X-Ray Spectroscopy and Atomic Data

Ehud Behar and Steven M. Kahn

Columbia Astrophysics Laboratory, Columbia University, New York, NY

ABSTRACT

The Laboratory Astrophysics program employing the Lawrence Livermore National Laboratory (LLNL) Electron Beam Ion Trap (EBIT) has been providing useful atomic data in support of the x-ray missions *Chandra* and *XMM-Newton*. Major achievements have been made for Fe-L ions in hot, collisional plasmas, relevant to stellar coronae, supernova remnants, elliptical galaxies, and galaxy clusters. Measurements for L-shell ions of other cosmically important elements are also required, some of which are in the LLNL EBIT pipeline. On the other hand, data for inner-shell excited lines relevant to photoionized plasmas near accretion sources are largely lacking. Even the wavelengths of these lines are only poorly known, which severely limits their use for diagnostics, despite the great potential.

1.   Introduction

With the advent of the high-resolution grating spectrometers on board the *Chandra* and *XMM-Newton* x-ray observatories, spectroscopy has taken central stage in the rejuvenated field of X-Ray Astronomy. The increasing role of spectroscopy as a tool for astrophysical measurements has naturally drawn attention to the relevant atomic physics and atomic data, as those are directly linked with the ability to draw meaningful conclusions from observed spectra. The x-ray wavelength band between 1 and 100 Å, which is covered complementally by the *Chandra* and *XMM-Newton* spectrometers, contains a rich forest of spectral lines emitted by highly charged ions that form at electron temperatures of $kT_e = 0.1$ - 3 keV. At these high temperatures, most of the cosmically abundant elements from C to Ni are stripped down to their K shell (i.e., $n = 1$, n being the principal quantum number) and emit relatively few spectral lines. Additionally, many strong lines of L-shell ($n = 2$) ions of Si to Ni fall in this wavelength range. The eight ionization stages in the L shell can provide more precise information on the temperature structure of the source, independent of elemental abundances, than the two K-shell ionization stages are capable of.

2.   The Central Role of Fe-L

The most prominent L-shell x-ray lines in astrophysical spectra are those of iron. Fe-L line emission has been used recently to probe the hot temperature structure in stellar coronal sources (Brinkman et al. 2001), supernova remnants (van der Heyden et al. 2002) and galaxy clusters (Peterson et al. 2001). In general, transitions for both K- and L- shell ions can be calculated with available atomic codes. Nevertheless, in the past, the Fe L-shell lines have been considered highly uncertain, presumably because of the increasing complexity of multi-electron atoms.

To address these issues, an x-ray spectroscopy laboratory-astrophysics program was initiated 11 years ago and is still active today (PI's Kahn and Beiersdorfer). The program is built around the unique capabilities of the LLNL EBIT to measure electron ion interactions. Driven by the astrophysical motivation, these efforts have, so far, focused primarily on the Fe L-shell complex. In particular, high-precision measurements for wavelengths and collisional excitation cross-sections (relative and absolute) have been published. Also, peculiar line ratios of Fe^{16+} that have been puzzling astronomers for years have been investigated and much of the atomic uncertainty has been disentangled from the real astrophysical issues. For more details see the contribution of Brown et al. in these proceedings.

As the first astrophysical grating spectra became available, our team has made a systematic attempt to test the ability of existing models to reproduce the observed emission line intensities. It was found (Behar, Cottam, & Kahn 2001a; Brinkman et al. 2001) that, generally, the observed line intensities could be fairly well reproduced by state-of-the-art distorted wave calculations. For instance, the HULLAC code (Bar-Shalom, Klapisch, & Oreg 2001) was used in those works. In particular, the Fe-L line intensities in the model including all of the high-n lines were found to fare quite well, implying that the calculated excitation rates and ensuing radiative cascades are fairly adequate. Conversely, accurate wavelengths still needed to be incorporated from laboratory measurements. A more detailed confrontation of the atomic calculations with Fe-L spectra of stellar coronae can be found in Behar et al. (2001b), where a comparison of calculations with the spectra of bright stellar coronal sources such as Capella and HR 1099 confirms that, just like the K-shell atomic data, the Fe L-shell data, when calculated correctly, are highly reliable and therefore very useful. The latest versions of the widely used databases MEKAL (Mewe, Kaastra, & Liedahl 1995) and APEC (Smith et al. 2001) now incorporate similar HULLAC data calculated by D. Liedahl with an earlier version of the code.

3.   Remaining Atomic Data Issues

Although in general state-of-the-art models are doing well, several discrete, but nonetheless important, discrepancies still remain. In particular, the ratios of the 3s - 2p line intensities relative to those of the 3d - 2p transitions of the same charge state were found to be anomalously high for both Fe^{16+} and Fe^{17+}. Similar effects have also been found in other late-type stars, and even in elliptical galaxies (Xu et al. 2002), which suggests that they are not associated with the astrophysical conditions. The origin of this discrepancy has been recently studied in many theoretical and experimental works (Laming et al. 2001; Doron & Behar 2002; Beiersdorfer et al. 2002), not all in agreement with each other. Additionally, the L-shells of other cosmically abundant elements remain largely unexplored. Major efforts in this direction are being conducted by Lepson et al. (see these proceedings).

This is the place to note that the atomic processes associated with the x-ray emission depend on the type of plasma at the source. So far, we have focused on hot, collisional plasmas that are governed by electron impact ionization and excitation. This type of plasma is relevant to x-ray observations of stellar coronae, supernova remnants, the hot ISM in old galaxies, the intergalactic haloes of galaxy clusters, and to the overwhelming majority of laboratory plasmas. On the other hand, x-ray sources that are ionized and excited by an external radiation field, e.g., active galactic nuclei (AGN) and x-ray binaries, require atomic data for photon impact processes. Modeling line emission from photoionized sources involves radiative recombination and photoexcitation rates (i.e., oscillator strengths). In a recent paper on the x-ray spectrum of the type II AGN NGC 1068 (Kikhabwala et al. 2002), we have shown that the available data for the K-shell ions are very good for reproducing the x-ray spectra from photoionized sources. Work is in progress to test the status of the Fe-L data for photoionized sources. Similar data are routinely used to model absorption spectra. In absorption, however, one also observes inner-shell transitions, for which until recently there was an enormous lack of atomic data.

4.   The Urgent Need for Inner-shell Absorption Measurements

The grating spectrometers on board *Chandra* and *XMM-Newton* enable us for the first time to detect x-ray absorption lines due to inner-shell photoexcitation. Since gas under almost any conditions absorbs x-rays, these lines are ubiquitous to x-ray absorption observations. Detections of inner-shell absorption lines have been reported mostly in the ionized outflows of AGN, but also for absorption by neutral ISM. Inner-shell absorption can probe the entire range of ionization states from neutral up to highly-ionized Li-like species. Consequently, these lines impose unprecedented, strict, constraints on the ionization structure

in the absorbing medium. Since there were extremely few relevant atomic data for these features in the literature, in order to analyze the spectra, we had to embark on a tedious endeavor of calculating numerous lines (Behar, Sako, & Kahn 2001c; Behar & Netzer 2002). Calculations by the Ohio State team have also contributed to this effort (Pradhan 2000; Nahar, Pradhan, & Zhang 2001; Pradhan et al. 2002). None of these calculations have been benchmarked in the laboratory.

One case where x-ray absorption lines (including inner-shell excited) are particularly useful is for measuring velocities in AGN outflows (e.g., Kaspi et al. 2001). For these measurements, the rest frame wavelengths of the lines need to be known to very high accuracies. The case of inner-shell Kα absorption by oxygen is particularly interesting because it could potentially relate the traditional x-ray absorber (O^{6+} and O^{7+}) with absorbers of other wavebands (e.g., O^{5+} in the UV). Whether these absorbers represent the same kinematic systems or not has been debated in the AGN community for some time now. The correct diagnostics of the inner-shell absorption lines could potentially provide a conclusive answer to this interesting astrophysical question. In order to demonstrate the large uncertainties of the currently available atomic data, in Table 1 we present three different calculations for the wavelengths of the strongest Kα inner-shell absorption lines of O^{1+} through O^{5+} and also compare them with the deduced wavelengths from the *Chandra*/LETGS observation of NGC 5548, courtesy of J. Kaastra. The wavelengths from NGC 5548 were obtained assuming that the outflow velocities of all ions are similar to that of the well-calibrated O^{6+} velocity. Actually, this assumption may not be valid, but with the lack of laboratory measurements, it provides a rough idea of where to expect these lines.

Although one might have expected the R-matrix method to be the most rigorous, it is clear from the table that at the current state of the atomic data, it is virtually impossible to determine what are the correct rest frame positions of these lines. The discrepancies among the various methods reach 50 mÅ, which corresponds here to uncertainties in the measured velocities of 700 km/s. This uncertainty is of the same order as typical outflow velocities in nearby active galaxies, implying that these lines are practically useless for this purpose, despite their great potential. Laboratory measurements are desperately needed.

