetric explosion, we calculated the profiles of the [Fe II] blend and of [O I]. The [Fe II] and [O I] profiles viewed at an angle of 15° from the jet direction and those for a spherical model are compared to the observed spectrum of 26 Nov 1998 in Figure 3 (right panel). In a spherical explosion (model F) oxygen is located at

higher velocities than iron, and the [O I] line is too broad. This is because of the deficiency of oxygen with small velocity along the line of sight. Therefore the observed line profiles are not explained with a spherical model.

In aspherical explosion models Fe is distributed preferentially along the jet direction, and so a larger ratio of the Fe and O line widths can be obtained. When the degree of asphericity is high and the explosion is viewed from near the jet direction, the component lines in the [Fe II] blend have double-peaked profiles, the blue- and red-shifted peaks corresponding to matter situated in the two opposite lobes of the jet, where Fe is mostly produced. Because of the high velocity of Fe, the peaks are widely separated, and the blend is wide (Fig. 3, model *C*). In contrast, the [O I] line is narrower and has a sharper peak, because O is produced mostly in the r-direction, at lower velocities and with a less aspherical distribution.

4 Black hole binary GRO J1655-40 (X-ray Nova Sco)

X-ray Nova Sco (GRO J1655-40), which consists of a stellar mass black hole and a low mass companion (e.g., Brandt et al. 1995; Nelemans et al. 2000), also exhibits what could be the nucleosynthesis signature of a hypernova explosion. The companion star is enriched with Ti, S, Si, Mg, and O but not much Fe (Israelian et al. 1999). This is compatible with heavy element ejection from a black hole progenitor. In order to eject large amount of Ti, S, and Si and to have at least $\sim 4\ M_\odot$ below mass cut and thus form a massive black hole, the explosion would need to be highly energetic (Figure 1; Israelian et al. 1999; Brown et al. 2000; Podsiadlowski et al. 2002). A hypernova explosion with the mass cut at large M_r ejects a relatively small mass Fe and would be consistent with these observed abundance features.

Podsiadlowski et al. (2002) modeled the pollution of the secondary. Their model handles the binary evolution, which follows the changes in binary parameters (i.e., the primary's mass, the secondary's mass, and the orbital separation) in different evolutional stages (i.e., the pre-supernova mass lose, the supernova explosion, and the post-supernova mass transfer) as well as the pollution of the secondary by the supernova explosion using several supernova models. The progenitor models are helium stars of 6 to 16 $M_\odot$ (which correspond to main sequence stars of 20 to 40 $M_\odot$) and the explosion energies range from E_{51} =1 to 30.

The main results are summarized as follows (see Podsiadlowski et al. 2002 for details),

1) Helium star models of 10 and 16 $M_\odot$ (whose main sequence masses are 30 and 40 $M_\odot$, respectively) are most probable.

2) Hypernova models are preferred over standard supernova models, which tend not to produce enough S and Si. These results 1) and 2) strengthen the connection among a massive star, a supernova which leaves a black hole as a compact remnant, and a hypernova.

3) The overabundance of Ti in the secondary star is difficult to explain in current models (at least by spherical models). Aspherical explosion models can reproduce the observed abundance of Ti when we assume that there is complete lateral mixing between the ejecta in the equatorial plane and in the polar direction. If such extreme mixing, however, can be realized or not needs further numerical studies.

5 Contribution of Hypernovae to Chemical Evolutions of Galaxies

5.1 Starburst Galaxy M82

X-ray emissions from the starburst galaxy M82 were observed with ASCA and the abundances of several heavy elements were obtained (Tsuru et al. 1997). Tsuru et al. (1997) found that the overall metalicity of M82 is quite low, i.e., O/H and Fe/H are only 0.06 - 0.05 times solar, while Si/H and S/H are $\sim$ 0.40 - 0.47 times solar. This implies that the abundance ratios are peculiar, i.e., the ratio O/Fe is about solar, while the ratios of Si and S relative to O and Fe are as high as $\sim$ 6 - 8. These ratios are very different from those ratios in SNe II. Compared with normal SNe II, the important characteristic of hypernova nucleosynthesis is the large Si/O, S/O, and Fe/O ratios. Figure 4 shows the good agreement between the hypernova model ($E_{51} = 30$) and the observed abundances in M82.

Hypernovae could also produce larger E per oxygen mass than normal SNe II, as required for M82. We therefore suggest that hypernova explosions may make important contributions to the metal enrichment and energy input to the interstellar matter in M82. The age of starburst activity is estimated to be $\lesssim 10^7$ years (Stickland 2002), which is so young that only massive stars ($M > 25\ M_\odot$) contributed to nucleosynthesis in M82.

5.2 Early Galactic Chemical Evolution

The abundance pattern of metal-poor stars with [Fe/H] < -2 provides us with very important information on the formation, evolution, and explosions of massive stars in the early evolution of the galaxy.

In the early galactic epoch when the galaxy is not yet chemically well-mixed, [Fe/H] may well be determined by mostly a single SN event (Audouze & Silk 1995). The formation of metal-poor stars is supposed to be driven

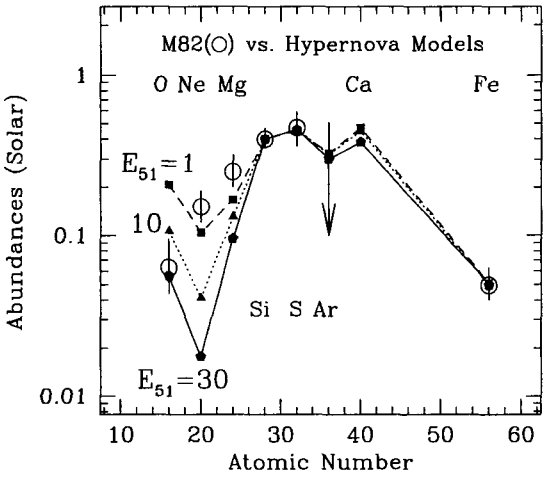


Figure 4: Abundance patterns in the ejecta of $25M_{\odot}$ metal-free SN II and hypernova models compared with abundances (relative to the solar values) of M82 observed with ASCA (Tsuru et al. 1997). Here, the open circles with error bars show the M82 data. The filled square, triangle, and pentagons represent E_{51} ($= E/10^{51}$ergs) =1, 10, and 30 models, respectively (Umeda et al. 2002).

by a supernova shock, so that [Fe/H] is determined by the ejected Fe mass and the amount of circumstellar hydrogen swept-up by the shock wave (Ryan, Norris, & Beers 1996). Then, hypernovae with larger E are likely to induce the formation of stars with smaller [Fe/H], because the mass of interstellar hydrogen swept up by a hypernova is roughly proportional to E (Ryan et al. 1996; Shigeyama & Tsujimoto 1998) and the ratio of the ejected iron mass to E is smaller for hypernovae than for canonical supernovae.

The observed abundances of metal-poor halo stars show quite interesting pattern. There are significant differences between the abundance patterns in the iron-peak elements below and above [Fe/H]~ -2.5 - -3.

1) For [Fe/H]$\lesssim -2.5$, the mean values of [Cr/Fe] and [Mn/Fe] decrease toward smaller metalicity, while [Co/Fe] increases (McWilliam et al. 1995; Ryan et al. 1996).

2) [Zn/Fe]~ 0 for [Fe/H] $\simeq -3$ to 0 (Sneden, Gratton, & Crocker 1991), while at [Fe/H] < -3.3, [Zn/Fe] increases toward smaller metalicity (Figure 5; Primas et al. 2000; Blake et al. 2001).

These trends cannot be explained with the conventional chemical evolution model that uses previous nucleosynthesis yields.

The larger [(Zn, Co)/Fe] and smaller [(Mn, Cr)/Fe] in the supernova ejecta

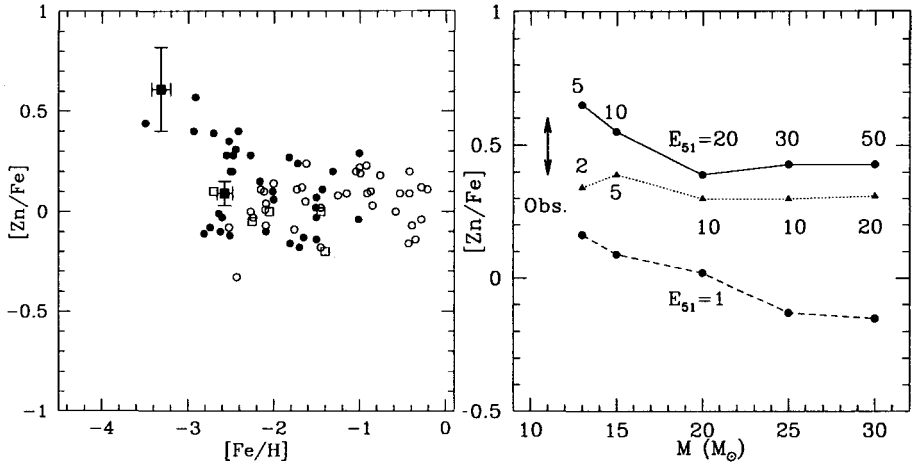


Figure 5: Left: Observed abundance ratios of [Zn/Fe]. These data are taken from Primas et al. (2000) (filled circles), Blake et al. (2001) (filled square) and from Sneden et al. (1991) (others) (Umeda & Nomoto 2002).
Right: The maximum [Zn/Fe] ratios as a function of M and $E_{51} = E/10^{51}$ergs (Umeda & Nomoto 2002). The observed large [Zn/Fe] ratio in very low-metal stars ([Fe/H] < -2.6) found in Primas et al. (2000) and Blake et al. (2001) are represented by a thick arrow.

can be realized if the mass ratio between the complete Si burning region and the incomplete Si burning region is larger, or equivalently if deep material from complete Si-burning region is ejected by mixing or aspherical effects. This can be realized if (1) the mass cut between the ejecta and the collapsed star is located at smaller M_r (Nakamura et al. 1999), (2) E is larger to move the outer edge of the complete Si burning region to larger M_r (Nakamura et al. 2001b), or (3) asphericity in the explosion is larger.

Also a large explosion energy E results in the enhancement of the local mass fractions of Zn and Co, while Cr and Mn are not enhanced (Umeda & Nomoto 2002). Therefore, if hypernovae made significant contributions to the early Galactic chemical evolution, it could explain the large Zn and Co abundances and the small Mn and Cr abundances observed in very metal-poor stars.

The dependence of [Zn/Fe] on M and E for spherical models is summarized in Figure 5. Models with $E_{51} = 1$ do not produce sufficiently large [Zn/Fe]. To be compatible with the observations of [Zn/Fe] ~ 0.5, the explosion energy must be much larger, i.e., $E_{51} \gtrsim 20$ for $M \gtrsim 20M_\odot$, i.e., hypernova-like explosions of massive stars ($M \gtrsim 25M_\odot$) with $E_{51} > 10$ are responsible for the

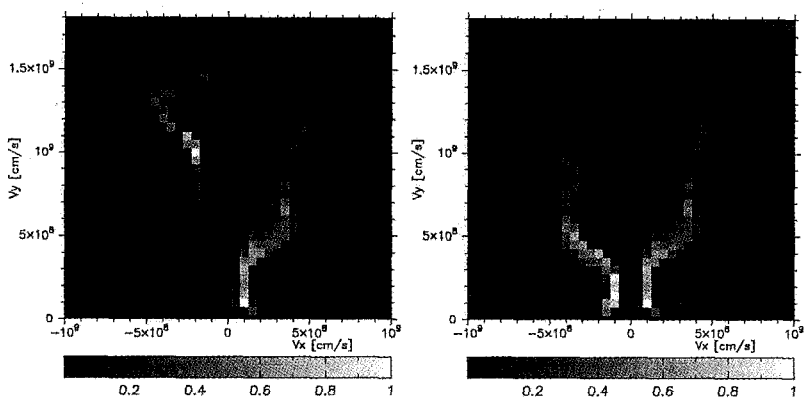


Figure 6: The density distributions of Zn (left panel: left half) and Fe (right half) and that of Mn (right panel: left half) and Fe (right half). The densities of each element are represented with linear scale, from 0 to the max density of each element.

production of Zn.

Iron-peak elements are produced in the deep region near the mass cut, so that their production is strongly affected by asphericity of the explosion. Figure 6 shows the 2D density distribution of ^{64}Ge (which decays into ^{64}Zn) and that of ^{55}Co (which decays into ^{55}Mn). ^{64}Zn, which is produced by the strong α-rich freezeout, is distributed preferentially in the z-direction. Moreover, production of ^{64}Zn is strongly enhanced compared with a spherical model, since the post-shock temperatures along the z-direction are much higher than that in spherical models. On the other hand, ^{55}Mn, which is produced by incomplete silicon burning, surrounds ^{56}Fe and located preferentially in the r-direction. Accordingly, materials with higher [Zn/Fe] and lower [Mn/Fe] are ejected along the Z-direction with higher velocities, and materials with lower [Zn/Fe] and higher [Mn/Fe] are ejected toward the r-direction with lower velocities, which would be easier to fall back to a central remnant.

In this way, larger asphericity in the explosion leads to larger [Zn/Fe] and [Co/Fe], but to smaller [Mn/Fe] and [Cr/Fe]. Then, if the degree of the asphericity tends to be larger for lower [Fe/H], the trends of [Zn, Co, Mn, Cr/Fe] follow the observed ones.

It is likely that in the real situations, the effects of asphericity, mixing, and the position of the mass-cut (Nakamura et al. 1999) work to make observed trends. Then, the restrictions encountered by each of them (e.g., requirement of large Y_e, large amounts of mixing) may be relaxed.

6 Summary

We investigated explosive nucleosynthesis in hypernovae. The characteristics of hypernova yields compared to those of ordinary core-collapse supernovae are summarized as follows:

1) Complete Si-burning takes place in more extended region, so that the mass ratio between the complete and incomplete Si burning regions is generally larger in hypernovae than normal SNe II. As a result, higher energy explosions tend to produce larger [(Zn, Co)/Fe], smaller [(Mn, Cr)/Fe], and larger [Fe/O]. If hypernovae made significant contributions to the early Galactic chemical evolution, it could explain the large Zn and Co abundances and the small Mn and Cr abundances observed in very metal-poor stars.

2) In hypernovae, Si-burning takes place in lower density regions, so that the effects of α-rich freezeout is enhanced. Thus ^{44}Ca, ^{48}Ti, and ^{64}Zn are produced more abundantly than in normal supernovae. The large [(Ti, Zn)/Fe] ratios observed in very metal poor stars strongly suggest a significant contribution of hypernovae.

3) Oxygen burning also takes place in more extended regions for the larger explosion energy. Then a larger amount of Si, S, Ar, and Ca ("Si") are synthesized, which makes the Si/O ratio larger. The abundance pattern of the starburst galaxy M82, characterized by abundant Si and S relative to O and Fe, may be attributed to hypernova explosions if the IMF is relatively flat, and thus the contribution of massive stars to the galactic chemical evolution is large.

We also examined aspherical explosion models. The explosive shock is stronger in the jet-direction. The high explosion energies concentrated along the jet result in hypernova-like nucleosynthesis. On the other hand, the ejecta orthogonal to the jet is more like ordinary supernovae. Therefore, asphericity in the explosions strengthens the nucleosynthesis properties of hypernovae as summarized above in 1) and 2) but not 3).

Then we investigated the nature of hypernova explosions and suggested 1) aspherical explosion models for hypernovae, and 2) the connection among hypernovae, stellar mass black holes, and massive progenitors, through the modelings of individual objects (SN1998bw and Nova Sco).

Properties of hypernova nucleosynthesis suggest that hypernovae of massive stars may make important contributions to the Galactic (and cosmic) chemical evolution, especially in the early low metalicity phase. This may be consistent with the suggestion that the IMF of Pop III stars is different from that of Pop I and II stars, and that more massive stars are abundant for Pop III (e.g., Nakamura & Umemura 1999; Omukai & Nishi 1999; Bromm, Coppi

& Larson 1999).