5. Suggested Z Pinch Measurements

The currently best available facility for producing inner-shell absorption lines and measuring their wavelengths and optical depths is the z pinch at Sandia National Laboratory. The powerful z pinch experiments (x-ray fluxes reaching 10^{19} erg s^{-1} cm^{-2}) produce long-lived (6 ns), steady- state, photoionized gas (Bailey et al. 2001), which can be studied with

high resolution spectrometers, both in emission and in absorption simultaneously. The ionizing spectrum can be characterized rather accurately by a blackbody spectrum. The control over the position and density of the targets in these experiments provides a sensitive handle on the ionization state and column density in the absorbing medium. A demonstration of these capabilities was given by Bailey et al. (2001), where absorption by photoionized Ne has been measured. The spectrum obtained with a crystal spectrometer in that experiment shows many individual lines that are nicely resolved, which allows for accurate wavelength and equivalent-width measurements.

6.   Summary and Prospects

Many atomic data needs for hot, collisional plasmas relevant to x-ray spectroscopic observations, now regularly obtained with the gratings on board *Chandra* and *XMM-Newton*, have been provided by the ongoing LLNL EBIT programs of Kahn and Beiersdorfer. Particularly, these programs have provided the most important data for Fe-L emission by hot gas and work is in progress to measure many more high-quality data for other L-shell systems. Now that we actually have high-resolution cosmic spectra, we can determine better than before which atomic data are the most crucial. Thus, the most urgent needs of the x-ray astronomy community for collisional plasmas will continue to be addressed with the EBIT Laboratory Astrophysics Program. Photoionized gases have received less attention as they are more rare in nature and very few laboratory experiments have sufficient x-ray flux to produce them. Particularly missing are measurements for inner-shell absorption lines. The current uncertainties in the positions of these lines considerably limit our ability to use them for diagnostics. In the future, we intend to try to use z pinch experiments to remedy this deficiency.

REFERENCES

Bailey, J., et al. 2001, J. Quant. Spectr. Radiat. Transfer, 71, 157

Bar-Shalom, A., Klapisch, M., & Oreg, J. 2001, J. Quant. Spectr. Radiat. Transfer, 71, 169

Behar, E., Cottam, J., & Kahn, S.M. 2001a, ApJ, 548, 966

Behar, E., Cottam, J., Peterson, J., Sako, M., Kahn, S.M., Bar-Shalom, A., Klapisch, M., & Brinkman A.C. 2001b in *X-Ray Astronomy 2000* ASP Conference Series Vol. 234, ed. R. Giaconni, S. Serio, & L. Stella (ASP, San Francisco), page 85

Behar, E., Sako, M., & Kahn, S.M. 2001c, ApJ, 563, 497

Behar, E. & Netzer, H. 2002, ApJ, 570, 165

Beiersdorfer, P., et al. 2002, ApJ Letters, submitted

Brinkman, A.C., et al. 2001, A&A, 365, L324

Doron, R., & Behar, E. 2002, ApJ, 574 July 20 issue

Kaspi, S., et al. 2001, ApJ, 554, 216

Kinkhabwala, A., et al. 2002, ApJ, in press (astro-ph/0203290)

Laming, J.M., et al. 2000, ApJ, 545, L161

Mewe, R., Kaastra, J.S., & Liedahl, D.A. 1995, Legacy 6, 16

Nahar, S.N., Pradhan, A.K., & Zhang, H.L. 2001 Phys. Rev. A, 63, 060701(R)

Peterson, J.R., et al. 2001, A&A, 365, L104

Pradhan, A.K. 2000 ApJ, 545, L165

Pradhan, A.K., Chen, G.X., Delahaye, F, Nahar, S.N., & Oelgoetz, J. 2002, astro-ph/0204116

Smith, R.K., Brickhouse, N.S., Liedahl, D.A., & Raymond, J.C. 2001, ApJ, 556, 69

van der Heyden, K., Behar, E., Vink, J., Rasmussen, A.P., Kaastra J.S., Bleeker, J.A.M., & Kahn, S.M., A&A, in press (astro-ph/0203160)

Xu, H., et al. 2002, ApJ, submitted (astro-ph/0110013)

Table 1. Wavelengths in Å for the strongest 1s - 2p inner-shell absorption lines in O ions.

Ion	R-matrix [a]	Cowan's Code [b]	HULLAC [c]	NGC 5548 [d]
O^{1+}	23.27	23.31	23.30	$\cdots$
O^{2+}	23.08	23.10	23.11	23.17 ± 0.01
	23.02	23.08	23.05	23.00 ± 0.02
	22.93	23.01	22.98	$\cdots$
O^{3+}	22.73	22.77	22.73	22.74 ± 0.02
	22.67	22.76	22.78	$\cdots$
	22.67	22.76	22.73	$\cdots$
O^{4+}	22.35	22.38	22.33	22.38 ± 0.01
O^{5+}	22.05	22.05	22.00	22.01 ± 0.01
	21.87	21.85	21.79	$\cdots$

[a]From Pradhan et al. (2002)

[b]Raassen & Kaastra, private communication

[c]Present work

[d]Kaastra, private communication

X-ray astronomy in the new Millenium.
A Summary

By Roger Blandford

Caltech, Pasadena, CA 91125, USA

Recent X-ray observations have had a major impact on topics ranging from proto-stars to cosmology. They have also drawn attention to important and general physical processes that currently limit our understanding of thermal and nonthermal X-ray sources. These include unmeasured atomic astrophysics data (wavelengths, oscillator strengths *etc.*), basic hydromagnetic processes (*e.g.* shock structure, reconnection), plasma processes (such as electron-ion equipartition and heat conduction) and radiative transfer (in disks and accretion columns). Progress on these problems will probably come from integrative studies that draw upon observations, throughout the electromagnetic spectrum, of different classes of source. X-ray observations are also giving a new perspective on astronomical subjects, like the nature of galactic nuclei and the evolution of stellar populations. They are contributing to answering central cosmological questions including the measurement of the matter content of the universe, understanding its overall luminosity density, describing its chemical evolution and locating the first luminous objects. X-ray astronomy has a healthy future with several international space missions under construction and in development.

Keywords: X-rays, accretion, black holes, galaxies, cosmology

1. Introduction

Out of the nearly seventy octaves of electromagnetic spectrum that have been opened up to astronomical observation, X-ray astronomers can lay claim to roughly ten (as opposed to the single octave explored by optical astronomers!). Although no one would pretend that all octaves are equally interesting in terms of physics, there are some special reasons why the X-ray band is peculiarly informative. It includes the K- and L-shell transitions of all the post-big bang elements. It is where to find thermal emission from gas with sound speed $\gtrsim 300$ km s^{-1}, typical of the intergalactic medium, galaxies and stars. It lies right below $\sim m_e c^2$, which is a characteristic energy scale for many nonthermal processes.

X-ray astronomy began forty years ago with the discovery of Sco X-1 [Giaconni *et al.* 1962, Pounds] and was enthusiastically developed surprisingly soon after the dawn of the space age, perhaps because radio astronomy had, by this time, revealed a universe of sources with strength and properties that were completely unexpected on the basis of optical observations. By the time Sco X-1 was identified, quasars and the microwave background had been discovered and pulsars were soon to follow. These four discoveries ushered in modern astronomy.

TEX Paper

As this Discussion Meeting celebrates, X-ray astronomy has come a long way. It is, arguably, ceasing to exist as a separate observational subfield of astronomy, so central have X-ray observations become to the study of essentially all classes of cosmic sources. The most improbable objects – brown dwarfs, all types of protostar and the moon for example – have been detected in X-rays [Güdel]. However, in spite of the fascination of this history, it is to the present and the future that we must turn and earlier speakers have taken stock of where we are after a couple of years of full operation of Chandra and XMM-Newton and what the prospects are for the future. Even this has turned out to be too ambitious to cover in a two day meeting and the papers presented here (and *a fortiori* this brief summary) have had to be quite selective and I shall defer to the other contributors for more representative bibliographies.

As the only non-observer speaking at this meeting, I have organized my commentary around three themes that are somewhat "orthogonal" to the preceding, source-centered talks. These are the physical processes that are ultimately responsible for X-ray emission, the peculiar importance of X-ray observations in rounding out our view of the structure and evolution of stars and galaxies and the under-acknowledged role of X-ray astronomy in defining the cosmological world model to which we have been led in recent years and which is now starting to raise some very important questions concerning what actually happened in the first Gyr of the life of the universe. I conclude with a brief listing of the proposed next generation of X-ray observatories.

2. Physical Processes

The description of many high temperature and nonthermal sources is dependent upon some poorly undestood physical processes. It is striking how often the same questions are asked of quite different physical environments. How important is thermal conduction? How effective is magnetic reconnection in heating plasma? And so on. The optimistic view, that links several of the talks, is that we really only have to solve these problems, *e.g.* the structure of Mach 30 shocks, once and we should be prepared to combine astrophysical, space physical observational with laboratory experiments and computational work to develop some confident answers.

(*a*) *Spectroscopy and its Interpretation*

The gratings on XMM-Newton and Chandra have unprecedented spectroscopic capability, and the ability to utilise them so effectively has benefitted from over thirty years of hard (and largely unacknowledged) work measuring and computing wavelengths, oscillator strengths and so on for transitions of little terrestrial interest. Of particular importance are the Helium-like triplets which combine observations of permitted, intercombination and forbidden transitions (Gabriel & Jordan 1969, [Kahn]) so as to provide density and temperature diagnostics. The low densities that allow forbidden transitions to be so important in cosmic sources are hard to work with experimentally. Conversely, the high radiation densities that allow radiative ionisation equilibria to be established are only just being achieved using powerful lasers.

By now, most of the important wavelengths, oscillator strengths, collision integrals *etc.* for the strongest lines have been measured, although these measurements are still lacking for the majority of the weaker lines that can also be observed in the brightest sources. Observations of H-like and He-like transitions of the more common ions provides useful diagnostics of the density, temperature, abundance, ionization equilibrium and velocity in the emission regions and we have seen here many examples of what can be done in a wide variety of sources including accretion disks (stellar [Done] and AGN [Fabian]), stars and protostars [Güdel], clusters of galaxies [Mushotzky] and polars [Cropper]. The power of X-ray spectroscopy is most clearly brought out by the detailed observations of bright supernova remnants [Canizares]. Here, it is possible to use the ~ 100 km s^{-1} velocity resolution to make three dimensional abundance maps of the expanding debris and forensic analyses of the initial explosions.

Unfortunately, most other sources are unresolved and most direct analyses of the data are often limited to one zone models and some very primitive radiative transfer. Now, it is possible, at least in principle, to include anisotropy, inhomogeneity and peculiar geometric effects in theoretical models of these sources. The problem is that we really have no clear idea of the disposition and structure of the emission region and the medium through which the radiation is propagating. A particularly important example is provided by the ongoing debate concerning the soft X-ray spectra of Seyfert galaxies. Does the power-law, continuum source that is reflected by the disk arise in a local corona or at high altitude so that it can illuminate the whole disk? Are the carbon and oxygen lines produced by reflection (Sako *et al.* 2002) or in a dusty, warm absorber (Lee *et al.* 2001) again at some distance? Why does there appear to be no sign of a reverberative response in Seyfert galaxies with variable X-ray continua? Answering these questions using a more detailed analysis of line formation in the two cases, is tantamount to understanding the source geometry. However, as broad emission lines are now being reported from binary X-ray sources like Cyg X-1 [Fabian] and J1650-500 (Miller *et al.* 2002), it is reasonable to suppose a model that works for Seyferts should also work for black hole binaries.

A second example is provided by stellar coronae, where hot, coronal gas is excited by twisted, magnetic loops. (Temperatures as high as 40 million degrees are now reported associated with the Galactic Cepheid, YY Mon [Güdel].) Although we can image similar activity in the sun, we still do not understand it at all well. Again, we do not have an agreed story as to the sequence of events that leads to coronal lines being emitted and, consequently, how to convert raw line strengths from distant stars into physical conditions in their coronae. The observation of coronal activity from late M stars that are thought to be cool enough not to possess surface convection zones suggests that other, non-magnetic processes could be at work. Again, there is probably a general theory that can be inferred by combining solar and stellar observations.

A third case, where we probably do understand the geometry and can compute the emissivity and opacity, is provided by cyclotron line formation in accreting white dwarfs [Cropper]. Although eclipse observations have confirmed the expected strongly inverted temperature gradient, the radiative transfer is quite subtle and a far more detailed theoretical treatment is likely to be necessary for us to reproduce the observed spectra as well as their polarisation and time-dependence. (Monte

Carlo techniqes will surely continue to play a major role here and some of the experience gained from working with Tokamak plasmas may be relevant.)

(*b*) *The Plasma Impasse*

X-ray sources are, inevitably, fully ionized gases. It is therefore unfortunate that X-ray astronomers have been resistant to learning the principles of plasma physics and incorporating them into their science, preferring instead to limit their purview to gas dynamics and atomic physics. This evasion can no longer be excused. There are now several sources where progress awaits the answers to fundamental plasma physics questions. Again, I only have space for a few examples.

The first question is "How fast do ions with some temperature T heat cooler electrons?". There is a minimal and standard heating rate resulting from two body Coulomb scattering. However, there is an abundance of wave-particle interactions that might, for example, be excited by streams of fast particles. The empirical evidence comes from the observations of high Mach number, heliospheric and supernova remnants shock fronts, which appear to transmit thermal electrons with a temperature well below the equipartition value so that the ions do not quickly attain collisional ionization equilibrium [Canizares]. This suggests that collective effects are not that important. This view is supported by observations of slowly accreting black holes which are also best interpreted in terms of a minimal, Coulomb heating rate. The answer to this question is of direct relevance to the debate about the efficacy of electron heat conduction. There are really two issues. What is the mean free path of the electrons along the magnetic field and how quickly do the field lines wander in response to an imposed turbulence spectrum? As discussed below, our best laboratories are clusters of galaxies.

A related plasma question concerns the efficiency of strong shocks for accelerating cosmic rays. There is a linear theory which appears to have some validity, again based upon heliospheric measurements and X-ray observations of supernova remnants [Canizares]. However, in order to model the observations, we have to understand how the back reaction associated with the cosmic ray pressure moderates the acceleration and what controls the rate of both ion and electron injection. Theoretically, it may soon be possible to perform $3 + 3$ dimensional kinetic simulations with sufficient resolution to address these questions. Observationally, GLAST should provide measurements of the energetically-dominant GeV ions. As relativistic electrons are scattered by the same Alfvénic turbulence that operates on the ions, the combination of γ-ray and X-ray observations should enable us to infer the injection rate and perhaps determine the scaling with Mach number.

Another, quite controversial feature of shock fronts is their role in amplifying magnetic field. Theoretically, there is no very good reason why simple gas dynamical shocks should do any more than compress the pre-shock magnetic field. In addition, radio observations of many (though not all) supernova remnants seem to show that the emissivity and, presumably, the magnetic field strength only increase in the interaction region between the shocked ejecta and the swept up interstellar medium. Conversely, it is possible that the hydromagnetic turbulence that is invoked to scatter the cosmic rays leads to an overall increase in the rms magnetic field strength. This is highly relevant to the late-time evolution of γ-ray burst afterglows.

Even greater uncertainty surrounds our understanding of relativistic shocks,

which are throught to be the primary acceleration site for X-ray-emitting relativistic electrons and magnetic field amplification in γ-ray bursts, pulsar wind nebulae and extragalactic jets. However, the diffusive mechanism for particle acceleration that operates nonrelativistically is kinematically precluded. There is a promising relativistic variant (Achterberg *et al.* 2001). However, this assumes that the cosmic rays can move far enough upstream from the shock to scatter off the background flow and it is not clear how this can happen if there is an oblique magnetic field. Neither is it clear how the scatterers can be generated. The magnetic field itself is also believed to be strongly magnified at the shock front, though no good explanation of how this happens has been found. Indeed, the very existence of sudden, collisionless discontinuities, as opposed to a slow sharing of momentum between two fluids, has been questioned. X-ray observations should be especially instructive because they permit us to resolve these putative shocks, say in the Crab Nebula and jets like M87, at energies where the emitting, relativistic electrons quickly cool. The observation of what is presumably X-ray synchrotron radiation well away from the supposed strong shocks implies that relativistic electrons have to be accelerated *in situ*, rather than at strong shocks. These observations further raise the possibility, discussed elsewhere, (Blandford 2002) that the observed sources including their "shocks" are actually relativistic, electromagnetic structures and are not well-described by gas dynamics.

Another general process is magnetic reconnection which has been invoked, for example, in explaining the energisation of accretion disk coronae. However, the manner in which it operates remains quite controversial. Most existing discussions (*e.g.* Priest & Forbes 2000) have been essentially hydromagnetic except within a small region where the magnitude of the field gradient becomes very large and where a scalar (and usually "anomalous") resistivity is invoked. A recent development, which has serious implications for the topological behaviour, is that the resistivity might be dominated by non-dissipative, Hall terms (Bhattarcharjee, Ma & Wang 2001). These embellishments of MHD are now finding their way into numerical simulations and it will be interesting to see what are their implications for X-ray sources.