Acknowledgments

We would like to thank N. Patat, Ph. Podsiadlowski, B. Schmidt, S. Ryan, C. Kobayashi, M. Shirouzu, F. Thielemann, I. Hachisu for useful discussion. KM thanks W. Hillebrandt for his support during the stay at MPA in November, 2001, when most part of the manuscript was written. This work has been supported in part by the grant-in-Aid for Scientific Research (07CE2002, 12640233) of the Ministry of Education, Science, Culture, and Sports in Japan.

References

1. Audouze, J., & Silk, J. 1995, ApJ, 451, L49
2. Blake, L.A.J., Ryan, S.G., Norris, J.E., & Beers, T.C. 2001, Nucl.Phys.A.
3. Blinnikov, S., Lundqvist, P, Bartunov, O., Nomoto, K., & Iwamoto, K. 2000, ApJ, 532, 1132
4. Branch, D. 2001, in "Supernovae and Gamma Ray Bursts," eds. M. Livio, et al. (Cambridge: Cambridge University Press), 96
5. Brandt, W.N., Podsiadlowski, P., Sigurdssen, S. 1995, MNRAS, 277, L35
6. Bromm, V., Coppi, P.S., & Larson, R. B. 1999, ApJ, 527, L5
7. Brown, G.E., Lee, C.-H., Wijers, R.A.M.J., Lee, H.K., Israelian, G., & Bethe, H.A. 2000, New Astronomy, 5, 191
8. Cappellaro, E., Turatto, M., & Mazzali, P. 1999, IAU Circ. 7091
9. Galama, T.J., et al. 1998, Nature, 395, 670
10. Germany, L.M., Reiss, D.J., Schmidt, B.P., Stubbs, C.W., & Sadler, E.M. 2000, ApJ, 533, 320
11. Hashimoto, M., Nomoto, K., & Shigeyama, T. 1989, A&A, 210, 5
12. Israelian, G., Rebolo, R., Basri, G., Casares, J., & Martin E.L. 1999, Nature, 401, 142
13. Iwamoto, K., Mazzali, P., Nomoto, K., et al. 1998, Nature, 395, 672
14. Iwamoto, K., Nakamura, T., Nomoto, K., et al. 2000, ApJ, 534, 660
15. MacFadyen, A.I. & Woosley, S.E. 1999, ApJ 524, 262
16. McWilliam, A., Preston, G.W., Sneden, C., & Searle, L. 1995, AJ, 109, 2757
17. Maeda, K., Nakamura, T., Nomoto, K., Mazzali, P.A., & Hachisu, I. 2002, ApJ, 565, in press
18. Mazzali, P.A., Iwamoto, K., & Nomoto, K. 2000, ApJ, 545, 407
19. Mazzali, P.A., Nomoto, K, Patat, F., & Maeda, K. 2001, ApJ, 559, 1047
20. Nakamura, F., & Umemura, M. 1999, ApJ, 515, 239

21. Nakamura, T., Mazzali, P. A., Nomoto, K., & Iwamoto, K. 2001a, ApJ, 550, 991
22. Nakamura, T., Umeda, H., Iwamoto, K., Nomoto, K., Hashimoto, M., Hix, R.W., & Thielemann, F.-K. 2001b, ApJ, 555, 880
23. Nakamura, F., & Umemura, M. 1999, ApJ, 515, 239
24. Nelemans, G., Tauris, T.M., van den Heuvel, E.P.J. 2000, A&A, 352, 87
25. Nomoto, K., Suzuki, T., Shigeyama, T., Kumagai, S., Yamaoka, H., Saio, H. 1993, Nature, 364, 507
26. Nomoto, K., Yamaoka, H., Pols, O. R., van den Heuvel, E. P. J., Iwamoto, K., Kumagai, S., & Shigeyama, T. 1994, Nature, 371, 227
27. Nomoto, K., et al. 2001a, in "Supernovae and Gamma Ray Bursts," eds. M. Livio, et al. (Cambridge: Cambridge University Press), 144 (astro-ph/0003077)
28. Nomoto, K., et al. 2001b, in " The influence of Binaries on Stellar Population Studies," ed. D. Vanbeveren (Kluwer), 507, (astro-ph/0105127)
29. Nomoto, K., Maeda, K., Umeda, H. 2001c, in "Chemical Enrichment of Intracluster and Intergalactic Medium," ASP Conference Series, in press (astro-ph/0110528)
30. Omukai, K., & Nishi, R. 1999, ApJ, 518, 64
31. Paczynski, B. 1998, ApJ, 494, L45
32. Patat, F., et al. 2001, ApJ, 555, 900
33. Podsiadlowski, Ph., Nomoto, K., Maeda, K., Nakamura, T., Mazzali, P.A., & Schmidt, B. 2002, ApJ, 567, in press
34. Primas, F., Reimers, D., Wisotzki, L., Reetz, J., Gehren, T., & Beers, T.C. 2000, in The First Stars, ed. A. Weiss, et al. (Springer), 51
35. Ryan, S.G., Norris, J.E. & Beers, T.C. 1996, ApJ, 471, 254
36. Shigeyama. T., & Tsujimoto, T. 1998, ApJ, 507, L135
37. Sneden, C., Gratton, R.G., & Crocker, D.A. 1991, A&A, 246, 354
38. Sollerman, J., Kozma, C., Fransson, C., Leibundgut, B., Lundqvist, P., Ryde, F., & Woudt, P. 2000, ApJ 537, L127
39. Stickland, D. 2002, in "Chemical Enrichiment of Intracluster and Intergalactic Medium," ASP Conference Series, in press (astro-ph/0107116)
40. Thielemann, F.-K., Nomoto, K., & Hashimoto, M. 1996, ApJ, 460, 408
41. Tsuru, T. G., Awaki, H., Koyama K., Ptak, A. 1997, PASJ, 49, 619
42. Turatto, et. al. 2000, ApJ, 534, L57
43. Umeda, H., Nomoto, K. Tsuru, T., et al. 2002, ApJ, submitted
44. Umeda, H., Nomoto, K. 2002, ApJ, 565, in press (astro-ph/0103241)
45. Wheeler, J. C., Yi, I., Höflich, P. A., & Wang, L. 2000, ApJ, 537, 810
46. Woosley, S.E., Eastman, R.G., & Schmidt, B.P. 1999, ApJ, 516, 788
47. Woosley, S.E. 1993, ApJ, 405, 273

48. Woosley, S. E., & Weaver, T. A. 1995, ApJS, 101, 181

TWO-DIMENSIONAL ACCRETION FLOW DRIVEN BY POYNTING FLUX

Hyun Kyu LEE[a]
Department of Physics, Hanyang University, Seoul 133-791, Korea
and
Asia Pacific Center for Theoretical Physics, Pohang 790-784, Korea

Using a simplified model of a two dimensional accretion disk, the accretion flow driven by the Poynting flux is formulated relativistically in the background of Kerr geometry.

1 Introduction

The accretion flow driven by the Poynting flux [1] [2] has been suggested for the accretion disks with the ordered magnetic fields. In contrast to the hydrodynamic jets in which the energy and the angular momentum are carried by the kinetic flux of the matter, Poynting flux is characterized by the outflow of the energy and the angular momentum carried predominantly by the electromagnetic field.

In the non-relativistic formulation, Blandford[2] suggested an axisymmetric and stationary solution for the Poynting outflow assuming a force-free magnetosphere surrounding an accretion disk. The poloidal field configuration for a black hole in a force-free magnetosphere has been discussed recently by Ghosh [3] in the relativistic formulation, in which the possible forms of the poloidal configurations of the magnetic field are suggested. The developments of the ordered magnetic field in the disk and the Poynting outflow from the disk have been studied by many authors [4] [5] [6]. Recently Ustyugova et al. [7] performed an axisymmetric magnetohydrodynamical simulation to show that the quasi-stationary and approximately force-free Poynting jet from the inner part of the accretion disk is possible.

Poynting flux in a system of black hole-accretion disk recently has also been studied in connection with the gamma ray bursts [8] [9] [10] [11]. One of the advantages of cosidering Poynting flux is that it carries very small baryonic component, which is essential for powering the gamma ray bursts [12]. The evolution of the system is also found to be largely depend on the Poynting outflow from the disk [13] [14]. The relativistic effects on the accreting flows

[a]e-mail : hklee@hepth.hanyang.ac.kr

Talk presented at 7th Italian-Korean Symposium on Relativistic Astrophysics, Kimhae and Ganghwa, Korea, July 23 - 28, 2001 and at 2nd KIAS Workshop on Astrophysics, Seoul, Sept. 3-8, 2001

close to a black hole has been discussed by Lasota[15], Abramowicz et al.[16] and recently by Gammie and Popham[17] in which the relativistic effects on the slim-disk with the vertical structures averaged are discussed in detail with viscous stresses. However the effect of relativity on the accreting flow dominated by Poynting flux has not been discussed well in depth so far.

In this work we present the relativistic formulation of the accretion flow driven dominantly by the Poynting flux [18]. We consider a toy model for a magnetically dominated thin accretion disk, which is assumed to be a two-dimensional disk located on the equatorial plane with a black hole at the center. To see the magnetic effect transparently it is also assumed that there is no viscous stress tensor in the disk and there is no radiative transfer from the disk. We develop a relativistic description of a two-dimensional model for a Poynting-flux-dominated thin accretion disk in the background metric of a Kerr black hole.

2 Surface charge and current on a Two-dimensional Accretion Disk

The structure of electromagnetic field with an accretion disk which lies on $\theta = \pi/2$ plane can be reproduced by assigning the surface charge and current on the plane in addition to the bulk charge and current distributions outside the disk.

Adopting the procedure used by Damour[19], we define a conserved current as

$$\mathcal{J}^{\mu} = J^{\mu}_{+}\Theta(\pi/2 - \theta) + J^{\mu}_{-}\Theta(\theta - \pi/2) + j^{\mu} \tag{1}$$

where $J^{\mu}_{\pm}$ is the bulk current densities(defined for the upper and lower space respectively) for the Maxwell equation,

$$F^{\mu\nu}_{\ \ ;\nu} = 4\pi J^{\mu}_{\pm}, \tag{2}$$

which terminates on the disk and j^{μ} is the surface current density defined on the two dimensional accretion disk. From the current conservation ,

$$0 = \mathcal{J}^{\mu}_{\ ;\mu} \tag{3}$$

and the current conservation of bulk current density

$$J^{\mu}_{\pm\ ;\mu} = 0 \tag{4}$$

we can identify

$$j^{\mu} = \frac{1}{4\pi}(F^{\theta\mu}_{+} - F^{\theta\mu}_{-})\delta(\theta - \pi/2) \tag{5}$$

which is an analogous expression as in Damour [19].

Now the charge density on the disk can be defined by

$$\tilde{\rho}_e = -j^\mu U_\mu = \alpha j^0 = \alpha \frac{[F^{\theta 0}]}{4\pi} \delta(\theta - \pi/2) \tag{6}$$

where

$$[A] \equiv A_+ - A_- \tag{7}$$

Defining a surface charge density σ_e as

$$\int \tilde{\rho}_e dV = \int \tilde{\rho}_e \frac{\rho^2 \varpi}{\sqrt{\Delta}} dr d\theta d\phi \equiv \int \sigma_e \frac{\rho}{\sqrt{\Delta}} \varpi dr d\phi \tag{8}$$

we get surface charge density given by

$$\sigma_e = \frac{\alpha}{4\pi} [F^{\theta 0}] \rho = -\frac{[E^{\hat{\theta}}]}{4\pi} \tag{9}$$

which is the Gauss' law on the surface of the accretion disk. Throughout this work, the electromagnetic field is defined by ZAMO(zero angular momentum observer) [20]

Similarly the current density is defined by

$$\tilde{j}^\mu = j^\mu - \rho_e U^\mu \tag{10}$$

Since $j^\theta = 0$ by construction we get

$$\tilde{j}^r = j^r \tag{11}$$

$$\tilde{j}^\phi = j^\phi + j^0 \beta \tag{12}$$

The rate of the electric charge carried out of the volume by the conserved current is given by

$$-(\frac{dQ}{dt})_V = \int (\alpha \tilde{j}^i)_{|i} dV = \int \alpha \tilde{j}^i d\Sigma_i \tag{13}$$

where $d\Sigma_i$ is the corresponding surface element.

Hence the surface current density $K^{\hat{i}}$ on the disk can be defined by

$$\int \alpha \tilde{j}^i d\Sigma_i \equiv \int \alpha K^{\hat{i}} \varpi d\phi. \tag{14}$$

Then from the radial part

$$\int \alpha \tilde{j}^r d\Sigma_r = \int \alpha \tilde{j}^r \sqrt{\gamma} d\theta d\phi = -\frac{1}{4\pi} \int \alpha [B^{\hat{\phi}}] \varpi d\phi \tag{15}$$

we obtain

$$K^{\hat{r}} = -\frac{1}{4\pi} [B^{\hat{\phi}}]. \tag{16}$$

Similarly the surface current density in $\phi-$ direction $K^{\hat{\phi}}$ on the disk,

$$K^{\hat{\phi}} = \frac{1}{4\pi} [B^{\hat{r}}], \tag{17}$$

can be obtained by identifying

$$\int \alpha \tilde{j}^{\phi} d\Sigma_{\phi} = \int \alpha (j^{\phi} + \beta j^0) \frac{\rho^2 \varpi}{\sqrt{\Delta}} dr \; d\theta = \int \alpha \frac{[B^{\hat{r}}]}{4\pi} \frac{\rho}{\sqrt{\Delta}} dr \tag{18}$$

It can be summarized as the Ampere's law on the accretion disk:

$$[\vec{B}] = -4\pi \vec{K} \times \hat{\theta}. \tag{19}$$

Then the derivatives with respect to θ near the accretion disk give

$$\partial_\theta E^{\hat{\theta}} = 4\pi \sigma_e \delta(\theta - \pi/2) \tag{20}$$

$$\partial_\theta B^{\hat{\phi}} = 4\pi K^{\hat{r}} \delta(\theta - \pi/2) \tag{21}$$

$$\partial_\theta B^{\hat{r}} = -4\pi K^{\hat{\phi}} \delta(\theta - \pi/2) \tag{22}$$

3 Conservation of energy - momentum tensor

The conservation of energy momentum tensor,

$$T^{\mu\nu}{}_{;\mu} = 0 \tag{23}$$

has four equations, one for each ν . For the steady and axial symmetric case, there are two Killing vectors and therefore correspondingly the energy and the angular momentum(azimuthal component) of the system are conserved. The combinations of equations for $\nu = 0$ and ϕ are responsible to the energy and the angular momentum conservation. The radial component of the equations, $\nu = r$, determines the orbital motion of the fluid.

3.1 Energy conservation

Using the Killing vector in t – *direction*,

$$\xi^{\mu} = (1, 0, 0, 0), \tag{24}$$

we can define the energy flux $\mathcal{E}^{\mu}$ from the energy momentum tensor, $T^{\mu\nu}$

$$\mathcal{E}^{\mu} = -T^{\mu\nu}\xi_{\nu}. \tag{25}$$

And the energy conservation can be written as

$$\mathcal{E}^{\mu}{}_{;\mu} = 0 \tag{26}$$

Energy flux is decomposed into two contributions from matter($\mathcal{E}^{\mu}_{m}$)and electromagnetic ($\mathcal{E}^{\mu}_{EM}$) parts:

$$\mathcal{E}^{\mu} = \mathcal{E}^{\mu}_{m} + \mathcal{E}^{\mu}_{EM} \tag{27}$$

Since we are interested in steady and axisymmetric case, the derivatives with respect to t and ϕ vanishes and hence $\mathcal{E}^{r}$ and $\mathcal{E}^{\mu}$ are only considered:

$$\mathcal{E}^{\mu}{}_{;\mu} = \frac{1}{\sqrt{-g}}\partial_{\mu}(\sqrt{-g}\mathcal{E}^{\mu}) = \frac{1}{\sqrt{-g}}[\partial_{r}(\sqrt{-g}\mathcal{E}^{r}) + \partial_{\theta}(\sqrt{-g}\mathcal{E}^{\theta}) \tag{28}$$

where

$$\mathcal{E}^{r}_{m} = -\rho_{m}u_{0}u^{r}, \qquad \mathcal{E}^{\theta}_{m} = 0 \tag{29}$$

$$\mathcal{E}^{r}_{EM} = -\frac{1}{4\pi}\frac{\sqrt{\Delta}}{\rho}(\alpha E^{\hat{\theta}} + \beta\varpi B^{\hat{r}})B^{\hat{\phi}}, \quad \mathcal{E}^{\theta}_{EM} = -\frac{1}{4\pi\rho}(-\alpha E^{\hat{r}} + \beta\varpi B^{\hat{\theta}})B^{\hat{\phi}} \tag{30}$$

Since we assume the idealized thin disk, we do not expect any substantial radial flow of electromagnetic energy and angular momentum unless there is any singular structure of electromagnetic field on the disk. Compared to the matter part which has delta function singularity,

$$\mathcal{E}^{r}_{m} = -\rho_{m}u_{0}u^{r} = -\frac{\sigma_{m}}{r}u_{0}u^{r}\delta(\theta - \pi/2) \tag{31}$$

the radial flow $\partial_{r}(\sqrt{-g}\mathcal{E}^{r}_{EM})$ has no delta-type singularity so that radial flow can be ignored in this thin disk approximation.