In addition to numerical simulation and *in situ* observation of space plasmas, it is becoming posssible to address some of these questions using the growing field of laboratory experimentation. It is now possible to create relativistic plasmas - both ionic and pair plasmas - using powerful lasers, electron beams and magnetic pinches. Temperatures as high as 100 MeV, energy fluxes of ~ 100 ZW m^{-2} and ~ 1 MT magnetic field strengths are all attainable. "High energy density" investigations are likely to become much more versatile in the coming years (*e.g.* Takabe 2001).

(c) *Black Hole Accretion*

The problem of accretion onto a compact object, specifically a black hole, is generally well-posed but has also not had a confident solution over the past thirty years. However, through a combination of theoretical arguments and direct observation of accreting sources, it has been possible to make a lot of progress recently. The greatest excitement has probably centred around the occasional observation of broad iron lines from selected, low luminosity AGN and, as reported here, a couple of Galactic binary X-ray sources [Fabian, Done]. (We now know that broad

lines are not seen as commonly as once thought and that their formation must be more complicated than envisaged in early models. The prospects for performing useful "reverberation mapping" do not look good.) In those sources, where these features are undoubtedly seen, we can say that there is evidence that the second parameter that characterizes a classical black hole – the spin – is responsible for the line width. Indeed, it has even been argued that the role of the black hole is not just the passive one of allowing stable orbits from which highly redshifted photons can be observed, but is an active one in which a magnetic connection of the gas to the spinning hole leads to an enhanced emissivity from the innermost, and most redshifted orbits (Wilms *et al.* 2002).

Another very promising line of investigation is epitomized by the observations of Sgr A* which show that the black hole is a strikingly underluminous X-ray source with an apparent luminosity $\sim 10^{-8}$ L_{Edd} and a radiative efficiency of $\sim 10^{-7}c^2$ relative to the inferred mass accretion rate of $\sim 10^{22}$ g s^{-1}. There have been several explanations put forward, but most of these require that the rate of electron-ion equilibration be slow, as discussed above. It no longer seems possible that all of the mass supplied can accrete onto the hole and either most of the mass is lost (Blandford & Begelman 1999) or the accretion backs up to the Bondi radius at $\sim 10^7$ gravitational radii. The whole matter has been made more interesting through the discovery of surprisingly rapid X-ray variability in Sgr A* (Baganoff *et al.* 2001) and the even more remarkable suggestion that the radio variation may be periodic (Zhao, Bowers & Goss 2001). These observations open up many more possibilities and will undoubtedly be quite constraining once the observational situation is clarified.

Galactic black holes provide more immediate gratification for observers than massive black holes, both on account of their larger fluxes and also because of their much more rapid variability timescales [Done]. There is now a fairly convincing, qualitative explanation of the low and high states. The former arise when the luminosity is $\gtrsim 0.03L_{\mathrm{Edd}}$ and a thin (or slender) disk extends down to r_{ms}; the latter when there is a central hole filled by gas that cannot cool and radiate efficiently and where a nonthermal spectrum is created by Comptonisation. It is not clear that all of this gas accretes onto the black hole.

Many of the questions raised by these observations are issues of theoretical principle that are still being debated. The approaches that will be necessary to address these questions are both observational and theoretical. For the former, the angular resolution of Chandra can be put to great advantage resolving the accretion radii in nearby, dormant galactic nuclei. These observations are helping us to define the physical conditions and perhaps to deduce the rate of gas supply to the central black hole. Theoretically, there are opportunities for carrying out 3D numerical fluid dynamical and MHD/electromagnetic (including general relativisitc) simulations of disks and outflows.

(*d*) *Nonthermal Emission*

The capability to perform arcsecond imaging at X-ray wavelengths is revolutionizing our view of nonthermal emission. Surely, the most famous instance of this is the discovery of a pair of axial jets in the Crab Nebula (Weisskopf *et al.* 2000) as well as other Pulsar Wind Nebulae. This was relatively unexpected and shows

that accretion disks are not necessary for "jet" formation. However, it may suggest something even more fundamental and to explain this, I should return to one of the first models of a pulsar, the Goldreich-Julian, axisymmetric model. Here, it was proposed that a spinning, magnetised neutron star acts as a unipolar inductor and generates an EMF, $\mathcal{E} \sim 30$ PV in the case of the Crab pulsar and that this drives a current ~ 300 TA around a quadrupolar circuit. (We now know, thanks to Ulysses, that the heliospheric electrical circuit is of this form although the EMF and current are only ~ 100 MV and ~ 1 GA repectively.)

Now real pulsars are, by definition, non-axisymmetric and the electromagnetic field just beyond the light cylinder will contain both an AC and a DC component. The interaction between these two components is unclear, but it has generally been assumed that essentially all of the electromagnetic Poynting flux is quickly converted into the kinetic energy of a plasma-dominated, outflowing wind. In other words, the electrical circuit is completed quite close to the pulsar. In this case, a hypersonic wind is created which, it has been supposed, passes through a strong shock front with Lorentz factor $\sim 10^6$ close to the famous "wisps", where particle acceleration and nonthermal emission can occur. However, the X-ray image really shows no evidence for this shock front except perhaps along the poles and the equator. The moving features that are seen optically also appear to be confined to the equatorial plane.

These observations suggest a different interpretation (Blandford 2002), specifically that the AC electromagnetic component dies away very quickly, perhaps non-dissipatively while the DC component persists all the way into the nebula. In this case, the X-ray emission that is observed largely delineates the quadrupolar current flow. More specifically, there is no strong, reverse shock and relatively little of the current completes close to the pulsar. The emission that is seen may well reflect MHD instability in the magnetic configuration - pinches and current sheets are notoriously unstable - and be a manifestation of ohmic dissipation. In this case, the observed Crab Nebula would be magnetically-, rather than particle-dominated except, perhaps, in the emission region where the relativistic electron energy density might build up to an equipartition value. These ideas should be testable by examining the spectral index gradients and the polarization map.

This viewpoint has implications for ultrarelativistic jets and gamma ray bursts (GRB), which are now also widely acknowledged to have an electromagnetic origin but where it is also supposed that the Poynting flux is quickly converted to fluid form (most commonly as an optically thin pair plasma in the former case and as a radiation-dominated fireball in the latter) and that the ultimate emitters are strong, relativistic shock fronts. By contrast, under the electromagnetic hypothesis, it is supposed that the energy released remains in an electromagnetic form all the way to the emission region and that the particle acceleration is a direct result of wave turbulence. There should be ample potential difference available for particle acceleration to take place. In the case of a quasar jet, the EMF is ~ 100 EV and the current is ~ 1 EA. (For GRBs the estimates are now ~ 10 ZV and ~ 100 EA, rather lower than in the past.)

3. Astronomical Questions

(a) *Nuclear Power*

High angular resolution, X-ray observations have, in an almost literal sense, transformed our view of active galactic nuclei (AGN). They have amply confirmed the finding from optical observations, that our classification of these sources is strongly aspect-dependent. The simple geometrical model of an AGN invokes a thick torus that will absorb UV continuum and emission lines and all but the hardest X-rays and γ-rays and then re-emit the absorbed energy in the thermal infrared. This description has received impressive confirmation with ASCA, XMM and Chandra measurements of hard X-ray spectra [Matt] which clearly exhibit the effects of absorption with hydrogen column densities that can approach $\sim 10^{24}$ cm^{-2}. This, in turn, implies that a significant fraction of the infrared background, as well as the bolometric luminosity density of the young universe, be associated with AGN. In round numbers, the energy density associated with the observed X-ray backgound (mostly in the energy interval $\sim 20 - 40$ keV) is 3×10^{-17} erg cm^{-3}, a fraction ~ 0.003 of the energy density measured in both the optical and in the far infrared backgrounds (and $\sim 10^{-4}$ of the microwave background energy density). In other words, if the intrinsic, integrated, UVX power of an AGN is, on average, thirty times the HX power and this is a reasonable guess based upon observations of local AGN, then AGN must account for ~ 10 percent of the infrared background, with the remainder being presumably associated with stars. Estimates in the literature based upon more detailed assumptions about the mean AGN spectrum and the redshift distribution of the sources range from $\sim 3 - 30$ percent. However, this quantity may not be very well-defined because much of the luminosity associated with galactic nuclei could be attributed to starbursts as opposed to accretion onto massive black holes [Ward].

The characterisation of the obscuring material as a torus is problematic, at least from a theoretical viewpoint. The difficulty is that it is very hard to see how a thick, cold ring of molecular gas can be supported. Individual clouds should collide inelastically and a torus would quickly deflate. A more plausible alternative (*eg* Sanders *et al.* 1989) is that there is a locally thin, though strongly warped disk, where the thickness is maintained through the marginal growth of gravitational instabilities. However there is still confusion about the size of this disk. Combined X-ray and SIRTF observations should greatly improve our understanding of AGN obscuration.