On the other hand the electromagnetic energy flow in the θ direction has singularity on the disk due to the discontinuity of the electromagnetic fields. Ignoring nonsingular parts, we obtain

$$\partial_{\theta}(\sqrt{-g}\mathcal{E}^{\theta}_{EM}) = \cdots + rK^{r}(-\alpha E^{\hat{r}} + \beta\varpi B^{\hat{\theta}}) \tag{32}$$

Then the energy conservation leads

$$-\partial_r(u_0)r\sigma_m u^r + rK^r(-\alpha E^{\hat{r}} + \beta\varpi B^{\hat{\theta}}) = 0 \tag{33}$$

or

$$\partial_r(u_0)\dot{M}_+ + 2\pi rK^r(-\alpha E^{\hat{r}} + \beta\varpi B^{\hat{\theta}}) = 0 \tag{34}$$

It can be derived from the condition that there is no accumulation of energy in the steady accretion due to magnetic braking[18]. In the idealized thin disk considered in this work, there is no radiative transfer or viscous interaction which balance the radial flow of the energy and angular momentum of the accreting matter. Hence the electromagnetic field anchored on the disk is responsible for the conservation of the total energy and angular momentum.

Let us consider a circular strip at r with width dr. The energy flux of electromagnetic field out of the disk is given by

$$\Delta E_{EM} = -2\int_r^{r+dr} \alpha\mathcal{E}^{\theta}_{+,EM}\sqrt{\gamma}d\Sigma_\theta \tag{35}$$

$$= -2\int_r^{r+dr} \alpha\frac{1}{4\pi\rho}(-\alpha E^{\hat{r}} + \beta\varpi B^{\hat{\theta}})2\pi K^{\hat{r}}\sqrt{\gamma}drd\phi \tag{36}$$

where factor 2 is introduced to take account of the other side of the disk. The energy flux out of the disk is the energy carried into the strip by the accretion flow given by

$$\Delta E_m = -(-u_0)\dot{M}_+|_r^{r+dr} = \partial_r u_0\dot{M}_+ dr. \tag{37}$$

The energy conservation on this strip,

$$\Delta E_{EM} = \Delta E_m, \tag{38}$$

can be written

$$\partial_r u_0\dot{M}_+ + 2\pi rK^{\hat{r}}(-\alpha E^{\hat{r}} + \beta\varpi B^{\hat{\theta}}) = 0. \tag{39}$$

which is identical to eq.(33)

3.2 Angular momentum conservation

Using the Killing vector in $\phi - direction$ for the axial symmetric case,

$$\eta^\mu = (0,\ 0,\ 0,\ 1) \tag{40}$$

we can define the angular momentum flux $\mathcal{L}^{\mu}_{(\phi)}$ from the energy momentum tensor, $T^{\mu\nu}$,

$$\mathcal{L}^{\mu}_{(\phi)} = -T^{\mu\nu}\eta_{\nu} \tag{41}$$

where

$$\mathcal{L}^{\mu}_{(\phi)} = \mathcal{L}^{\mu}_{m(\phi)} + \mathcal{L}^{\mu}_{EM(\phi)}. \tag{42}$$

Since we are interested in a steady and axisymmetric case, the derivatives with respect to t and ϕ vanishes and hence $\mathcal{L}^{r}$ and $\mathcal{L}^{\mu}$ are only considered:

$$\mathcal{L}^{\mu}_{(\phi);\mu} = \frac{1}{\sqrt{-g}}\partial_{\mu}(\sqrt{-g}\mathcal{L}^{\mu}_{(\phi)}) = \frac{1}{\sqrt{-g}}[\partial_{r}(\sqrt{-g}\mathcal{L}^{r}_{(\phi)}) + \partial_{\theta}(\sqrt{-g}\mathcal{L}^{\theta}_{(\phi)})] \tag{43}$$

where

$$\mathcal{L}^{\mu}_{m(\phi)} = \rho_{m}u_{\phi}u^{\mu}, \quad \mathcal{L}^{r}_{EM(\phi)} = -\frac{1}{4\pi}\frac{\varpi\sqrt{\Delta}}{\rho}B^{\hat{\phi}}B^{\hat{r}} \tag{44}$$

and the flux into θ – *direction* (normal to the disk on the equatorial plane), $\mathcal{L}^{\theta}_{EM(\phi)}$,

$$\mathcal{L}^{\theta}_{EM(\phi)} = -\frac{1}{4\pi}\frac{\varpi}{\rho}B^{\hat{\phi}}B^{\theta} \tag{45}$$

It is the torque exerted on the surface current density K^{r}. Using eq.(45), The Poynting flux out of the disk, Eq.(30), can be written as

$$\mathcal{E}^{\theta}_{EM} = \frac{1}{4\pi\rho}[\alpha\hat{E}\times\hat{B}|_{\theta} - \beta\mathcal{L}^{\theta}_{EM(\phi)}] \tag{46}$$

The product of $B^{\hat{\phi}}B^{\hat{r}}$ has no discontinuity and can be ignored compared to the matter part and

$$\partial_{\theta}\mathcal{L}^{\theta}_{EM(\phi)} = -\frac{1}{4\pi}\frac{\varpi}{\rho}B^{\theta}\partial_{\theta}B^{\hat{\phi}} = -\frac{\varpi}{r}B^{\theta}K^{\hat{r}} \tag{47}$$

Then the angular momentum conservation can be written as

$$\partial_{r}(u_{\phi})r\sigma_{m}u^{r} - rK^{r}\varpi B^{\hat{\theta}} = 0 \tag{48}$$

or

$$\partial_{r}(u_{\phi})\dot{M}_{+} + 2\pi rK^{r}\varpi B^{\hat{\theta}} = 0 \tag{49}$$

As observed in the previous section, the angular momentum carried out by Poynting flux is the angular momentum accumulated otherwise by the accreting flow on the disk.

Using eq.(45) the angular momentum flux of electromagnetic field into the θ-direction is given by

$$\Delta L_{EM} = 2\int_{r}^{r+dr} \alpha \mathcal{L}^{\theta}_{EM}\sqrt{\gamma}d\Sigma_{\theta} \tag{50}$$

$$= -\int B^{\hat{\theta}}B^{\hat{\phi}}\rho\varpi dr \tag{51}$$

where factor 2 is introduced to take account of the other side of the disk. For an infinitesimally small dr,

$$\Delta L_{EM} = -B^{\hat{\theta}}B^{\hat{\phi}}\rho\varpi dr = 2\pi K^{r}B^{\hat{\theta}}r\varpi dr \tag{52}$$

It is balanced by the change of the angular momentum of matter in the strip given by

$$\Delta L_{m} = -u_{\phi}\dot{M}_{+}|_{r}^{r+dr} = -\frac{du_{\phi}}{dr}\dot{M}_{+}dr. \tag{53}$$

From the angular momentum conservation,

$$\Delta L_{m} = \Delta L_{EM}, \tag{54}$$

we get

$$-\frac{du_{\phi}}{dr}\dot{M}_{+} = 2\pi K^{r}B^{\hat{\theta}}r\varpi, \tag{55}$$

which is identical to eq.(49). And we get

$$\dot{M}_{+} = \frac{B^{\hat{\theta}}B^{\hat{\phi}}\rho\varpi}{du_{\phi}/dr}. \tag{56}$$

In the non-relativistic limit u_{ϕ} can be identified as non-relativistic Keplerian angular momentum: $\Omega_{K} = \sqrt{M/r^{3}}$ and we get

$$\dot{M}_{+} \to \frac{B^{\hat{\theta}}B^{\hat{\phi}}r^{2}}{r\Omega_{K}/2} = 2r\frac{B^{\hat{\theta}}B^{\hat{\phi}}}{\Omega_{K}} \tag{57}$$

which is identical to those in Lee and Kim [13] and Ustyugova et al. [7].

3.3 Radial equation

The radial component of the energy-momentum conservation

$$T^{\mu r}{}_{,\mu} = 0 \tag{58}$$

determines the angular velocity of the fluid.

For example the Keplerian angular velocity can be obtained from the matter part only,

$$T_m^{\mu r}{}_{;\mu} = \rho_m(u^r \partial_r u^r + \Gamma^r_{\mu\nu} u\mu u^\nu) \tag{59}$$

where

$$\Gamma^r_{\mu\nu} u^\mu u^\nu = \frac{g^{rr}}{2}[\partial_r g_{rr} u^r u^r - \partial_r g_{\mu\nu} u^\mu u^\nu|_{0,\phi}] \tag{60}$$

$|_{0,\phi}$ refers $\mu, \nu = 0, \phi$. For a circular orbit with $u^r = 0$, we get

$$\partial_r g_{\mu\nu} u^\mu u^\nu|_{0,\phi} = (u^0)^2 \partial_r g_{\phi\phi}[\frac{u^\phi}{u^0} - \Omega_K^+][\frac{u^\phi}{u^0} - \Omega_K^-] = 0 \tag{61}$$

where $\Omega_K^\pm$ are the Keplerian angular velocities[21] defined by

$$\Omega_K^\pm = [-\partial_r g_{0\phi} \pm \sqrt{(\partial g_{0\phi})^2 - \partial_r g_{\phi\phi} \partial_r g_{00}}]/\partial_r g_{\phi\phi} \tag{62}$$

From the normalized velocity,

$$u_\mu u^\mu = 1, \quad \partial_r (u_\mu u^\mu) = 0 \tag{63}$$

we get

$$u^r \partial_r u_r = -(u^0 \partial_r u_0 + u^\phi \partial_r u_\phi) - u_\mu \partial_r u^\mu \tag{64}$$

$$u_\mu \partial u^\mu = \frac{1}{2} u_\mu u_\nu \partial g^{\mu\nu} \tag{65}$$

Using the identities

$$u^r \partial_r u_r = -u_r u_r \partial_r g^{rr} + u^r \partial_r u^r / g^{rr}, \quad u_\mu \partial_r u^\mu = \frac{1}{2} u_\mu u_\nu \partial_r g^{\mu\nu} \tag{66}$$

we get

$$u^r \partial_r u^r = -g^{rr}(u^0 \partial_r u_0 + u^\phi \partial_r u_\phi) - \Gamma^r_{\mu\nu} u^\mu u^\nu. \tag{67}$$

and

$$T_m^{\mu r}{}_{;\mu} = \rho_m u^r \partial u^r + \rho_m \Gamma^r_{\mu\nu} u^\mu u^\nu \tag{68}$$

$$= -\rho g^{rr} u^0 \partial(u_\phi)[\frac{u^\phi}{u^0} + \frac{\partial_r u_0}{\partial_r u_\phi}]. \tag{69}$$

Hence the radial equation with only matter part gives an identity

$$\frac{u^\phi}{u^0} = -\frac{\partial_r u_0}{\partial_r u_\phi} \tag{70}$$

which can be satisfied automatically for the fluid in Keperian orbits.

The radial equation is modified when the electromagnetic interaction is introduced. The electromagnetic part of the radial equation is given by

$$T_{EM}^{\mu r}{}_{;\mu} = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}T_{EM}^{\mu r}) + \Gamma^r_{\mu\lambda}T_{EM}^{\mu\lambda} \tag{71}$$

Since the second term has no singular part, its effect can be ignored in the thin accretion disk limit. For the same reason, the radial derivative of the first term gives no effect and we get

$$T_{EM}^{\mu r}{}_{;\mu} = \frac{1}{\sqrt{-g}}\partial_\theta(\sqrt{-g}T_{EM}^{\theta r}) + \cdots \tag{72}$$

where $\cdots$ refers nonsingular parts and for $\theta = \pi/2$ we get

$$T_{EM}^{\mu r}{}_{;\mu} = -\frac{\sqrt{\Delta}}{4\pi r^2}\partial_\theta[(E^{\hat\theta}E^{\hat r} + B^{\hat r}B^{\hat\theta})]. \tag{73}$$

Since we are interested only in the singular part, the derivatives on $E^{\hat r}$ and $B^{\hat\theta}$ which are continuous on the disk have no effect on the radial equation.

$$T_{EM}^{\mu r}{}_{;\mu} = -\frac{\sqrt{\Delta}}{r^2}[\sigma E^{\hat r} - K^{\hat\phi}B^{\hat\theta})]\delta(\theta - \pi/2). \tag{74}$$

Hence the radial equation,

$$T^{\mu r}{}_{,\mu} = T_m^{\mu r}{}_{;\mu} + T_{EM}^{\mu r}{}_{;\mu} = 0 \tag{75}$$

can be written as

$$\frac{\sigma_m}{r}(\frac{g^{rr}}{2}[\partial_r(g_{rr}(u^r)^2) - (u^0)^2\partial_r g_{\phi\phi}[\frac{u^\phi}{u^0} - \Omega_K^+][\frac{u^\phi}{u^0} - \Omega_K^-]])$$

$$-\frac{\sqrt{\Delta}}{r^2}[\sigma E^{\hat r} - K^{\hat\phi}B^{\hat\theta}] = 0 \tag{76}$$

or after making use of the velocity normalization condition,

$$-\frac{\sigma_m}{r}g^{rr}u^0\partial(u_\phi)[\frac{u^\phi}{u^0}+\frac{\partial_r u_0}{\partial_r u_\phi}]-\frac{\sqrt{\Delta}}{r^2}[\sigma E^{\hat{r}}-K^{\hat{\phi}}B^{\hat{\theta}}]=0. \tag{77}$$

4 Basic equations

We have now three basic equations for the two dimensional acrretion flow dominated by the Poynting flux given by

$$(I) \quad -\partial_r(u_0)r\sigma_m u^r+rK^r(-\alpha E^{\hat{r}}+\beta\varpi B^{\hat{\theta}})=0 \tag{78}$$

or

$$\partial_r(u_0)\dot{M}_+ + 2\pi rK^r(-\alpha E^{\hat{r}}+\beta\varpi B^{\hat{\theta}})=0 \tag{79}$$

$$(II) \quad \partial_r(u_\phi)r\sigma_m u^r - rK^r\varpi B^{\hat{\theta}}=0 \tag{80}$$

or

$$\partial_r(u_\phi)\dot{M}_+ + 2\pi rK^r\varpi B^{\hat{\theta}}=0 \tag{81}$$

$$(III) \quad \frac{\sigma_m}{r}(\frac{g^{rr}}{2}[\partial_r(g_{rr}(u^r)^2)-(u^0)^2\partial_r g_{\phi\phi}[\frac{u^\phi}{u^0}-\Omega_K^+][\frac{u^\phi}{u^0}-\Omega_K^-]]) - \frac{\sqrt{\Delta}}{r^2}[\sigma E^{\hat{r}}-K^{\hat{\phi}}B^{\hat{\theta}}]=0 \tag{82}$$

or

$$-\frac{\sigma_m}{r}g^{rr}u^0\partial(u_\phi)[\frac{u^\phi}{u^0}+\frac{\partial_r u_0}{\partial_r u_\phi}]-\frac{\sqrt{\Delta}}{r^2}[\sigma E^{\hat{r}}-K^{\hat{\phi}}B^{\hat{\theta}}]=0 \tag{83}$$

These equations provide the information on how the fluid motion in the accretion flow is determined by the electromagnetic field. For example the accretion rate and the angular and radial velocities of the fluid.

5 Disk power

The power out of the disk between r_{in} and r_{out}, P_{disk}, can be calculated by integrating eq.(36) from r_{in} to r_{out}

$$P_{disk} = -2\pi \int_{r_{in}}^{r_{out}} \alpha \frac{1}{\rho}(-\alpha E^{\hat{r}} + \beta \varpi B^{\hat{\theta}}) K^{\hat{r}} \sqrt{\gamma} dr \tag{84}$$

Using the energy conservation, eq.(79), we get

$$P_{disk} = -2 \int_{r_{in}}^{r_{out}} \partial_r u_0 \dot{M}_+ dr = \Delta u_0 \dot{M}_+ \tag{85}$$

where

$$\Delta u_0 = u_0(r_{in}) - u_0(r_{out}). \tag{86}$$

This is the expected result from the energy conservation. In the magnetic field dominated accretion disk, the power out of the disk carried by Poynting flux is just the gravitational energy of the accreting fluid.

Eq.(85) indicates that the power out of the disk in this toy model is not dependent strongly on the details of the electromagnetic field configuration. The power can be calculated once $\dot{M}_+$ is known at one point r on the disk. For example, suppose that the exact solution of the relativistic equation can be approximated at a large distance r_0 by the configuration suggested by Blandford [2], then we can use the relation, for example,

$$B^{\hat{\phi}} = 2\Omega_F r B^{\hat{\theta}}, \tag{87}$$

to get the accretion rate expressed in terms of the magnetic field component perpendicular to the disk,

$$\dot{M}_+ = 4(B^{\hat{\theta}}(r_0) r_0)^2. \tag{88}$$

Then the power for a disk with Poynting flux extended to $r_{out} = \infty$ can be given by

$$P_{disk} = 4(1 - [-u_0(r_{in})])(B^{\hat{\theta}}(r_0) r_0)^2, \tag{89}$$

where we take $-u_0(\infty) = 1$.

Acknowledgments

This work was supported by grant No. 1999-2-11200-003-5 from the Korea Science & Engineering Foundation.

References

1. R.V.E. Lovelace, Nature 262, 649(1976)
2. R. D. Blandford, MNRAS 176, 465(1976)
3. P. Ghosh, MNRAS 315, 89(2000)
4. P. Ghosh and M.A. Abramowicz, MNRAS 292, 887(1997)
5. M. Livio, G.I. Ogilvie and J.E. Pringle, ApJ **512**, 100(1999)
6. C. Fendt, Astron. Astrophys. **319**, 1025(1997)
7. G.V. Ustyugova, R.V.E. Lovelace, M.M. Romanova, H. Li and S.A. Colgate, ApJ **541**, L21(2000)
8. H.K. Lee, R.A.M.J. Wijers and G.E. Brown, Phys. Rep. **325**, 83(2000)
9. H.K. Lee, G.E. Brown and R.A.M.J. Wijers, ApJ **536**, 416(2000)
10. L.-X. Li, Phys. Rev. **D61**, 084016(2000)
11. H.K. Lee and H.K. Kim, in preperation (2001)
12. T. Piran, Phys. Rep. **314**, 575(1998)
13. H.K. Lee and H.K. Kim, J. Korean Phys. Soc. **36**, 188(2000)
14. L.-X. Li and B. Paczynski, ApJ **534**, L197(2000)
15. J.-P. Lasota, in "Theory of Accretion Disks 2" ed. W.J. Duschl et al.(Dordrecht:Kluwer, 1994)
16. M.A. Abramowicz, X.-M. Chen, M. Granath and J.-P. Lasota, ApJ **471**, 762(1996)
17. C.F. Gammie and R. Popham, ApJ **498**, 313(1998)
18. H.K. Lee, Phy. Rev. D64, 043006(2001)
19. T. Damour, Phys. Rev. **D 18**, 3598(1978)
20. K.S. Thorne, R.H. Price, and D.A. Macdonald, "Black Holes; The Membrane Paradigm" (Yale University Press, New Haven and London, 1986)
21. S. Shapiro and S.A. Teukolsky,"Black Holes, White Dwarfs, and Neutron Stars" (John Wiley & Sons, New York, 1983)

PART 5

Black Hole Astrophysics

This page is intentionally left blank

BLACK HOLE MASS–PERIOD CORRELATION IN SOFT X-RAY TRANSIENTS

CHANG-HWAN LEE
School of Physics, Korea Institute for Advanced Study, Seoul 130-012, Korea
chlee@kias.re.kr

We investigate the soft X-ray transients with black hole primaries which may have been the sources of gamma-ray bursts and hypernovae earlier in their evolution. For systems with evolved donors we are able to reconstruct the preexplosion periods and find that the black-hole mass increases with the orbital period of the binary. This correlation can be understood in terms of angular-momentum support in the helium star progenitor of the black hole, if the systems with shorter periods had more rapidly rotating primaries prior to their explosion: centrifugal support will then prevent more of its mass from collapsing into the black hole on a dynamical time.

1 Introduction

Recent observations strongly suggest a connection between gamma-ray bursts and supernovae, with indication that the supernovae in question are especially energetic and of type Ib/c, i.e., core collapses of massive stars which have lost their hydrogen envelope (see van Paradijs et al. 2000, and references therein). This supports suggestions by Woosley (1993) and Paczyński (1998) for the origin of gamma-ray bursts in stellar core collapses. The hydrodynamics of a jet escaping from a star and causing its explosion was explored in detail by MacFadyen & Woosley (1999), who showed that contrary to accepted wisdom, a fairly baryon-free, ultra-relativistic jet could plough through the collapsing star and emerge with large Lorentz factors. The powering of the outflow by coupling of high magnetic fields to the rotation of the black hole (Blandford & Znajek 1977), first suggested by Paczyński (1998) in the context of gamma-ray bursts, was worked out in detail by van Putten (1999, 2001). Li has also discussed the extraction of energy from a black hole in a recent series of papers (e.g., Li 2000a).

Building on these thoughts, we have modelled both the powering of a gamma-ray burst by black-hole rotation and the stellar evolution pathways that set up favorable conditions for that mechanism Brown et al. (2000). An essential ingredient in this model is a rapidly rotating black hole (Lee et al. 2001). As suggested by MacFadyen & Woosley (1999), a massive star in a close binary will spin faster than in a single star for a number of reasons: first,

when the hydrogen envelope is lifted off by spiral-in, it will cease to serve as a sink of angular momentum for the core. Second, the tidal friction concomitant of the spiral-in process will spin up the inner region, giving it a larger angular momentum than the same region in a single star Rasio & Livio (1996). Third, tidal coupling in the close binary will tend to bring the primary into corotation with secondary. This latter process is not very efficient in the short post spiral-in life of the binaries we consider, but its effect does probably matter to the outer layers of the helium star, which are crucial for our work. With its more rapid rotation, the helium star then forms a black hole with a large Kerr parameter, which immediately after its formation (in a few seconds) begins to input power into its surroundings at a very high rate. This, then, powers both a gamma-ray burst (e.g., Brown et al. 2000) and the expulsion of the material that was centrifugally prevented from falling into the black hole. In fact, van Putten (1999, 2001) estimates that the power input into that material exceeds that into the GRB and (Li 2000b) also finds that more energy can be extracted by the disk than by the GRB. It should be noted that an initially less rapidly rotating black hole could be spun up by disk accretion quite rapidly, and start a similar process after some accretion has taken place (MacFadyen & Woosley 1999, Brown et al. 2000).

2 Gap in orbital periods

It is clear that the present orbital period is not the quantity germane to the helium star spin before explosion. Rather, we should consider the orbital period just before explosion. As we now explain, we are able to do this only for binaries with evolved companions, although we can say something about those with main sequence ones.

Brown et al. (2001b) evolved the SXTs with main-sequence companions. Seven of these have unevolved K or M-star companions with short periods $\lesssim$ 0.5 days (Lee et al. 2001). In the short-period SXTs the low-mass companion K or M stars lose mass by magnetic braking and gravitational waves. We label these AMLs for angular momentum loss.

We find a "period gap" in which binaries with ZAMS companion masses less than $2M_\odot$ shorten their periods by AML, the black hole accreting most of the ZAMS mass down to the present $\sim 0.5M_\odot$, whereas binaries with more massive companions increase in period as the donor evolves. The AML erases the history of the binaries inside the period gap, so we are unable to reconstruct the preexplosion period. Also, the last $\sim 1 - 2M_\odot$ of the present black hole mass can have been accreted from the donor, so it is somewhat larger than at the time of formation. We should therefore reconstruct the preexplosion

orbits only for the binaries with evolved companions, including Il Lupi and Nova Scorpii (Lee et al. 2001).

3 The evolution of soft X-ray transients

Brown et al. (1999), Wellstein & Langer (1999) and Brown et al. (2001a) showed that massive He stars could evolve into high-mass black holes only if they were covered with hydrogen during most of their He core burning era (Case C mass transfer in binaries). In case A or B mass transfer in binaries (Roche Lobe overflow in main-sequence or red-giant stage) the Fe core that was left was too low in mass to go into a high-mass black hole. Brown et al. (2001a) showed that high-mass black holes could be formed only if the mass was taken off of the black hole progenitor after helium core burning was finished; i.e. Case C mass transfer. We limit our evolutionary scenarios here to case C mass transfer with primaries of masses of 20–35 $M_{\odot}$, and substantially lower-mass secondaries.

A major uncertainty in the evolution of all compact X-ray binaries is the phase of spiral-in that occurred in their evolution: these binaries are initially very wide, and when the primary fills its Roche lobe and transfers mass to the secondary, the mass transfer leads to instability, resulting in the secondary plunging into the primary's envelope. Next, dissipation of energy of the secondary causes the primary's envelope to be ejected, and the orbit to shrink. Following the original work by Webbink (1984), Brown et al. 2001b write the standard formula for common envelope evolution as

$$\frac{GM_pM_e}{\lambda R} = \frac{GM_pM_e}{\lambda r_L a_i} = \alpha_{ce}\left(\frac{GM_{\rm He}M_d}{2a_f} - \frac{GM_pM_d}{2a_i}\right) \qquad (1)$$

where M_p is the total mass of the BH progenitor star just before the common envelope forms, M_e is the mass of its hydrogen envelope, $M_{\rm He}$ is the mass of its core, $a_{\rm i}$ and $a_{\rm f}$ are the initial and final separation, before and after the common envelope, respectively, and r_L is the dimensionless Roche-lobe radius. This equation essentially relates the loss of orbital energy of the secondary to the binding energy of the ejected envelope.

Given the parameters of the system at first Roche contact, when spiral-in starts, the final separation is determined by the product of λ and α_{ce}, the efficiency of the energy conversion. In general, these parameters are only the simplest recipe prescription for the complex hydrodynamical interaction during spiral-in. While we therefore cannot predict the value of $\lambda\alpha_{ce}$ from first principles, we can try to find its value from constraints in some systems, and then assume it is the same for all similar systems. The general properties of

SXTs and the specific cases of V4641 Sgr and IL Lupi favor $\lambda\alpha_{ce} \sim 0.2$, which we therefore adopt as a general efficiency for the evolution of other transient sources (Lee et al. 2001).

From the above theory, certain regularities follow for the system behaviour as a function of companion mass. First, the binding energy relation for spiral-in (Eq. 1) shows that very nearly $a_f \propto M_d$, with not much variation due to other aspects of the systems (see previous section). Furthermore, the relation between Roche lobe radius and donor mass when $M_d \ll M_{BH}$ implies that $R_L/a_f \propto M_d^{1/3}$ (e.g., Eggleton 1983). As a result, the Roche lobe radius of the donor just after spiral-in will scale with donor mass as $R_L \propto M_d^{4/3}$. On the other hand, the donor radius itself depends on its mass only as $R_d \propto M_d^{0.8}$. Therefore, a low-mass donor is more likely to fill its Roche lobe immmediately after spiral-in. We thus expect there to be some minimum donor mass for SXTs, below which the spiral-in is so deep that the spiral-in ends in merger (because the donor overfills its Roche lobe before the end of spiral-in), and no black-hole binary is formed. Above this minimum mass, there will be a range of donor masses that are close enough to filling their Roche lobes after spiral-in that they will be tidally locked and will come into contact via AML. Above this, there will be a range of mixed evolution, where both AML and nuclear evolution play a role. Finally, for the most massive donors, the post-spiral-in orbits will be too wide for AML to shrink them much, so mass transfer will be initiated only via nuclear expansion of the donor. Of course, the ranges of case C radii of stars and variations of primary masses will ensure that the boundaries between these regions are not sharp: near the boundaries the fate of the system depends on its precise initial parameters.

4 Reconstructing the pre-explosion orbits

The most extensive evolutionary studies have been made for Nova Scorpii. Starting from the work of Regös et al. (1998) who make the case that the companion is in late main sequence evolution, Beer & Podsiadlowski (2001) carry out extensive numerical calculations of the evolution, starting with a preexplosion mass of $2.5M_\odot$ and separation of $\sim 6R_\odot$. More schematically we arrived at a preexplosion mass of $1.91M_\odot$ and separation of $5.33R_\odot$ (Lee et al. 2001). We, therefore, have an 0.4 day preexplosion period. With $\sim 6M_\odot$ mass loss in the explosion (Nelemans et al. 1999), nearly half the system mass, the binary period increases to 1.5 day, well beyond the period gap. We believe that this very high mass loss explains many of the special features of Nova Scorpii. After explosion the binary evolves to its present period by nearly conservative mass transfer. Our estimate is that $0.41M_\odot$ is transferred from the donor to

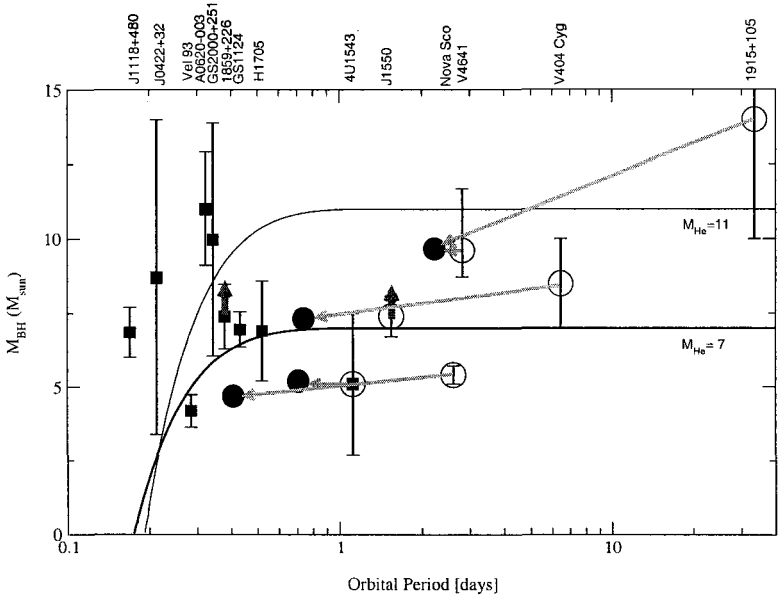


Figure 1: Reconstructed pre-explosion orbital period vs. black hole masses of SXTs with evolved companion (Lee et al. 2001). The reconstructed pre-explosion orbital periods & black hole masses are marked by filled circles, and the current observation of binaries with evolved companions are marked by big open circles. The solid lines indicate the possible ranges of black hole masses with polytropic index $n = 3$ (radiative), for given pre-explosion spin periods which are assumed the same as the pre-explosion orbital period. Here we used $R_{He} = 0.22(M_{He}/M_{\odot})^{0.6} R_{\odot}$.

the black hole.

We were able to evolve V4641 Sgr from the same ZAMS $M_{giant} = 30M_{\odot}$ as Nova Scorpii. We found that $a_f = 18R_{\odot}$ and the present black hole mass are good estimates of the preexplosion properties (Lee et al. 2001).