Obscuration is not the only way to produce beaming. The X-ray study of rela-tivistic jets, especially in blazars and nearby sources like M87 (which would probably be classified as a blazar were it pointed at us) is becoming more sophisticated. How-ever we are still not yet able to answer the quantitative questions about beaming fraction, the angular and radial variation of jet Lorentz factors *etc.* . These ques-tions will undoubtedly be a focus of future X-ray research, especially after GLAST is launched.

The luminosity of galactic nuclei is not exclusively radiative. There are "super-winds" which are driven by nuclear activity [Ward]. In addition, roughly ten percent of optically-selected quasars exhibit (X-ray quiet) broad absorption line outflows. Even if these flows only represent a minor fraction of the nuclear power budget, they can still have a large impact on the host galaxies because the momentum flux

of a wind scales inversely with the outflow speed. Not only can these outflows drive away much of the interstellar medium, they may even influence the overall galaxy morphology by, for example, inhibiting or limiting disk formation.

(*b*) *Stellar Populations*

X-ray observations are also presenting us with a complementary perspective to that obtained using optical and infrared observations on how the stellar populations of galaxies evolve. This is because they allow us to witness the endpoints of stellar evolution as opposed to the beginnings. So far most attention has been on nearby galaxies, both spirals and ellipticals [Ward]. The primary targets are young supernova remnants and binary X-ray soures. Comparison with radio and infrared observations is particularly important in the former case as it allows us to derive some useful empirical correlations. As an example of what can be learned in this way, it has been reported that the star formation rate declines less rapidly than the supernova rate [Ward].

Probably the biggest new discovery is the ultraluminous X-ray sources which are now showing up quite regularly in all types of galaxy [Ward]. They are defined by having X-ray luminosity in excess of the Eddington limit for a ~ 30 $M_\odot$ black hole. They have been associated with intermediate mass black holes, conceivably relics from the Population III era. In this case the fuel supply is a bit of a puzzle. They must either be short-lived binaries or single stars moving slowly through molecular clouds and, consequently, a major constituent of the universe, overall. Alternatively, they could be the long-sought stellar blazars, although here the absence of a radio emission is surprising. A third possibility is that they are normal mass black hole binaries with super-Eddington luminosity which may be physically possible in the highly clumped, radiation-dominated fluid that is expected in the innermost regions of accretion disks orbiting ~ 10 $M_\odot$ black holes, is responsible. Again, this phase would have to be quite short-lived. The heterogeneity of the observed properties of ultraluminous X-ray sources suggests that more than one of these explanations could be correct.

4. Cosmological Issues

(*a*) *Clusters*

X-ray observations continue to be of central importance to the study of rich clusters of galaxies [Allen, Mushotzky]. The most basic information concerns the shape and depth of the gravitational potential well. The various methods that have been used to determine this now seem to be in fairly good agreement. Arguably the most reliable is gravitational lensing - strong in the core and weak in the outer parts - especially when there are reliable source redshifts. As demonstrated here, this technique works well for clusters that appear to be nearly circular and which are dynamically relaxed. The total mass density can then be derived from Poisson's equation. (The smoothness of the known arcs assures us that the potential is also quite smooth.) It is then possible to use imaging spectroscopy to measure the baryon mass distribution, as roughly 85 percent of this is believed to be in the form of hot gas. (In practice this is carried out by model-fitting rather than an unbiased inversion.) This procedure may be problematic near the center of the cluster but

is far safer in the outer parts of the cluster where most of the mass resides and where we are most likely to sample fairly the cosmological mass distribution. If we are prepared to trust the baryon density derived from the measured deuterium (plus other light elements) abundance and the theory of big bang nucleosynthesis ($\Omega_b = 0.04$), then we can deduce the contemporary matter fraction of the universe and a value $\Omega_0 = 0.31 \pm 0.03$ was quoted (assuming a Hubble constant $H_0 = 70$ km s^{-1} Mpc^{-1}) [Allen]. This argument, which preceded the more publicised type Ia supernova determination and which is less subject to systematic error, appears to be holding up very well. Repeating these measurements with larger redshift clusters, probably in conjunction with Sunyaev-Zel'dovich measurements, should allow us to infer the expansion history of the universe, at least in principle. To date, this translates into an unsurprising bound on Ω_Λ.

It is also possible to explore the thermal history of the intergalactic gas by comparing the measured mass – temperature relation with the expectations of adiabatic simulations. It should not be too surprising that the agreement here is less good [Mushotzky, Allen]. The entropy history of the gas which affects this determination is likely to be quite complicated. The gas temperature, immediately after re-ionization by the first stars and quasars is only ~ 10 km s^{-1}. However as structure grows, this gas will acquire speeds of several hundred km s^{-1}. Strong, large scale shock fronts are likely to form and increase the entropy of the gas. However, as noted above, the post-shock electron temperatures will be significantly lower than the ion temperatures (There is no observational evidence for strong accretion shocks surrounding observed clusters; it appears, quite reasonably, that the gas is heated much earlier in the merger hierarchy.) The gas will also be mixed with cooler gas swept out of galaxies. It is very hard to compute the influence of these and other easily-imagined effects from first principles. The dark matter, mass-velocity relation, which is probably best determined by gravitational lensing, is more likely to furnish a robust measure of the growth of large scale structure.

There is an inescapable implication of the cluster gas having been heated by strong shock fronts, whatever their provenance. This is that the hot gas will be accompanied by cosmic ray ions. The speeds and Mach numbers of the shock fronts are quite similar to those in the interplanetary and interstellar media and we expect that the post-shock, GeV cosmic ray partial pressure will lie somewhere in the range 0.2-0.5 of the total pressure. This fraction will decrease slightly as the gas is adiabatically compressed but if the gas starts to cool appreciably then cosmic rays may dominate the pressure and inhibit further cooling. This may be part of the explanation of the suprising results from X-ray spectroscopy by XMM of a few well-studied clusters [Mushotzky]. These seem to show that the gas starts to cool as it flows towards the central cD galaxy and then appears to vanish – the lines expected from gas with temperature below ~ 2 keV are absent. By resisting further compression, the cosmic rays will make it easier to keep the gas hot. Other factors that have been invoked to explain the failure to observe cooling flows include variable metallicity, thermal conduction and supernova heating. There is a further implication of having these cosmic rays present and this is that they may contribute to the heating and, in particular, may create γ-rays through pion production. The predicted γ-ray flux from nearby clusters should be detectable by GLAST and, under extreme assumptions, could contribute to the γ-ray background.

X-ray observations are also providing new information on the chemical history of

the universe as we try to reconstruct the history of clusters of galaxies [Mushotzky]. C, N, O, S, Fe, Ni have all been measured in a large sample and abundance gradients in a smaller number. The iron abundance has now stabilized at $[Fe/H] \sim 0.3$ (possibly increasing with the size of the cluster) and is consistent with a Type Ia origin. The supernovae may have occured mostly in clusters with the processed gas being driven out in superwinds. The correlations of the relative abundances with the cluster properties as well as the radius are starting to become quite diagnostic of the evolutionary history of the cluster gas.

(b) X-ray Background

Another great success for Chandra and XMM-Newton has been the resolution of $\sim 80 - 90$ percent of the X-ray background into $\sim 3 \times 10^8$ discrete sources [Brandt, Hasinger]. As discussed above, most of the energy density appears to derive from black hole accretion in low power AGN. The redshift distribution of these sources is controversial. On the one hand the $\sim 10^8$ sources which contribute most of the background appear to have modest redshifts $z < 1$ [Hasinger]; on the other, the faintest sources are seen out to $z \sim 3$ [Brandt]. In addition, the surveys are so sensitive that nearby normal galaxies and luminous quasars with $z \gtrsim 5$ are also found to be minor contributors to the background. I suspect that these statements are approximately true and not in contradiction.

The notion that most of these X-ray sources are obscured receives support from their association with ISO sources and bodes well for SIRTF observations. If, following the example above, we suppose that AGN account for ~ 10 percent of the infrared background then the energy produced, allowing for the expansion of the universe, corresponds to $\sim 10^7$ $M_\odot c^2$ per *local L^** galaxy. If we assume that black holes grow with ~ 10 percent radiative efficiency, then we deduce a mean black hole mass per L^* galaxy of $\sim 10^8$ $M_\odot$, roughly compatible with what is measured. In addition, as there are $\sim 3 \times 10^9$ of these locally-specified L^* galaxies out to $z \sim 2$, where most of the black hole mass is grown, we conclude that a typical nucleus of one of these galaxies is active for $\sim 10^8$ yr, consistent with the Salpeter time.

A possible problem with this neat explanation is that if the spectrum below an observed energy ~ 30 keV really does come from obscured sources, then they must have low redshifts as it is hard to see how an absorption turnover could occur at a much higher energy than ~ 40 keV when scattering dominates any plausible opacity. If this is true, then, when Swift or EXIST identify the sources that contribute most of the hard X-ray background, they should find that they have low redshifts. An alternative possibility is that the hard X-ray sources are mostly at $z \sim 2 - 3$ and we are observing hard, Comptonised spectra in a corona with temperature ~ 100 keV. Ultimately, we should like to extend this analysis all the way up to γ-ray energies and make smooth contact with studies of the γ-ray background.