Recently Greiner (2001) have determined the period and black hole mass of GRS 1915+105 to be 33.5 day and $14\pm4M_{\odot}$. Starting from a binary with the properties of V4641 Sgr and allowing for $\sim 4.6M_{\odot}$ to be transferred from the donor to the black hole in assumed conservative mass transfer, we position the preexplosion period and black hole mass of GRS 1915+105 at the same point as V4641 Sgr (Lee et al. 2001).

If we assume a ZAMS $M_d = 2.5M_{\odot}$ and a $7M_{\odot}$ black hole for IL Lupi the period would be 0.7 day. We will see from our Fig. 1 that we predict little

mass loss in the explosion for such a period. Conservative transfer of $0.4M_{\odot}$ from the companion to the black hole would bring the period to the present 1.123 day. Starting from an initially lower mass black hole, less mass transfer would be needed to bring the binary to its present period (Lee et al. 2001).

The black hole in V404 Cyg appears to be somewhat more massive than in Il Lupi, so we begin with a similar mass companion, but a $9M_{\odot}$ black hole, which would have a period of 0.63 day. Again, we neglect mass loss in the explosion, although a small correction for this might be made later. Conservative transfer of $\sim 1.5M_{\odot}$ from the donor to the black hole then brings the period close to the present one (Lee et al. 2001).

5 Angular momentum in black hole progenitors

It is, in general, a difficult and unsolved problem to calculate the angular momentum of a stellar core at any given time. Even if we make the usual assumption that the rotation is initially solid-body, and not very far away from the maximal stable rotation frequency, the viscous coupling between the various layers of the star as it evolves is poorly known, and thus it is hard to be very quantitative. The general trend, however, is that the core will shrink and the envelope expand. In absence of viscous coupling, every mass element retains its angular momentum, and hence the core spins up as the envelope spins down, setting up a strong gradient in rotation frequency between the core and the envelope. Viscosity will then act to reduce this gradient, transporting angular momentum from the core to the envelope. Spruit & Phinney (1998) computed the coupling for a specific set of assumptions, to investigate the initial spin periods and magnetic fields of pulsars. They found that for a long time the whole star is kept in corotation, but during the later stages of nuclear burning the viscous coupling cannot keep up with the rapidly changing stellar structure. Livio & Pringle (1998) then showed, using novae, that the coupling of Spruit & Phinney was too strong, and thus that viscosity would fail to keep up with the evolving star somewhat earlier.

In our scenario, a number of effects will increase the angular momentum of the core relative to a similar core of a single star: (1) during spiral-in, the matter somewhat inside the orbit of the secondary is spun up by tidal torques (Rasio & Livio 1996); (2) the removal of the envelope halts the viscous slowdown of the core by friction with the envelope; (3) during the post-spiral-in evolution, tidal coupling will tend to spin the helium star up even closer to the orbital period than was achieved by the first effect. This will not be a very strong effect because the duration of this phase is short, but it will affect the outer parts of the helium star somewhat, and this is the most important part

(see below).

The net result of all these effects will be that the helium star will spin fairly rapidly, especially its envelope. The core is not so crucial to our argument about the fraction of the star that can fall into the black hole, since the few solar masses in it will not be centrifugally supported even in quite short orbits. For the purpose of a definite calculation, we therefore make the following assumptions: (1) the helium star corotates with the orbit before explosion and is in solid-body rotation; (2) the mass distribution of the helium star with radius is given by a fully radiative zero-age helium main sequence star. This latter approximation is, of course, not extremely good. However, what counts is the angular momentum as a function of mass, so the fact that the mass distribution has changed from helium ZAMS to explosion would be entirely inconsequential if no redistribution of angular momentum had taken place in the interim. As we saw above, any redistribution of angular momentum would take the form of angular momentum transport towards the outer layers. This means that relative to our ideal calculations below, a better calculation would find more angular momentum in the outer layers, and therefore somewhat smaller black hole masses than the ones we calculate.

An important result for our proposed relation between SXTs and hypernovae and gamma-ray bursts is shown in Fig. 2. This figure shows the expected Kerr parameter of the black hole formed in our model. We see that for the short-period systems, this Kerr parameter is very large, 0.7–0.9. This means that we are justified in adding only the mass that immediately falls in to the black hole, because as soon as the rapidly rotating black hole is formed, it will drive a very large energy flux in the manner described by Brown et al. (2000). This both causes a GRB and expels the leftover stellar envelope. The systems with longer orbital periods do not give rise to black holes with large Kerr parameters, and thus are presumably not the sites of GRBs.

6 Conclusions

We have shown that there is an observed correlation between orbital period and black-hole mass in soft X-ray transients. We have modelled this correlation as resulting from the spin of the helium star progenitor of the black hole: if the pre-explosion orbit has a short period, the helium star spins rapidly. This means that some part of its outer envelope is centrifugally prevented from falling into the black hole that forms at the core. This material is then expelled swiftly, leading to a black hole mass much less than the helium star mass. As the orbital period is lengthened, the centrifugal support wanes, leading to a more massive black hole. The reason for swift expulsion of material held

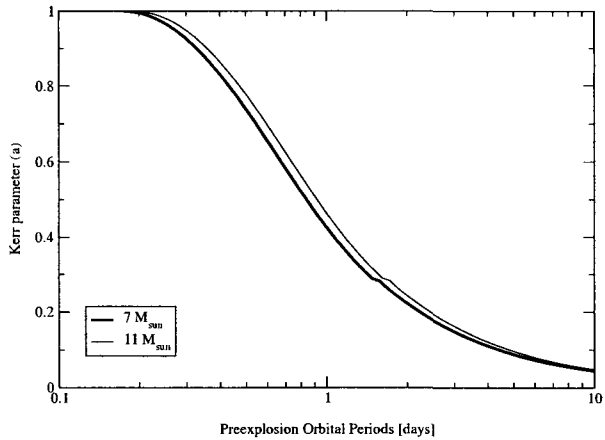


Figure 2: The Kerr parameter of the black hole resulting from the collapse of a helium star synchronous with the orbit, as a function of orbital period (Lee et al. 2001). Note that the result depends very little on the mass of the helium star, or on whether we use a simple polytrope or a more sophisticated model. The plot illustrates that rapidly rotating black holes needed for powering GRBs originate only from originally short-period SXTs.

up by a centrifugal barrier is the fact that black holes formed in our scenario naturally have high Kerr parameters (Fig. 2). This implies that they input very high energy fluxes into their surrounding medium via the Blandford-Znajek mechanism, and thus power both a gamma-ray burst and the expulsion of the material that does not immediately fall in.

References

1. Beer M. and Podsiadlowski Ph. 2001, MNRAS, submitted; astro-ph/0109136
2. Blandford R.D. and Znajek R.L. 1977, MNRAS, 179, 433
3. Brown G.E., Lee C.-H., Wijers R.A.M.J., Lee H.K., Israelian G., and Bethe H.A. 2000, New Astronomy, 5, 191
4. Brown G.E., Heger A., Langer N., Lee C.-H., Wellstein S., and Bethe H.A. 2001a, New Astronomy, 6, 457
5. Brown G.E., Lee C.-H., and Bethe H.A. 1999, New Astronomy, 4, 313
6. Brown G.E., Lee C.-H., and Tauris T. 2001b, New Astronomy, 6, 331

7. Eggleton P.P. 1983, ApJ, 268, 368
8. Greiner J. et al. 2001, Nature, 414, 522
9. Lee C.H., Brown G.E., and Wijers R.A.M.J. 2001, ApJ, submitted; astro-ph/0109538
10. Li L.X. (2000a), ApJ, 531, L111; 540, L17; 544, 375; astro-ph/0012469
11. Li L.X. (2000b), ApJ, 533, L115
12. Livio M. and Pringle J.E. 1998, ApJ, 505, 339
13. MacFadyen A. I. and Woosley S. E. 1999, ApJ, 524, 262
14. Nelemans G., Tauris T. M., and van den Heuvel E. P. J. 1999, A&A, 352, L87
15. Paczyński B. 1998, Proceedings of 4th Huntsville Gamma-Ray Bursts Symposium, Huntsville, Sep. 1997, eds. C.A. Meegan, R.D. Preece and T.M. Koshut, AIP conference proceedings 428, p. 783
16. Rasio F. A. and Livio M. 1996, ApJ, 471, 366
17. Regös E., Tout C.A., and Wickramasinghe D. 1998, ApJ, 509, 362
18. Spruit H. and Phinney E. S. 1998, Nature, 393, 139
19. Tauris T.M. and Dewi J.D.M. 2001, A&A, 369, 170
20. van Paradijs J., Kouveliotou C., and Wijers R.A.M.J. 2000, ARA&A, 38, 379
21. van Putten M.H.P.M. 1999, Science, 284, 115
22. van Putten M.H.P.M. 2001, Physics Report, 345, 1
23. Webbink R.F. 1984, ApJ, 277, 355
24. Wellstein S. and Langer N. 1999, A&A, 350, 148
25. Woosley S. E. 1993, ApJ, 405, 273

RADIATION HYDRODYNAMICAL MODEL FOR THE SUPERMASSIVE BLACK HOLE FORMATION

Nozomu KAWAKATU, Masayuki UMEMURA
Center for Computational Physics, University of Tsukuba, Tsukuba, Ibaraki 305, Japan
E-mail: kawakatu@rccp.tsukuba.ac.jp, umemura@rccp.tsukuba.ac.jp

We quantitatively scrutinize the effects of the radiation drag arising from the radiation fields in a galactic bulge in order to examine the possibility that the radiation drag could be an effective mechanism to extract angular momentum in a spheroidal system like a bulge and allow plenty of gas to accrete onto the galactic center. For this purpose, we numerically solve the relativistic radiation hydrodynamical equation coupled with accurate radiative transfer, and quantitatively assess the radiation drag efficiency. As a result, we find that in an optically thick regime the radiation drag efficiency is sensitively dependent on the density distributions of the interstellar medium (ISM). Hence, if the bulge formation begins with a star formation event in a clumpy ISM, the radiation drag will effectively work to remove the angular momentum and the accreted gas may form a supermassive black hole. Taking into account of the realistic chemical evolution, the present radiation hydrodynamical model can account for the observational results quantitatively.

1 Introduction

The recent complilation of the kinematical data for galactic centers in both inactive and active galaxies has revealed that the mass of a supermassive black hole (BH) does correlate linearly with the mass of the hosting bulge. (It is noted that the term of a bulge is used to mean a whole galaxy for an elliptical galaxy in this paper as often so.) The demography of supermassive BHs has shown a number of intriguing relations: (1) The BH mass exhibits a linear relation to the bulge mass for a wide range of BH mass with the median BH mass fraction of $f_{\rm BH} \equiv M_{\rm BH}/M_{\rm bulge} = 0.001 - 0.006$. [19,30,22,10,7,26] (2) The BH mass correlates with the velocity dispersion of bulge stars with a power-law relation as $M_{\rm BH} \propto \sigma^n$, $n = 3.75$ [10] or 4.72. [7,26,27] (3) $f_{\rm BH}$ tends to grow with the age of youngest stars in a bulge until 10^9 yr. [25] (4) In disc galaxies, the mass ratio is significantly smaller than 0.001 if the disc stars are included. [31,33] (5) For quasars, $f_{\rm BH}$ is on a similar level to that for elliptical galaxies. [21,23,38] (6) $f_{\rm BH}$ in Seyfert 1 galaxies is under debate, which may be considerably smaller than 0.001 [37,11] or similar to that for ellipticals, [23,24,38] while the BH mass-to-velocity dispersion relation in Seyfert 1 galaxies seems to hold good in a similar way to elliptical galaxies. [11,29,6] These BH-to-bulge correlations suggest that the formation of a supermassive BH is physically connected with the formation of a galactic bulge.

Recently, as a possible mechanism to work in a spheroidal system, Umemura [34]has considered the effects of radiation drag. The radiation drag is a relativistic effect, which may extract the angular momentum effectively in a spheroidal system like a bulge, so that plenty of interstellar medium (ISM) could accrete onto the galactic center. If the radiation drag works efficiently in an optically thick medium, the rate of mass accretion induced by the radiation drag is maximally $L_{\rm bol}/c^2$, [35,36,9] where $L_{\rm bol}$ is the bolometric luminosity. Umemura [34]has found that, if the maximal drag efficiency is achieved, the resultant BH-to-bulge mass ratio is basically determined by the energy conversion efficiency of the nuclear fusion from hydrogen to helium, i.e., 0.007. However, it is not very clear whether this mechanism really works efficiently in realistic situations.

In this paper, we investigate in detail the efficiency of the radiation drag in an optically thick ISM to test whether the radiation drag model is promising to account for the putative BH-to-bulge correlations. The model for the chemical evolution of elliptical galaxies suggests that an elliptical galaxy is initiated by a starburst in the early stage of ($< 10^7$yr), and evolves passively after a galactic wind event at a few $\times 10^8$ yr. [1,39,17,28] Also, in nearby starburst galaxies that have been studied, the ISM is highly clumpy. [32,12] Thus, if we consider the radiation drag in the early phase of bulge evolution, we should consider an inhomogeneous optically thick ISM. In this paper, to elucidate the mutual effect between the clumpiness of the ISM and the optical depth on the radiation drag efficiency, we build up a simple model of bulge system and accurately solve the radiation transfer in a clumpy ISM.

2 Model

We assume that a spherical galactic bulge consists of three components, that is, dark matter, stars, and a dusty ISM. The bulge radius $r_{\rm b}$ is set to be 1-10kpc by taking account of the observed sizes of elliptical galaxies or bulges in spiral galaxies. Dark matter is distributed uniformly inside the bulge. The mass of dark matter component, M_{DM}, within the bulge size is equal to the stellar mass of the galactic bulge, $M_{\rm bulge}$. As for the stellar component, we assume star clusters with a specific stellar initial mass function (see below). The star clusters are distributed uniformly inside the galactic bulge. Hereafter, 'a star' in this paper means 'a star cluster'. Here, $N_*(= 100)$ stars are distributed randomly. For a dusty ISM, we consider the two cases: one is a uniformly distributed ISM, and the other is a clumpy ISM. In the case of clumpy ISM, $N_{\rm c}(= 10^4)$ identical clouds are distributed randomly. The density $\rho_{\rm gas}$ in a cloud is assumed to be uniform. The size of a gas cloud, $r_{\rm c}$, is a parameter. Then, the optical depth of a gas cloud is $\bar{\tau} = \chi \rho_{\rm gas} r_{\rm c}$, where $\chi = n_{\rm d}\sigma_{\rm d}/\rho_{\rm g}$ with

$n_{\rm d}$ and $\sigma_{\rm d}$ being the number density of dust grains and the cross-section of dust, respectively. We suppose that the stars and the ISM clouds corotate with the angular velocity corresponding to the angular momentum obtained by the tidal torque at the linear stage of density fluctuations. Quantitatively, the angular momentum is given by the spin parameter $\lambda = (J_{\rm T}|E_{\rm T}|^{1/2})/(GM_{\rm T}^{5/2}) = 0.05$, where $J_{\rm T}$, $E_{\rm T}$, and $M_{\rm T}$ are respectively the total angular momentum, energy, and mass. [2,13] Here, the rigid rotation is assumed.

The mass range of galactic bulges is postulated to be $10^{6-13}M_\odot$. (However, in the present analysis, it is not very important to specify $M_{\rm bulge}$, because the results are scaled with $M_{\rm bulge}$ as shown below.) The total mass of the ISM, $M_{\rm gas}$, and the mass of each gas cloud, $m_{\rm c}$, are parameters.

Finally, as for the stellar evolution, we assume a Salpeter-type initial mass function (IMF) as $\phi = A(m_*/M_\odot)^{-1.35}$ for a mass range of $[m_l, m_u]$. In the present analysis, we consider an initial starburst and the subsequent passive evolution of stars. The upper mass limit is inferred to be around $40M_\odot$ in starburst regions. [5] As for the lower mass limit, some authors suggest that the IMF in starburst regions is deficient in low-mass stars, with a cut-off of about $2 - 3M_\odot$. [4,3,14] In this paper, we assume $m_l = 2M_\odot$ and $m_u = 40M_\odot$. In order to incorporate the stellar evolution, we adopt the mass-luminosity relation $(\ell_*/L_\odot) = (m_*/M_\odot)^{3.7}$, and the mass-age relation $\tau_* = 1.1 \times 10^{10}(m_*/M_\odot)^{-2.7}$yr , where m_*, ℓ_*, and τ_* are respectively the stellar mass, luminosity, and age. [20]

3 Basic equations

We put the origin at the center of the bulge, and adopt cylindrical coordinate r, ϕ, and z, where z-axis is the rotation axis of stars and gas.

Then, we have the radiation hydrodynamical equations to $O(v/c)$ as

$$\frac{dv_{\rm r}}{dt} = \frac{v_\phi^2}{r} - f_{\rm g}^{\rm r} + \frac{\chi}{c}[F^{\rm r} - (E + P^{\rm rr})v_{\rm r}] \tag{1}$$

$$\frac{1}{r}\frac{d(rv_\phi)}{dt} = \frac{\chi}{c}[F^\phi - (E + P^{\phi\phi})v_\phi] \tag{2}$$

$$\frac{dv_{\rm z}}{dt} = -f_{\rm g}^{\rm z} + \frac{\chi}{c}[F^{\rm z} - (E + P^{\rm zz})v_{\rm z}] \tag{3}$$

where E is the radiation energy density, F^α is the radiation flux, and $P^{\alpha\beta}$ is the radiation stress tensor where the non-diagonal components are null owing to the present symmetry, and $f_{\rm g}^{\rm r}$, $f_{\rm g}^z$ are respectively r- and z- component of the gravitational force.

The azimuthal equation of motion (2) is the equation of angular momentum transfer. This equation implies that the radiation flux force (the first term on the right-hand side) makes fluid elements tend to corotate with stars, whereas the radiation drag (the second term on the right-side) works to extract the angular momentum from gas. Therefore, the gain and loss of the total angular momentum sre determined by equation (2)

4 Clumpy ISM

4.1 Radiation transfer in a clumpy ISM

In an optically thin regime, the radiation fields produced by a star are

$$dE_0 = \frac{1}{c}\frac{\ell_*}{4\pi r^2},\ \mathbf{dF}_0 = \frac{\ell_*}{4\pi r^3}\mathbf{r},\ dP_0^{\rm rr} = dE_0,\ dP_0^{\phi\phi} = dP_0^{zz} \simeq 0, \tag{4}$$

where $\mathbf{r} = \mathbf{r_i} - \mathbf{r_j}$ and $r = |\mathbf{r}|$ with $\mathbf{r_i}$ being the position of the i-th cloud and $\mathbf{r_j}$ being the position of the j-th star. The physical quantities with suffix 0 are the radiation fields without extinction. The corotating flux in the azimuthal direction and the flux contributing to the radiation drag are respectively

$$dF_0^{\rm rot} = (dE_0 + dP_0^{\phi\phi})V_*, \tag{5}$$

$$dF_0^{\rm drag} = -(dE_0 + dP_0^{\phi\phi})v_{\rm gas}, \tag{6}$$

where $V_* = r_*\omega$ and $v_{\rm gas} = r_{\rm gas}\omega$ are the rotational velocities of a star and a gas cloud, with r_* and $r_{\rm gas}$ being respectively the distances from the rotational axis to a star and to a gas cloud. Next, we consider the extinction by dust in clumpy gas clouds. We calculate the radiation fields by the direct integration of the radiation transfer (see Fig. 1), where the extinction by all intervening clouds from a star to a cloud is summed up. Then, the radiation flux $\mathbf{F}$, the radiation energy density E, and the radiation stress tensor P are respectively given by

$$\mathbf{F} = \sum_{j=1}^{N_*} \mathbf{dF}_{0,j}\exp(-\tau_j), \tag{7}$$

$$E = \sum_{j=1}^{N_*} dE_{0,j}\exp(-\tau_j), \tag{8}$$

$$P = \sum_{j=1}^{N_*} dP_{0,j}\exp(-\tau_j), \tag{9}$$

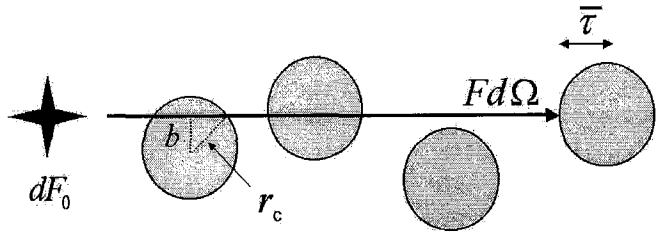


Figure 1: A schematic illustration of the radiative transfer calculations in a clumpy ISM. b and r_c are respectively the impact parameter and the radius of the clumpy gas clouds. $\bar{\tau}$ is the optical depth of a cloud.

where $\tau_j = 2\sum(1-(b/r_c)^2)^{1/2}\bar{\tau}$ is the optical depth for all intervening clouds along the light ray.

4.2 Angular momentum transfer in a clumpy ISM

By substituting (5)-(9) for (2), we can evaluate the total angular momentum loss rate as

$$\dot{J} = \frac{\chi}{c}\sum_{i=1}^{N_c} r_i(F_i^{\rm rot} - F_i^{\rm drag}), \tag{10}$$

where $F_i^\alpha = \sum_{j=1}^{N_*} dF_{0,j}^\alpha \exp(-\tau_j)$ with α being 'rot' or 'drag', and r_i is the distance from the rotational axis to a gas cloud. We assume here that the stripped gas falls onto a central massive object. Then, with the above angular momentum loss rate, we estimate the total mass of the dusty ISM assembled on to a central massive object which may evolve into a supermassive BH. By the equation (10) and the relation $\dot{M}_{\rm gas}/M_{\rm gas} = -\dot{J}/J_0$, the BH mass $M_{\rm BH}$ is assessed as

$$M_{\rm BH} = \int_0^{t(m_l)} \dot{M}_{\rm gas} dt = -\int_0^{t(m_l)} M_{\rm gas}\frac{\dot{J}}{J_0} dt, \tag{11}$$

where $t(m_l)$ is the age of a star with mass m_l.

Here, we introduce $\bar{N}_{\rm int}$ as a measure of the clumpiness of ISM. The $\bar{N}_{\rm int}$ is defined by

$$\bar{N}_{\rm int} = n_c \pi r_c^2 r_b = \frac{3}{4} N_c \left(\frac{r_c}{r_b}\right)^2, \tag{12}$$

where $n_c = N_c/\frac{4}{3}\pi r_b^3$ is the number density of gas clouds. $\bar{N}_{\rm int}$ means the average number of gas clouds that are intersected by a light ray over a bulge radius. The degree of clumpiness is larger for smaller $\bar{N}_{\rm int}$, and $\bar{N}_{\rm int} = \infty$

corresponds to a uniform distribution. With changing $\bar{N}_{\rm int}$, we investigate the relationship between the clumpiness of ISM and the radiation drag efficiency.

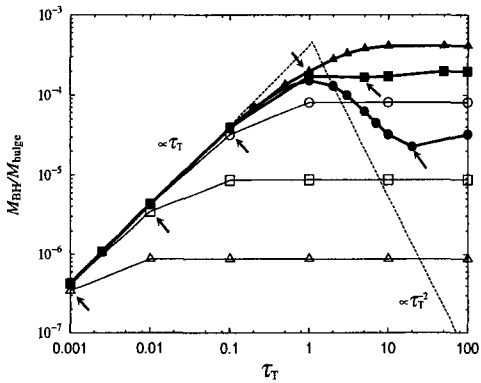


Figure 2: The BH-to-bulge mass ratio ($M_{\rm BH}/M_{\rm bulge}$) against the total optical depth ($\tau_{\rm T}$) of the bulge. The thick lines show the results for $\bar{N}_{\rm int} \geq 1$ and the thin lines are those for $\bar{N}_{\rm int} < 1$. Filled circles denote $\bar{N}_{\rm int} = 20$, filled squares $\bar{N}_{\rm int} = 5$, filled triangles $\bar{N}_{\rm int} = 1$, open circles $\bar{N}_{\rm int} = 0.1$, open squares $\bar{N}_{\rm int} = 0.01$, and open triangles $\bar{N}_{\rm int} = 0.001$. The dashed line is the analytic solution for a uniform ISM corresponding to $\bar{N}_{\rm int} \to \infty$, where $M_{\rm BH}/M_{\rm bulge} \propto \tau_{\rm T}$ in an optically thin regime and $M_{\rm BH}/M_{\rm bulge} \propto \tau_{\rm T}^{-2}$ in an optically thick regime. The arrows show the points where the optical depth of a cloud ($\bar{\tau}$) is unity. For $\bar{N}_{\rm int} \approx 1$, the radiation drag efficiency is maximal if $\bar{\tau} \geq 1$. For $\bar{N}_{\rm int} < 1$, the radiation drag efficiency is saturated when $\bar{\tau} \geq 1$.

In Fig. 2, for different $\bar{N}_{\rm int}$, we show the ratio of the resultant BH mass to the bulge mass ($M_{\rm BH}/M_{\rm bulge}$) against the total optical depth $\tau_{\rm T}$. In this figure, the dashed line is the analytic solution for a uniform ISM corresponding to $\bar{N}_{\rm int} = \infty$. The solid lines show the numerical results for a clumpy ISM, where the arrows represent $\bar{\tau} = 1$ for different $\bar{N}_{\rm int}$. The thick lines show the results for $\bar{N}_{\rm int} \geq 1$ and the thin lines are those for $\bar{N}_{\rm int} < 1$. In Fig. 2, several points are clear regarding the effects of the clumpiness on the radiation drag efficiency. First, if $\bar{N}_{\rm int} < 1$, the drag efficiency is saturated when $\bar{\tau} > 1$. Secondly, if $\bar{N}_{\rm int} \approx O(1)$, $M_{\rm BH}/M_{\rm bulge}$ grows with $\tau_{\rm T}$ and maintains a high level in optically thick ISM. Thirdly, when $\bar{N}_{\rm int}$ is much larger than unity, the drag efficiency falls down as the uniform case (see this paper for details). [15] To conclude, the combination effects of the high clumpiness of the ISM and the large covering factor of clouds are necessary for high drag efficiency.

5 Black hole-to-bulge mass ratio

The linear relation between the BH mass and the bulge mass is a direct consequence of the present radiation hydrodynamical mechanism. The possible mass accreted by the radiation drag is given by $M_{\rm max} = \int_0^{t(m_l)} L_{\rm bulge}(t)/c^2 dt$. Incorporating stellar evolution based on the present single starburst model, we find $M_{\rm max} = 1.2 \times 10^{-3} M_{\rm bulge}$. In a highly clumpy ISM where $\bar{N}_{\rm int}$ is smaller than several, the BH-to-bulge mass ratio $M_{\rm BH}/M_{\rm bulge}$ can be expressed by

$$\frac{M_{\rm BH}}{M_{\rm bulge}} = \eta \frac{M_{\rm max}}{M_{\rm bulge}} (1 - {\rm e}^{-\tau_T}) \tag{13}$$

$$= 1.2 \times 10^{-3} \eta (1 - {\rm e}^{-\tau_T}), \tag{14}$$

where η gives the radiation drag efficiency. From the above analysis, η is found to be maximally 0.34 in the optically thick limit, and then $M_{\rm BH}/M_{\rm bulge} = 4.1 \times 10^{-4}$.

The present model may be too simple in some respects to compare the results closely with the observations. For a more precise prediction of $M_{\rm BH}/M_{\rm bulge}$, it seems necessary to consider at least two effects which have not been incorporated in this simple model: one is the recycling of ISM for star formation, and the other is the QSO activity, which can enhance the mass ratio by maximally a factor of 1.7. [34]

To treat the former effect, we use an evolutionary spectral synthesis code 'PEGASE'. [8] And also, we adopt parameter sets to reproduce metallicity and color of a present-day bulge. [39] If the above effects are taken into account, the averaged mass ratio could be $M_{\rm BH}/M_{\rm bulge} \simeq 1.0 \times 10^{-3}$. This is compared with the recent observational data in Fig. 3, combined with the result by the model of the realistic chemical evolution (see this paper for details). [16] In Fig. 3, the hatched area is the prediction of the present analysis. As seen in this figure, the present radiation hydrodynamical model can account for the observational data quantitatively, although the prediction falls slightly short of the observed median mass ratio. The prediction is smaller by a factor of six than $\langle M_{\rm BH}/M_{\rm bulge} \rangle = 0.006$ by Magorrian et al. [22], while it is comparable to $\langle M_{\rm BH}/M_{\rm bulge} \rangle = 0.001$ by Merritt & Ferrarese. [26]

6 Conclusions

By assuming a simple model of a bulge, we have investigated the mutual effect between the clumpiness of the interstellar medium and the optical depth on the radiation drag efficiency for the angular momentum transfer. In a clumpy interstellar medium, we have accurately solved 3D radiation transfer to calculate

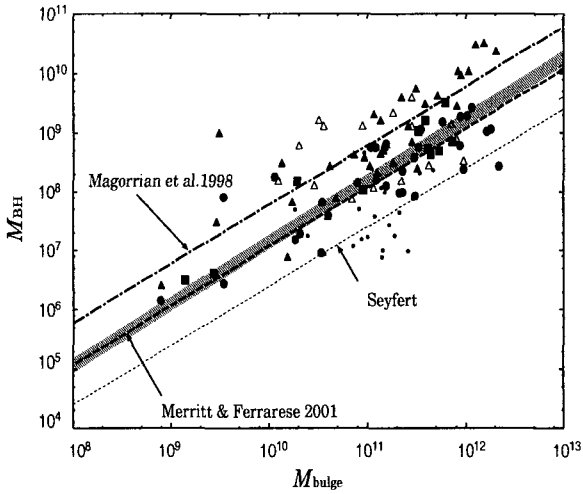


Figure 3: The relation between the BH mass and the bulge mass. The vertical axis is the BH mass, and the horizontal axis is the bulge mass, in units of $M_{\odot}$. Recent observational results are plotted by symbols. The filled triangles denote elliptical galaxies from Magorrian et al. (1998). The filled squares show elliptical galaxies from Ho (1999), Ferrarese & Merritt (2000), Kormendy & Ho (2000), and Sarzi et al. (2001). The filled circles denote elliptical galaxies from Merritt & Ferrarese (2001a). The small dots denote Seyfert galaxies from Ho (1999), Wandel (1999), and Gebhardt et al. (2000a), and the open triangles show QSOs from Laor (1998). The relation of Magorrian et al. (1998) is $M_{\rm BH} = 0.006 M_{\rm bulge}$, which is shown by a dot-dashed line; the relation of Merritt & Ferrarese (2001a) is $M_{\rm BH} = 0.001 M_{\rm bulge}$, which is shown by a dashed line; and the relation for Seyfert galaxies is $M_{\rm BH} = 2.5 \times 10^{-4} M_{\rm bulge}$ (Sarzi et al. 2001), which is shown by a thin dashed line. The hatched area shows the prediction of this papaer. The lower bound is the model of the realistic chemical evolution. The upper bound is the model incoporating the effect of the AGN activity.

the radiation drag force by the rotating bulge stars. Then, for the radiation drag to work effectively, it is necessary that the covering factor is larger than unity and that the distributions of ISM is highly clumpy. The present radiation hydrodynamical mechanism accounts for the linear relation between the mass of a supermassive BH and the mass of a galactic bulge. The range of the predicted mass ratio is from $M_{\rm BH}/M_{\rm bulge} \simeq 1 - 2 \times 10^{-3}$.

Acknowledgments

We thank M. Mori, T. Nakamoto, K. Ohsuga and H. Susa for helpful discussion. We are grateful to Prof. Ferrarese for many useful coments on this work. We also thank the organizers for this stimulating meeting in a very pleasant atmosphere. Numerical simulations were performed with facilities at Center of Computational Physics, University of Tsukuba. This work was supported in part by the Grant-in-Aid of the JSPS, 11640225.