(c) First Light

The famous penetrating power of hard X-rays makes them excellent candidates for probing the very early universe. Recent observations are encouraging. Quasars have been found with $z \sim 6.5$ and the X-ray powers suggest that the holes exceed a billion solar masses. This is a constraint on theories of the growth of black holes

and a timely reminder that quasar activity must be intimately related to galaxy formation and evolution during the ~ 0.5 Gyr between reionisation and $z \sim 6.5$ when the first quasars are seen.

In addition, GRB have already been seen out to $z \sim 4.5$, where the redshift actually helps by allowing an observer with a fixed response time to observe an earlier and brighter part of the evolution, at greater emitted photon energy. There is optimism that Swift will identify X-ray afterglows emitted earlier than the light from the first observed quasars.

5. Future Missions

Although we look forward to many years of active service from Chandra and XMM-Newton as well as the upcoming launches of INTEGRAL (2002), Swift (2003), AS-TRO E-2 (2005) and GLAST (2006), there are also longer range plans to construct more powerful telescopes like Constellation-X, EXIST, LOBSTER, Generation-X and XEUS. The problems discussed at this meeting are already rewriting the scientific case for the longer term missions and it is hoped that this will be reflected in further improvements in mission design and optimal use of over-subscribed international resources for space astronomy.

References

Achterberg, A. A., Gallant, Y. A., Kirk, J. G. & Guthmann, A. W. 2001 MNRAS 328 393

Balbus, S. A. & Hawley, J. F. 1998 RMP 70 1

Baganoff, F. K. *et al.* 2001 Nature 413 45

Bhattarcharjee, A., Ma, Z. W. & Wang, X. 2001 Phys. Plasmas 8 1829

Blandford, R. D. 2002 in Lighthouses of the Universe ed. R. Sunyaev *et al.* Berlin:Springer in press

Blandford, R. D. & Begelman, M. C. 1999 MNRAS 303 L1

Esin, A. *et al.* 2001 ApJ 555 483

Gabriel, A. H. & Jordan, C. 1969 MNRAS 145 241

Giacconi, R. *et al.* 1962 PRL 9 439

Lee, J. C. *et al.* 2001 ApJ 554 L13

Miller, J. M. *et al.* 2002 ApJ (submitted)

Priest, E. R. & Forbes, T. G. 2000 Magnetic Reconnection Cambridge: Cambridge University Press

Sako, M. *et al.* 2002 ApJ (submitted)

Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T. & Matthews, K. 1989 ApJ 347 29

Takabe, H. 2001 Prog. Theo. Phys. 143 202

Weisskopf, M. C. *et al.* 2000 ApJ 536 L81

Wilms, J. *et al.* 2001 MNRAS 328 L27

Zhao, J. H., Bower, G. C. & Goss, W. M. 2001 ApJ 547 29

THE FORMAL UNDERPINNINGS OF THE RESPONSE FUNCTIONS USED IN X-RAY SPECTRAL ANALYSIS

JOHN E. DAVIS

Center for Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139; davis@space.mit.edu

ABSTRACT

This work provides an in-depth mathematical description of the response functions that are used for spatial and spectral analysis of X-ray data. The use of such functions is well known to anyone familiar with the analysis of X-ray data where they may be identified with the quantities contained in the ancillary response file (ARF), the redistribution matrix file (RMF), and the exposure map. Starting from first principles, explicit mathematical expressions for these functions, for both imaging and dispersive modes, are arrived at in terms of the underlying instrumental characteristics of the telescope including the effects of pointing motion. The response functions are presented in the context of integral equations relating the expected detector count rate to the source spectrum incident upon the telescope. Their application to the analysis of several source distributions is considered. These include multiple, possibly overlapping, spectrally distinct point sources, as well as extended sources. Assumptions and limitations behind the usage of these functions, as well as their practical computation, are addressed.

Subject headings: methods: analytical — methods: data analysis — X-rays: general

1. INTRODUCTION

It is a basic tenet of X-ray spectral analysis that the source flux incident at the telescope is related to the observed count rate through an integral equation involving the effective area of the telescope. The most commonly accepted technique for dealing with this equation involves the use of spectral analysis programs such as *xspec* (Arnaud 1996). The effective area is input into these programs via a file called the ancillary response file (ARF). In addition, the energy resolution of the detector is specified by the redistribution matrix file (RMF). This work presents formal descriptions of the quantities embodied by the ARF and RMF in terms of the underlying instrumental responses, making a clear connection between the incident source flux and the observed count rate.

Roughly speaking, the effective area of an X-ray telescope composed of a mirror and a detector is more or less the product of the effective area of the mirror with the quantum efficiency (QE) of the detector. However, the ARF, which relates observed counts to a source flux, also depends upon observation-dependent quantities such as the detailed aspect history of the telescope, its point-spread function (PSF), and details of the analysis itself, e.g., the filtering and binning of the observed data.

With the advent of the *Chandra X-Ray Observatory* (CXC 2000), all of the subtleties introduced by the shape of the PSF and the telescope aspect motion were deemed important for the computation of an ARF. *Chandra* is well calibrated (Weisskopf et al. 2000), and with its unprecedented combined spectral and spatial response, a precise definition of the ARF that incorporates the effects of spacecraft motion and the PSF is necessary in order to perform spectral analysis at the highest resolution of the instrument.

The main goal of this work is to present an explicit first principles derivation of the ARF by including the proper treatment of telescope motion (e.g., dither) and the PSF, as well as any filtering of the data. The bottom-up approach taken here necessarily implies that a meaningful and consistent derivation can be achieved by considering the role of

the ARF in spectral analysis. As a result, the ARF is presented in the context of integral equations that connect the incident X-ray flux to an expected count rate.

Traditionally, the ARF and RMF have been used primarily for the analysis of spectral image data. An important aspect of this work is to extend this approach to the analysis of dispersive spectral data, such as data obtained by *Chandra* or *Newton*. To this end, definitions of an ARF and an RMF are presented that are suitable for dispersive data analysis and that may be utilized by existing spectral analysis software.

One of the original motivations for this work was the need to create a related object, an exposure map, for use in the analysis of data obtained by the *Chandra X-Ray Observatory*. This paper also gives a rigorous definition of the exposure map and discusses some of its uses and its limitations in spectral image analysis. The resulting definition is consistent with current use, intuition, and physics.

The next section contains a discussion of the general response of an X-ray telescope and also serves to introduce the notation and conventions used throughout this paper. Although originally inspired by the need to create ARFs and exposure maps for *Chandra*, the presentation has been kept as general as possible without focusing on any particular telescope or instrument. A derivation of the imaging ARF follows in § 3, where its application to several problems is considered. These include the problem of multiple overlapping point sources. Section 4 contains a definition of the exposure map and explores its use as well as its limitations in dealing with extended sources. The definition of a dispersive ARF and RMF that are suitable for use in the analysis of dispersive spectral data is given in § 5. Following the summary of the paper is an appendix that considers the practical computation of these objects.

2. TELESCOPE RESPONSE

Let $S(\lambda, \hat{p})d\lambda \, d\hat{p}$ represent the number of photons per unit area per unit time incident upon the telescope with directions that lie in the cone of directions between $\hat{p}$ and $\hat{p} + d\hat{p}$

and whose wavelengths lie between λ and $\lambda + d\lambda$. This source flux is assumed to be time independent.[1] Similarly, let $S_D(h, \sigma, t)d\sigma$ represent the expected number of counts per unit time, in pulse-height bin h, within a region of area $d\sigma$ at the position σ on the detector. (Although in this paper h is a discrete quantity that represents a pulse-height channel, one could easily generalize h to a continuous quantity such as a voltage by introducing the infinitesimal dh and replacing $\sum_h$ by $\int dh$.)

In general, the count rate will be time dependent even if the source flux does not vary with time. This time dependence arises from several effects, including but not limited to telescope pointing motion (e.g., pointing wobble or dither), detector electronics, telemetry saturation, and thermal expansion effects that cause the movement of individual telescope subsystems with respect to one another.

As used here, pointing is the measured, generally periodic movement of a coordinate system attached to the optical axis with respect to a coordinate system fixed with respect to the sky. These two coordinate systems are related by a time-dependent rotation matrix $\mathscr{R}(t)$ that completely characterizes the pointing. It is assumed that complete knowledge of $\mathscr{R}(t)$ is available from, e.g., an on-board aspect system, at some level of accuracy.