References

1. N. Arimoto and Y. Yoshii, A&A **164**, 260 (1986)
2. J. Barnes and G. Efstathiou, ApJ **319**, 575 (1987)
3. S. Charlot *et al*, ApJ **419**, L57 (1993)
4. J. S.Doane and W. G. Mathews, ApJ **419**, 573 (1993)
5. R. Doyon, P. J. Puxley and R. D. Joseph, ApJ **397**, 117 (1992)
6. L. Ferrarese *et al*, ApJ **555**, L79 (2001)
7. L. Ferrarese and D.Merritt, ApJ **539**, L9 (2000)
8. M. Fioc and B.R-.Volmerange, A&A **326**, 950 (1997)
9. J. Fukue, M. Umemura and S. Mineshige, PASJ **49**, 673 (1997)
10. K. Gebhardt *et al*, ApJ **539**, L13 (2000a)
11. K. Gebhardt *et al*, ApJ **543**, L5 (2000b)
12. K. Gordon *et al*, ApJ **487**, 625 (1997)
13. A. Heavens and J. Peacock, MNRAS **232**, 339 (1988)
14. J. K. Hill *et al*, ApJ **425**, 122 (1994)
15. N. Kawakatu and M. Umemura, MNRAS **329**, 572 (2002)
16. N. Kawakatu and M. Umemura, 2002, in preparation
17. T. Kodama and N. Arimoto, A&A **320**, 41 (1997)
18. J. Kormendy J. and L.C. Ho, Encyclopedia of Astronomy and Astrophysics (2000)
19. J. Kormendy and D. Richstone, ARA&A **33**, 581 (1995)
20. K.R. Lang, Astrophysical Data : Planets and Stars. Springer - Verlag, Berlin(1974)
21. A. Laor, ApJ **505**, L83 (1998)

22. J. Magorrian *et al*, ApJ **115**, 2285 (1998)
23. R.J. McLure and J.S. Dunlop, MNRAS (2001a), submitted
24. R.J. McLure and J.S. Dunlop, MNRAS **327**, 199 (2001b)
25. M. R. Merrifield, D. A. Forbes and A. Terlevich, MNRAS **313**, L29 (2000)
26. D. Merritt and L. Ferrarese, MNRAS **320**, L30 (2001a)
27. D. Merritt and L. Ferrarese, ApJ **547**, 140 (2001b)
28. M. Mori *et al*, ApJ **478**, L21 (1997)
29. C. H. Nelson, ApJ **544**, L91 (2000)
30. D. Richstone *et al*, Nat **395A**, 14 (1998)
31. P. Salucci *et al*, MNRAS **317**, 488 (2000)
32. D.B. Sanders *et al*, ApJ **325**, 74 (1988)
33. M. Sarzi *et al*, ApJ **550**, 65 (2001)
34. M. Umemura, ApJ **560**, L29 (2001)
35. M. Umemura, J. Fukue and S. Mineshige, ApJ **479**, L97 (1997)
36. M. Umemura, J. Fukue and S. Mineshige, MNRAS **299**, 1123 (1998)
37. A. Wandel, ApJ **519**, L39 (1999)
38. A. Wandele, ApJ (2001), submitted
39. Y. Yoshii and N. Arimoto, A&A **188**, 13 (1987)

MAGNETIC PROPERTIES OF NEUTRON STAR MATTER AND PULSED GAMMA EMISSION OF SOFT GAMMA REPEATERS

SERGEY BASTRUKOV
Center for High Energy Astrophysics and Isotope Studies
Ewha Womans University, Seoul 120-750, Korea
Center for High Energy Physics, Kyungpook National University,
Daegu 702-701, Korea
Joint Institute for Nuclear Research ,141980 Dubna, Russia
E-mail: bast@mm.ewha.ac.kr

JONGMANN YANG, MIYOUNG KIM
Center for High Energy Astrophysics and Isotope Studies
Ewha Woman University, Seoul 120-750, Korea
Center for High Energy Physics, Kyungpook National University,
Daegu 702-701, Korea
E-mail: jyang@mm.ewha.ac.kr; E-mail: MiyoungK@mm.ewha.ac.kr

DMITRY PODGAINY
Joint Institute for Nuclear Research, 141980 Dubna, Russia
E-mail: podgainy@cv.jinr.ru

We examine hypothesis that long-periodic pulsed emission of soft gamma repeaters (SGR's) in the quiescent regime of their radiation owe its origin to magneto-mechanical pulsations of magnetar triggered by gamma-bursting starquake. Two alternative models are discussed for a neutron star undergoing global differentially-rotational pulsations in the regime of strong coupling between mechanical displacements of nuclear material and seed magnetic field inherited by neutron star from its massive progenitor. First is the single-component magneto-hydrodynamic (MHD) model and second is magneto-elastodynamic (MED) model presuming permanent magnetization of neutron star matter. Based on the energy variational principle analytic estimates for periods of non-radial torsional MHD and MED modes are derived. Numerically we found that periods of MED mode fall into the realm of periods of pulsed emission of magnetars. This correspondence is interpreted as an indirect evidence that detected γ-pulses are caused by differentially-rotational vibrations of permanently magnetized neutron star in which self-gravity has been brought to equilibrium by degenerate nuclear matter in the state of paramagnetic saturation promoted by Pauli mechanism of alignment of spin magnetic moments of neutrons along the seed magnetic field.

1 Introduction

Resent years have seen resurgence of interest in magnetic properties of neutron star matter [1] which has been stimulated by identification of soft gamma

repeaters with magnetars [2,3] - young neutron stars endowed with magnetic field of ultastrong intensity [4,5]. The characteristic feature of radiative activity of this subclass of magnetic neutron stars is two regimes of gamma-emission: the irregular γ-bursting regime caused by starquakes [6,7] and quiescent regime characterized by long-periodic 5-8 second γ-pulses (markedly longer than those for radio pulsars, about 0.8 second) the origin of which can be attributed either to rotation or to pulsations triggered by gamma-bursting starquake. In this work we explore the latter possibility. In doing this we examine two models of differentially-rotational pulsations owing its existence to strong coupling between mechanical displacements and magnetic field. First is the single-component magneto-hydrodynamic (MHD) model for non-radial torsional Alfvén pulsations implying uniform distribution of charged particles which imparts to the neutron star matter features of an ideal magnetoplasma - a homogeneous infinitely conductive medium with heavily suppressed magnetic properties. The second is magneto-elastodynamic (MED) model of non-radial torsional magneto-elastic pulsations presuming strong magnetic polarizability of neutron star matter, the feature generic to ideal paramagnetic insulator - a non-conductive medium permanently magnetized to the state of paramagnetic saturation.

It is to be noted that over the years the torsional mode of elastic shear vibrations of the neutron star crust have been studied by several authors [8,9,10]. Also the seismic shear waves confined in the crustal region of neutron stars have been investigated in the starkquake model of γ bursts [6,7,11]. In this communication we focus on the bulk magneto-mechanical torsional modes associated with global long wavelength differentially-rotational pulsations of neutron star.

2 Non-radial elastic torsional pulsations of neutron star

The standard approach to the elastodynamics of oscillatory motions of incompressible isotropic nuclear matter implies the reduction of the Lamé equation $\mathbf{u}$ to Helmholtz equation for standing shear wave:

$$\rho \frac{\partial^2 u_i}{\partial t^2} + \frac{\partial \sigma_{ij}}{\partial x_j} = 0, \quad \sigma_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad \frac{\partial u_k}{\partial x_k} = 0, \tag{1}$$

$$\nabla^2 \mathbf{u} + k^2 \mathbf{u} = 0, \quad k^2 = \frac{\omega^2}{V_t^2}, \quad V_t^2 = \frac{\mu}{\rho}, \tag{2}$$

where $\mathbf{u}$ is the displacement field of material element of matter, σ_{ij} is the tensor of mechanical stresses subjected to Hookean law of elastic deformations with μ being the shear modulus; V_t is the speed of bulk shear wave. The frequency ω is fixed by boundary condition of free-from-stresses surface $n_i\, \sigma_{ij} = 0$, where n_i

stands for the unit vector normal to surface; the summation over repeated indices is presumed. This method has been extensively exploited in most of above mentioned studies of differentially-rotational pulsations of neutron stars with particular attention to non-radial surface t-mode[8] associated with the torsional standing wave confined in the infinitely thin surface layer. The frequency of vibrations of crustal material in this wave is given by $\omega = (c_t/R)[L(L+1)]^{1/2}$, where c_t is the speed of surface shear wave and R is the star radius.

Our goal here is to evaluate the frequency of bulk t-mode attributed to long wavelength differentially-rotational global pulsations of entire volume of neutron star. In the limit of long wavelengths, $\lambda = (2\pi/k) \to \infty$, the Helmholtz equation for standing shear wave, Eq.(2), is replaced by Laplace equation, $\nabla^2 \mathbf{u}(\mathbf{r},t) = 0$, since $k \to 0$. In this case the frequency can be computed on the basis of energy variational principle; such an approach has been used in the starkquake theory of gamma bursts[11]. In searching for periods of normal long wavelength pulsations it is convenient to introduce separable representation for displacement field: $\mathbf{u}(\mathbf{r},t) = \mathbf{a}(\mathbf{r})\alpha(t)$, where $\mathbf{a}(\mathbf{r})$ is the field of instantaneous displacements. It is obvious

$$\nabla^2 \mathbf{a}(\mathbf{r}) = 0, \qquad \nabla \cdot \mathbf{a}(\mathbf{r}) = 0. \tag{3}$$

The regular in origin solution of (3) corresponding to bulk torsional mode is given by toroidal vector field

$$\mathbf{a} = N_t \nabla \times \mathbf{r}\, r^L P_L(\eta) = [\boldsymbol{\phi}(\mathbf{r}) \times \mathbf{r}], \quad \boldsymbol{\phi}(\mathbf{r}) = N_t \nabla r^L P_L(\eta), \tag{4}$$

where $\boldsymbol{\phi}(\mathbf{r})$ is the angular vector field of torsion. This solution exhibits the non-radial character of displacements: the function r^L has no nodes, from which the name non-radial bulk t-mode is derived; hereafter $P_L(\eta)$ stands for the Legendre polynomial of the multipole degree L and $\eta = \cos\theta$. Figure 1 is a snapshot of torsional quadrupole, $L = 2$, and octupole, $L = 3$, deformations of neutron star. It is convenient to use the tensor form for fluctuating variables

$$u_i(\mathbf{r},t) = a_i(\mathbf{r})\,\alpha(t), \;\; \sigma_{ij}(\mathbf{r},t) = 2\mu\, a_{ij}(\mathbf{r})\,\alpha(t), \;\; a_{ij} = \frac{1}{2}\left(\frac{\partial a_i}{\partial x_j} + \frac{\partial a_j}{\partial x_i}\right), \tag{5}$$

where $a_{ij}(\mathbf{r})$ is the strain tensor. Multiplying (1) by $a_i(\mathbf{r})$ and integrating over the whole volume of the star we obtain

$$\frac{dH}{dt} = 0, \;\; H = \frac{M_e \dot{\alpha}^2}{2} + \frac{K_e \alpha^2}{2}, \;\; M_e = \int \rho\, a_i\, a_i\, dv, \;\; K_e = 2\mu \int a_{ij}\, a_{ij}\, dv, \tag{6}$$

where M_e and K_e stand for inertia and stiffness of elastic non-radial pulsations, respectively. The frequency $\omega = (K_e/M_e)^{1/2}$ of non-radial bulk t-mode is given

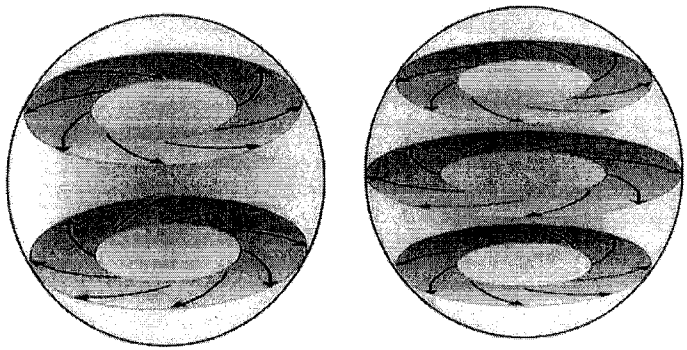


Figure 1: Geometrical illustration of torsional quadrupole and octupole vibrations.

by [12]

$$\omega_t^2 = \omega_E^2[(2L+3)(L-1)], \quad \omega_E^2 = \frac{V_t^2}{R^2} = \frac{\mu}{\rho R^2}, \tag{7}$$

where ω_E is the natural unit of frequency of elastic shear vibrations. For pure neutron matter condensed to the normal nuclear density $\rho = 2.8 \times 10^{14}$ g cm^{-3}, the shear modulus $\mu \approx 10^{33} - 10^{34}$ dyn cm^{-2}. The periods $P = \nu^{-1} = (2\pi\omega)^{-1}$ of lowest overtones fall in the interval 0.2-20.0 millisecond [13,14].

3 Non-radial torsional MHD pulsations

The standard argument for applicability of single-component hydromagnetic model is that it has to do with infinitely conducting medium. For the neutron star matter the latter property is motivated by the presence of electrons and protons. The MHD approach implies that multicomponent nuclear matter can be replaced by single-component perfectly conducting continuous medium whose oscillatory behavior is described in terms of velocity of material flow $\delta\mathbf{V}$, and magnetic flux density $\delta\mathbf{B}$. The linear MHD equations of dissipative-free oscillatory motions read

$$\rho \frac{\partial \delta\mathbf{V}(\mathbf{r},t)}{\partial t} = \frac{1}{4\pi}(\mathbf{B}\cdot\nabla)\,\delta\mathbf{B}(\mathbf{r},t), \quad \nabla\cdot\delta\mathbf{V} = 0, \tag{8}$$

$$\frac{\partial \delta \mathbf{B}(\mathbf{r},t)}{\partial t} = \nabla \times [\delta \mathbf{V}(\mathbf{r},t) \times \mathbf{B}], \quad \mathbf{B} = \text{const.} \tag{9}$$

The capability of transmitting transverse Alfvén waves is the most conspicuous feature of infinitely conducting medium. In this wave the material flow and magnetic flux undergo coupled oscillations whose propagation is characterized by dispersion-free law $\omega = V_A\, k$, where $V_A = B/\sqrt{4\pi\rho}$ is the Alfvén speed. In view of this it is often stressed that magnetic field frozen-in the flow imparts to the perfectly conducting medium features typical of Hookean elastic material capable of transmitting transverse wave of elastic shear. The frequency of bulk non-radial hydromagnetic pulsations of neutron star can be computed by the above expounded energy variational method. Multiplying (9) by $\delta\mathbf{V}$ and integrating over the star volume we get

$$\frac{\partial}{\partial t} \int \frac{\rho\, \delta \mathbf{V}^2}{2} d\, v + \frac{1}{4\pi} \int \delta \mathbf{V} \cdot [(\mathbf{B} \cdot \nabla) \delta \mathbf{B}]\, dv = 0. \tag{10}$$

Next step is to use separable form for fluctuating variables

$$\delta \mathbf{V}(\mathbf{r},t) = \mathbf{a}(\mathbf{r}) \dot{\alpha}(t), \qquad \delta \mathbf{B}(\mathbf{r},t) = \mathbf{b}(\mathbf{r}) \alpha(t), \quad \mathbf{b}(\mathbf{r}) = (\mathbf{B} \cdot \nabla) \mathbf{a}(\mathbf{r}). \tag{11}$$

The equation for $\mathbf{b}$ results from (9) after elimination of time derivative. Inserting (11) into (10), we arrive at the Hamiltonian of normal Alfvén vibrations

$$H = \frac{M_a \dot{\alpha}^2}{2} + \frac{K_a \alpha^2}{2}, \quad M_a = \int \rho\, a_i\, a_i\, dv, \quad K_a = \frac{1}{4\pi} \int b_i\, b_i\, dv. \tag{12}$$

Given the above mentioned analogy between hydromagnetic oscillatory behavior of magnetoplasma and Hookean elastic solid we take the field a_i in form of toroidal field corresponding to non-radial torsional vibrations of an elastic spherical mass, Eq.(4). The velocity field for differentially - rotational pulsations under consideration reads

$$\delta \mathbf{V}(\mathbf{r},t) = [\delta \mathbf{\Omega}(\mathbf{r},t) \times \mathbf{r}], \; \delta \mathbf{\Omega}(\mathbf{r},t) = \boldsymbol{\phi}(\mathbf{r})\, \dot{\alpha}(t), \; \boldsymbol{\phi}(\mathbf{r}) = N_l \nabla\, r^L P_L(\mu). \tag{13}$$

For the seed magnetic field pointing to z-direction we obtain [15]

$$\omega_a^2 = \omega_A^2 (L^2 - 1) \frac{3L+3}{2L-1}, \quad \omega_A^2 = \frac{V_A^2}{R^2} = \frac{B^2}{4\pi\rho R^2}, \tag{14}$$

where ω_A is natural unit of frequency of Alfvénic pulsations. Numerical estimates for the periods torsional a-mode are presented in Figures 2 and 3 and discussed in the last section.