In the coordinate system fixed with respect to optical axis, the source flux will appear to be time dependent according to

$$S(\lambda, \hat{p}, t) = S(\lambda, \mathscr{R}^{-1}(t) \cdot \hat{p}) . \tag{1}$$

This equation simply expresses the fact that an observer fixed to the telescope will see a time-dependent source and that this induced time dependence is a direct consequence of telescope motion. In the telescope coordinate system, the source flux and the count rate are related to one another via the equation

$$S_D(h, \sigma, t) = \int d\lambda \int d\hat{p}\, R_T(h, \lambda, \sigma, \hat{p}, t)S(\lambda, \mathscr{R}^{-1}(t) \cdot \hat{p}) , \tag{2}$$

which defines the total response $R_T(h, \lambda, \sigma, \hat{p}, t)$ of the telescope. It is an extremely complicated function that incorporates all elements of the telescope such as the detector and its electronics, the mirror, and diffraction gratings, if present. Since the units of $S_D(h, \sigma, t)$ are counts per unit detector area per unit time, and the units of $S(\lambda, \hat{p})d\lambda$ are photons per unit aperture area per unit time, it follows that the response $R_T(h, \lambda, \sigma, \hat{p}, t)$ is a unitless quantity (counts/photon times aperture area/detector area).

The mathematical form of equation (2) represents a linear mapping from S to S_D. Hence, strictly speaking, equation (2) is valid only when one can neglect nonlinear effects such as local gain depression or pile-up, i.e., a nonlinear effect caused by the finite time resolution of the detector. Pile-up can be treated, at least in principle, by making the response $R_T(h, \lambda, \sigma, \hat{p}, t)$ a function of the incident flux $S(\lambda, \hat{p})$. However, this topic is beyond the scope of the present work and will be addressed elsewhere (J. E. Davis 2000, in preparation).

The explicit time dependence of $R_T(h, \lambda, \sigma, \hat{p}, t)$ is due to effects associated with telemetry saturation, thermal expansion, and so on. It is assumed that all but the time dependence associated with internal movement of the telescope subsystems may be encapsulated in a function $T(\sigma, t)$ that factors out of $R_T(h, \lambda, \sigma, \hat{p}, t)$, i.e.,

$$R_T(h, \sigma, \lambda, \hat{p}, t) = T(\sigma, t)R(h, \sigma, \lambda, \hat{p}, t) . \tag{3}$$

In this equation the residual time dependence of the response function $R(h, \sigma, \lambda, \hat{p}, t)$ depends only upon the relative movements of the individual telescope subsystems. The function $T(\sigma, t)$ can be thought of as representing the so-called good time intervals (GTIs). However, $T(\sigma, t)$ could play a more general role than this because it could include the effects of bad pixels,[2] which themselves may be time dependent, as well as any dead-time factors associated with telemetry saturation. It should also be noted that although the incident source flux has been assumed to be independent of time, $T(\sigma, t)$ could encompass time-dependent variations in the source flux as long as the spectral shape itself does not depend upon time. In this case, only the amplitude of the flux is time dependent and this time dependence may be factored out of the incident source flux and into the function $T(\sigma, t)$. Further exploration of this possibility is left to the reader. Suffice it to say that the actual values that $T(\sigma, t)$ take on are not important in what follows. Hence, equation (2) can be written

$$S_D(h, \sigma, t) = T(\sigma, t) \int d\lambda \int d\hat{p}\, R(h, \sigma, \lambda, \hat{p}, t)$$
$$\times S(\lambda, \mathscr{R}^{-1}(t) \cdot \hat{p})$$
$$= T(\sigma, t) \int d\lambda \int d\hat{p}\, R(h, \sigma, \lambda, \mathscr{R}(t) \cdot \hat{p}, t)$$
$$\times S(\lambda, \hat{p}) , \tag{4}$$

where the last form follows from the change of variables $\hat{p} \to \mathscr{R}(t) \cdot \hat{p}$ and noting that because the matrix $\mathscr{R}(t)$ is orthogonal, the Jacobian of the transformation is unity.

Assuming that the response depends upon several independent telescope subsystems, it may be factored into subsystem-dependent pieces. The specific form of the factorization will depend upon the actual physical relationships between the various subsystems. For a prototypical X-ray telescope with a focusing mirror at the aperture of the telescope that focuses X-rays onto a position-sensitive detector, an appropriate factorization is given by

$$R(h, \sigma, \lambda, \hat{p}, t) = D(h, \sigma, \lambda) \int d\hat{p}'\, \delta(\sigma - \sigma(\hat{p}', t))$$
$$\times \mathscr{F}(\lambda, \hat{p}', \hat{p})M(\lambda, \hat{p}) . \tag{5}$$

A different factorization may have to be used to describe the response of some other type of telescope (see § 5 for the factorization needed to describe the presence of a diffraction grating). In this equation, $M(\lambda, \hat{p})$ is the off-axis effective area of the mirror and $D(h, \sigma, \lambda)$ represents the probability that a photon with wavelength λ at position σ on the detector will give rise to a pulse height h. The term involving the delta function symbolizes the passage of a photon with direction $\hat{p}'$ from the mirror to the position σ on the detec-

---

[1] If the source flux varies in time such that its time dependence is uncorrelated with the spatial and spectral shape, as is often the case, then one can always factor out the time dependence and handle it separately. This technique is discussed in more detail below. The treatment of more complex time-dependent sources that do not admit this factorization is beyond the scope of this paper.

[2] As described below, the effects of static bad pixels are assumed to be contained in the detector response function $D(h, \sigma, \lambda)$.

tor via the coordinate transformation $\sigma(\hat{p}', t)$. In general, this function is time dependent and represents any relative motion that exists between the mirror and the detector.

The PSF of the telescope is represented by the function $\mathcal{F}(\lambda, \hat{p}', \hat{p})$, which is assumed to satisfy the normalization condition

$$1 = \int d\hat{p}' \mathcal{F}(\lambda, \hat{p}', \hat{p}) . \qquad (6)$$

Its definition is based upon the idea (see Fig. 1) that the mirror itself (e.g., one of a typical type I Wolter design) may be modeled by the appropriate probability distribution for photons to enter and leave the mirror at a single point (the so-called mirror node), with an effective area given by $M(\lambda, \hat{p})$. As indicated in Figure 1, the PSF defined in this way will also depend (implicitly) upon the position of the detector and its relationship to the focal surface. For simplicity, in the following it will be assumed that the variations in the movement of the detector with respect to the mirror during the course of an observation are small enough that the PSF may be regarded as independent of time to sufficient accuracy. The reader should note that a perfect PSF defined in this sense is represented by

$$\mathcal{F}_{\text{perfect}}(\lambda, \hat{p}', \hat{p}) = \delta(\hat{p}' - \hat{p}) . \qquad (7)$$
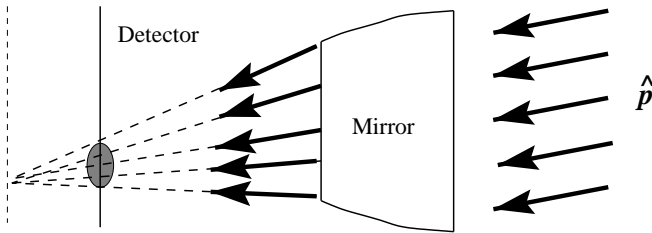
(a)
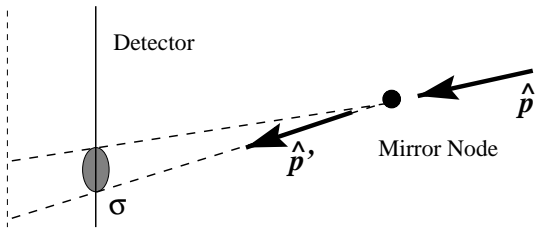
(b)

FIG. 1.—Figure illustrating how the PSF $\mathcal{F}(\lambda, \hat{p}', \hat{p})$ may be used to model a realistic mirror of finite extent accurately. The top portion of the figure shows a "realistic" mirror with the detector in a defocused position. In this position photons from a distant point source enter the mirror at the front and exit the mirror at the back and are focused toward a point behind the detector causing the point source to appear out of focus. The bottom part of the figure shows that this effect can be modeled in terms of the concept of a "mirror node," where rays enter and leave the same point with the appropriate redistribution $\mathcal{F}(\lambda, \hat{p}', \hat{p})$ of "incoming" photons with direction $\hat{p}$ into "outgoing" ones with direction $\hat{p}'$. In particular, note that the same PSF appropriate for the defocused position of the detector cannot be used to model the image when the detector is at the focal position; there a different PSF will be required. In other words, for $\mathcal{F}(\lambda, \hat{p}', \hat{p})$ to model the mirror accurately, it also depends *implicitly* upon the location of the detector with respect to the focal surface.