4 Non-radial MED pulsations of paramagnetic neutron star

In [16] arguments have been given that Pauli paramagnetism can promote permanent magnetization of neutron star matter due to alignment of neutron magnetic moments along the seed magnetic field inherited by neutron star from its massive progenitor. It has been shown that continuum mechanics of paramagentic nuclear matter (whose constitutive equation is given by $\mathbf{M} = \chi \mathbf{B}$, where χ is paramagnetic susceptibility) can be described by equations of magneto-elastic dynamics

$$\frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{V} = 0, \quad \rho \frac{d\mathbf{V}}{dt} = \frac{1}{2} \nabla \times [\mathbf{M} \times \mathbf{B}], \quad \frac{d\mathbf{M}}{dt} = [\Omega \times \mathbf{M}]. \tag{15}$$

Deserved for particular emphasis is the driving magneto-elastic bulk force $\mathbf{F} = (1/2)\, \nabla \times [\mathbf{M} \times \mathbf{B}]$, originating from magnetic torque density $\mathbf{T} = (1/2)\, [\mathbf{M} \times \mathbf{B}]$ which arises when magnetization deviates from the direction of equilibrium magnetic field of the star; evolution of $\mathbf{M}$ is described by equation of local precession with local angular velocity $\Omega = (1/2)[\nabla \times \mathbf{V}]$ around axis of magnetic anisotropy. The unique feature of paramagnetic nuclear matter is its capability of transmitting perturbation by transverse magneto-elastic wave [16]

$$\omega = V_M\, k, \qquad V_M = \sqrt{\frac{MB}{4\rho}} = \sqrt{\frac{\chi}{4\rho}}\, B, \tag{16}$$

in which magnetization and material flow undergo coupled oscillations; kinematically this wave has many features in common with Alfvén wave while the underlying physics is absolutely different. To compute the frequency of non-radial magneto-torsional mode one can again take advantage of energy variational principle. The equation of energy balance reads

$$\frac{\partial}{\partial t} \int \frac{\rho \delta \mathbf{V}^2}{2}\, dv = \int [\delta \mathbf{M} \times \mathbf{B}] \cdot \delta \Omega\, dv, \tag{17}$$

By substituting into (17) the factorized representation of fluctuating variables

$$\delta \mathbf{V}(\mathbf{r}, t) = \mathbf{a}(\mathbf{r}) \dot{\alpha}(t), \quad \delta \Omega(\mathbf{r}, t) = \frac{1}{2} [\nabla \times \mathbf{a}(\mathbf{r})]\, \dot{\alpha}(t), \tag{18}$$

$$\delta \mathbf{M}(\mathbf{r}, t) = \mathbf{m}(\mathbf{r})\, \alpha(t), \quad \mathbf{m} = \frac{1}{2} [[\nabla \times \mathbf{a}(\mathbf{r})] \times \mathbf{M}], \tag{19}$$

we obtain

$$H = \frac{M_m \dot{\alpha}^2}{2} + \frac{K_m \alpha^2}{2}, \; M_m = \int \rho a_i\, a_i\, dv, \; K_m = \chi^{-1} \int m_i\, m_i\, dv. \tag{20}$$

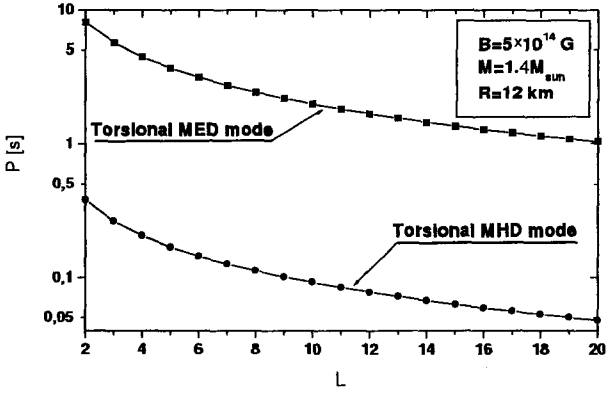


Figure 2: Period of bulk non-radial torsional magneto-hydrodynamic (MHD) and magneto-elastodynamic (MED) global pulsations of neutron star versus multipole degree L of vibration.

In the case of seed magnetic field directed along the polar axis the frequency is given by [16]

$$\omega^2 = \omega_M^2 \, (L^2 - 1)\frac{(2L+3)}{(2L-1)}, \qquad \omega_M^2 = \frac{V_M^2}{R^2} = \frac{MB}{4\rho R^2} = \frac{\chi B^2}{4\rho R^2}. \qquad (21)$$

where ω_M is the natural unit of frequency of magneto-torsional pulsations. Numerical values of periods are presented in Figures 2 and 3 and discussed in the next section.

5 Discussion

In Figure 2 we plot periods of MHD and MED modes as functions of multipole degree of vibration for homogeneous neutron star model with canonical parameters $M = 1.4M_\odot$ and $R = 12$ km corresponding to average density of normal nuclear matter $\rho = 2.8 \times 10^{14}$ g/cm; the paramagnetic susceptibility has been taken for pure neutron degenerate matter $\chi = (3/2)\rho\mu^2/\epsilon_F = 1.3 \times 10^{-4}$. These estimates suggest that both kinds of regarded magneto-mechanical pulsations can bear on pulsed emission of SGR'S while the periods of the MHD mode are shorter than those for MED mode. The most striking feature of general trend pictuted in Figure 2 is the gradual elongation of periods with decreasing multipole degree of vibration L. It seems quite plausible that just after starkquake the triggered vibrations of magnetar are characterized by large values of L, whereas the oscillatory state of more late stage is characterized by small values of L and, hence, more longer periods.

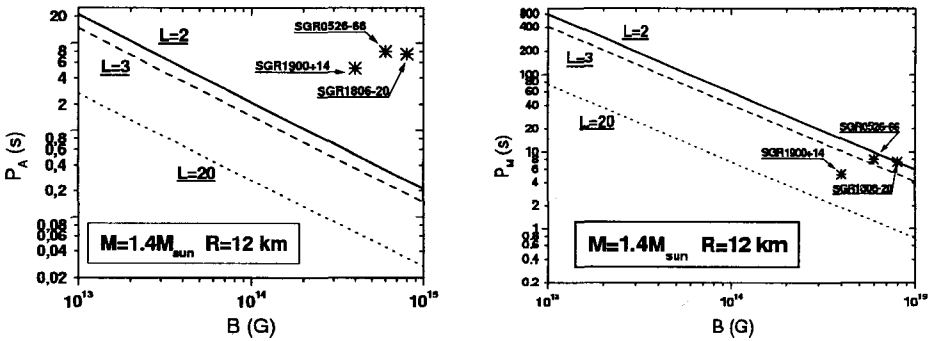


Figure 3: Periods of non-radial torsional Alfvén (left panel) and magneto-elastic (right panel) quadrupole and octupole bulk modes versus the magnetic field intensity; the data on pulsed gamma-emission of soft gamma repeaters is pictured by stars.

In Figure 3 the periods of non-radial quadrupole and octupole torsional modes versus magnetic field intensity B are plotted in juxtaposition with data on pulsed gamma emission of soft gamma repeaters. The fact that periods of MED mode are in the realm of data allows us to speculate about indirect evidence that considerable contribution to the ultrastrong magnetic field of magnetars belongs to permanent magnetization of neutron star material. Energetically the pulsed radiation of SGR's can be attributed to conversion of the energy of magneto-mechanical vibrations of paramagnetic neutron star into the energy of pulsed gamma-emission. It has been shown early that MED pulsations can produce magnetosphere with fluctuating both polarization charge and magnetic field lines frozen-in the body of neutron star [17]. It seems plausible, therefore, to expect that the radiation mechanism of gamma repeaters in the quiescent regime of their activity is fairly similar to that for magnetospheric γ-emission of pulsars attributed to curvature radiation [18]. The difference is that for these latter the pulsed emission is due to rigid-body rotation whereas for magnetars the gamma-pulses can owe its origin to the torsional magneto-mechanical pulsations; the elongation of period of gamma-pulses can be attributed to gradual decreasing of multipole degree of pulsations triggered by starquake.

Acknowledgments

This work is partly supported by Korea Research Foundation Grant (KRF-2001-041-D00052).

1. C.Y.Cardall, M. Prakash and J. Lattimer, *Astrophys. J.* **554**, 322 (2001).

2. C. Thompson and R.C. Duncan, *MNRAS* **275**, 255 (1995); *Astrophys. J.*, **473**, 322 (1996)
3. C. Kouveliotou, et al., *Astrophys. J.* **510**, L115 (1999).
4. R.C. Duncan and C. Thompson, *Astrophys. J.*, **392**, L9 (1992).
5. B. Paczyński, *Acta Astron.* **42**, 145 (1992).
6. O. Blaes, R. Blandford, P. Goldreich and P. Madau, *Astrophys. J.* **343**, 839 (1989).
7. R.C. Duncan, *Astrophys. J* **498**, L45 (1998).
8. H.M. Van Horn, *Astrophys. J.* **236**, 899 (1980). C.J. Hansen and D.F. Cioffi, *Astrophys. J.* **238**, 740 (1980).
9. P.N. McDermott, H.M. Van Horn and C.J. Hansen, *Astrophys. J.* **325** 725 (1988).
10. T.E. Strohmayer, *Astrophys. J.* **372**, 573 (1991).
11. O. Blaes, R. Blandford, P. Madau and S. Koonin *Astrophys. J.* **363**, 612 (1990).
12. S.I. Bastrukov and D.V. Podgainy, *Physica* **A 250**, 345 (1998), Sect.1.
13. S.I. Bastrukov, F. Weber and D.V. Podgainy, *J. Phys.* **G 25**, 107 (1999).
14. S.I. Bastrukov, I.V. Molodtsova, D.V. Podgainy, F. Weber and V.V. Papoyan, *Phys. Part. Nucl.* **30**, 436 (1999).
15. S.I. Bastrukov and D.V. Podgainy, *Phys. Rev.* **E 54**, 4465 (1996).
16. S. Bastrukov, J. Yang, D. Podgainy and F. Weber, in *Explosive Phenomena in Astrophysical Compact Objects*, eds. H. C. Chang, C. H. Lee, M. Rho, I. Yi, AIP Conf. Proc. 556., (Melville, New York, 2001) p.197.
17. S. Bastrukov, J. Yang, D. Podgainy and F. Weber, *Mem. Soc. Astron. It.* 2002, in print.
18. K. Hirotani, *MNRAS* **317**, 225 (2000); *Astrophys. J.* **549**, 495 (2001).

List of Participants

Sang-Hyeon Ahn	Seoul, Korea	sha@astro.snu.ac.kr
Kiehunn Bach	Yonsei, Korea	khback@galaxy.yonsei.ac.kr
Chang Hyun Baek	Pusan, Korea	chbaek@comet.es.pusan.ac.kr
J. M. Bai	Seoul, Korea	jmmbai@astro.snu.ac.kr
S. I. Bastrukov	JINR, Dubna, Russia	bast@mm.ewha.ac.kr
A. Beloborodov	Sweden	andrei@astro.su.se
R. Blandford	Caltech, USA	rdb@caltech.edu
A. Celotti	Trieste, Italy	celotti@sissa.it
Heon-Young Chang	KIAS, Korea	hyc@ns.kias.re.kr
Chul-Sung Choi	KAO, Korea	cschoi@kao.re.kr
T. Dotani	ISAS, Japan	dotani@astro.isas.ac.jp
L. Ferrarese	Rutgers, USA	lff@physics.rutgers.edu
Andre Fletcher	KAO, Korea	abfletch@kao.re.kr
Byoung-il Han	Seoul, Korea	hanbi@astro.snu.ac.kr
F. He	Ewha, Korea	fhedoc@netease.com
Kouichi Hirotani	NASA/GSFC, USA	hirotani@milkyway.gsfc.nasa.gov
P. A. Höflich	Austin, Texas, USA	pah@astro.as.utexas.edu
Ho-seong Hwang	Seoul, Korea	galllaxy@astro.snu.ac.kr
Siek Hyung	KAO, Korea	hyung@kao.re.kr
Yu-Kyung Jung	Kyungbook, Korea	elara1020@hotmail.com
Hyesung Kang	Pusan, Korea	kang@uju.es.pusan.ac.kr
Nozomu Kawakatsu	Tsukuba, Japan	kawakatu@rccp.tsukuba.ac.jp
Chung Wook Kim	KIAS, Korea	cwkim@kias.re.kr
D. H. Kim	Ajou, Korea	
Hee Il Kim	KIAS, Korea	khi@alpha.kias.re.kr
Hui Kyung Kim	Hanyang, Korea	tarzan@theory.hanyang.ac.kr
Sang Chul Kim	Seoul, Korea	sckim@astro.snu.ac.kr
Seong-Jin Kim	Seoul, Korea	kim4230@astro.snu.ac.kr
Soon-Wook Kim	Seoul, Korea	skim@astro6.snu.ac.kr
J. H. Krolik	Johns Hopkins, USA	jhk@pha.jhu.edu
Chang-Hwan Lee	KIAS, Korea	chlee@kias.re.kr
Hee-Won Lee	Sejong, Korea	hwlee@kunja.sejong.ac.kr
Hyun Kyu Lee	Hanyang, Korea	hklee@hepth.hanyang.ac.kr
Hyung Mok Lee	Seoul, Korea	hmlee@astro.snu.ac.kr
Joong Jea Lee	Korea	
Seong-Jae Lee	Chungnam, Korea	seong@canopus.chungnam.ac.kr
Tae Hoon Lee	Soongsil, Korea	thlee@physics.soongsil.ac.kr
Keiichi Maeda	Tokyo, Japan	maeda@astron.s.u-tokyo.ac.jp
K. Menou	Princeton, USA	kristen@astro.princeton.edu
Shin Mineshige	Yukawa I., Japan	minesige@yukawa.kyoto-u.ac.jp
Sang Hoon Oh	Seoul, Korea	kori94@plaza1.snu.ac.kr
J.-H. Park	Hanyang, Korea	
M.-G. Park	Kyungpook, Korea	mgp@knu.ac.kr

S. J. Park	KAO, Korea	sjpark@kao.re.kr
M. van Putten	MIT, USA	mvp@schauder.mit.edu
J.-H. Rho		
D.-G. Rho	Yonsei, Korea	rrdong9@galaxy.yonsei.ac.kr
Jung-hi Shin	KAIST, Korea	
M. Sikora	Warsaw, Poland	sikora@camk.edu.pl
J. H. Sung		
Rock-Hee Sung	Kyungbook Nat'l U.	onggy3@honmail.net
E. T. Vishniac	Johns Hopkins, USA	ethan@pha.jhu.edu
J. Yang	Ewha, Korea	jyang@ewha.ac.kr
JaiYul Yoo	Seoul, Korea	jyyu@astro.snu.ac.kr
Il-sang Yoon	Seoul, Korea	omnia76@snu.ac.kr
Suk-Jin Yoon	Yonsei, Korea	sjyoon@csa.yonsei.ac.kr
Kwang Yun Youm	Hanyang, Korea	

www.worldscientific.com
4895 hc

ISBN 981-02-4870-9
9 789810 248703