Generally speaking, the detector response $D(h, \sigma, \lambda)$ can be factored into a QE $Q(\sigma, \lambda)$ and a redistribution function $D_R(h, \sigma, \lambda)$:

$$D(h, \sigma, \lambda) = D_R(h, \sigma, \lambda)Q(\sigma, \lambda) . \qquad (8)$$

The function $D_R(h, \sigma, \lambda)$ is known as the RMF (George et al. 2000) and represents a mapping, or redistribution, from wavelength to pulse height by the detector. Without any loss of generality, it is assumed to be normalized to unity via

$$1 = \sum_h D_R(h, \sigma, \lambda) . \qquad (9)$$

Alternatively, the quantum efficiency may be defined by

$$Q(\sigma, \lambda) = \sum_h D(h, \sigma, \lambda) . \qquad (10)$$

In general, as indicated here the RMF varies with position on the detector, although many applications assume a spatially constant RMF. In contrast, the QE function $Q(\sigma, \lambda)$ is assumed to contain the effects of (static) bad pixels, detector dead regions, detector boundaries, and so on, all of which cause it to vary spatially.

Using the response function given by equation (5), the expected count rate from the source $S(\lambda, \hat{p})$ is seen to be

$$S_D(h, \sigma, t) = T(\sigma, t) \int d\lambda \int d\hat{p} \, R(h, \sigma, \lambda, \hat{p}_t, t)S(\lambda, \hat{p})$$

$$= T(\sigma, t) \int d\lambda \, D(h, \sigma, \lambda) \int d\hat{p}' \, \delta(\sigma - \sigma(\hat{p}', t))$$

$$\times \int d\hat{p} \, \mathcal{F}(\lambda, \hat{p}', \hat{p}_t)M(\lambda, \hat{p}_t)S(\lambda, \hat{p}) , \qquad (11)$$

where, for notational simplicity, $\hat{p}_t$ symbolizes $\mathcal{R}(t) \cdot \hat{p}$.

Telescope pointing motion will cause events to appear spatially mixed together when expressed as an image in detector coordinates. For this reason, it is preferable to work in the sky coordinate system where one can remove the effects of the motion by projecting the events to the sky in the appropriate manner. Thus, we define an aspect-corrected count rate via

$$S_A(h, \hat{p}, t) = J(\sigma(\hat{p}_t, t), \hat{p}_t, t)S_D(h, \sigma(\hat{p}_t, t), t) , \qquad (12)$$

where $J(\sigma, \hat{p}, t)$ is the instantaneous Jacobian of the transformation from the detector coordinate σ to the sky coordinate $\hat{p}$ via the inversion of $\sigma(\hat{p}_t, t)$. Physically, the Jacobian represents the stretching or magnification of an element of area on the detector as it appears in the sky.

By making use of the well-known change of variable formula for delta functions expressed in the form

$$\delta(\sigma(\hat{p}_t, t) - \sigma(\hat{p}', t)) = \frac{\delta(\hat{p}' - \hat{p}_t)}{J(\sigma(\hat{p}_t, t), \hat{p}_t, t)} , \qquad (13)$$

one can show that the aspect-corrected count rate is

$$S_A(h, \hat{p}', t) = T(\sigma(\hat{p}_t', t), t) \int d\lambda \, D(\sigma(\hat{p}_t', t), h, \lambda)$$

$$\times \int d\hat{p} \, \mathcal{F}(\lambda, \hat{p}_t', \hat{p}_t)M(\lambda, \hat{p}_t)S(\lambda, \hat{p}) . \qquad (14)$$

The above equation assumes that complete knowledge of the aspect history is available in order to perform the aspect correction. In general, there will be uncertainties in the

aspect solution, which in turn leads to spatial uncertainties in the aspect-corrected count rate. Mathematically, this will manifest itself as a broadening or smearing of the delta function in equation (13) by an amount that depends upon the aspect uncertainties. The most straightforward way to handle this effect is to absorb the uncertainties into the PSF itself. For this reason, in the following sections $\mathcal{F}_A(\lambda, \hat{p}', \hat{p})$ will denote the PSF that includes the effect of the aspect uncertainties.

3. DERIVATION OF THE ARF

The total number of expected counts with pulse height h over an observation interval τ in some region Ω can be computed by integrating t over the observation interval and $\hat{p}'$ over the sky region Ω, i.e.,

$$C_\Omega(h) = \int_\Omega d\hat{p}' \int_0^\tau dt \, S_A(h, \hat{p}', t) \tag{15}$$

$$= \int d\lambda \int_\Omega d\hat{p}' \int_0^\tau dt \, T(\sigma(\hat{p}'_t, t), t) D(\sigma(\hat{p}'_t, t), h, \lambda)$$

$$\times \mathcal{F}_A(\lambda, \hat{p}'_t, \hat{q}_t) M(\lambda, \hat{q}_t) S(\lambda, \hat{p}) .$$

By using equation (8) and assuming for the moment that the RMF does not vary with position, the previous equation can be rewritten as

$$C_\Omega(h) = \tau_{\rm eff} \int d\lambda \, D_R(h, \lambda) \int d\hat{p} \, A_\Omega(\lambda, \hat{p}) S(\lambda, \hat{p}) , \tag{16}$$

where

$$A_\Omega(\lambda, \hat{p}) = \frac{1}{\tau_{\rm eff}} \int_\Omega d\hat{p}' \int_0^\tau dt \, T(t) Q(\lambda, \sigma(\hat{p}'_t, t))$$

$$\times \mathcal{F}_A(\lambda, \hat{p}'_t, \hat{p}_t) M(\lambda, \hat{p}_t) \tag{17}$$

and the effective exposure time is given by

$$\tau_{\rm eff} = \int_0^\tau dt \, T(t) . \tag{18}$$

For simplicity it has been assumed that the GTI function $T(t)$ does not depend upon detector position.

Equation (17) defines the ARF. It has the units of area times counts per photon and depends upon the region Ω, wavelength λ, and sky position $\hat{p}$. However, for the purposes of point-source analysis, knowledge of the ARF is required only at the position of the point source, where it may be regarded as depending only upon wavelength with the understanding that it is valid only for that source position and region. But for arbitrary sources, it should be regarded as an explicit function of $\hat{p}$.

Armed with the ARF, equation (16) is an integral equation that may be "solved" to yield the source distribution $S(\lambda, \hat{p})$ from the observed aspect-corrected counts $C_\Omega(h)$. Actually, because of the integration over the sky region Ω, much of the spatial dependence in $S(\lambda, \hat{p})$ will be lost and in practice one will have to assume a known spatial dependence; the examples below illustrate this point more fully. One should also realize that the kernel of equation (16) is really a probability distribution and that the observed number of counts will most likely differ from the expected number of counts predicted by the equation. This implies that equation (16) does not really have a unique solution (for a finite observation time), and any method of "solving" should allow for fluctuations in the number of counts. One

must also take into account any external background sources as well as any internal background produced by, say, noise in the detector. Techniques for treating this equation are beyond the scope of this paper and may be found elsewhere (Arnaud 1996; Kahn & Blissett 1980).

In the next few sections, equation (16) is considered in the context of various source distributions. The problem of the practical computation of the ARF is taken up in Appendix A.

3.1. A Single Point Source

A point source located at position $\hat{q}$ in the sky with a spectrum $s(\lambda)$ may be represented using the source distribution

$$S(\lambda, \hat{p}) = s(\lambda)\delta(\hat{p} - \hat{q}) . \tag{19}$$

Substituting this equation into equation (16) yields

$$C_\Omega(h) = \tau_{\rm eff} \int d\lambda \, D_R(h, \lambda) A_\Omega(\lambda, \hat{q}) s(\lambda) . \tag{20}$$

This equation is essentially the integral equation that the popular spectral analysis program *xspec* (Arnaud 1996) is designed to solve. As noted above, the ARF is required to be computed only at the position $\hat{q}$ of the point source; because of the time dependence of the telescope motion, the integration over t in equation (17) does however sample various detector regions and variations in the off-axis effective area.

3.2. Multiple Point Sources

Multiple point sources may be represented by a source distribution of the form

$$S(\lambda, \hat{p}) = \sum_i s_i(\lambda)\delta(\hat{p} - \hat{q}_i) , \tag{21}$$

where the position of the ith source is given by $\hat{q}_i$ and its spectrum is $s_i(\lambda)$. Insertion of this distribution into equation (16) produces

$$C_\Omega(h) = \tau_{\rm eff} \int d\lambda \, D_R(h, \lambda) \left[\sum_i A_\Omega(\lambda, \hat{q}_i) s_i(\lambda) \right] . \tag{22}$$

If all of the sources have an identical source spectrum such that $s_i(\lambda) = s(\lambda)$, then the resulting integral equation reduces to the case of a single point source, i.e.,

$$C_\Omega(h) = \tau_{\rm eff} \int d\lambda \, D_R(h, \lambda) \left[\sum_i A_\Omega(\lambda, \hat{q}_i) \right] s(\lambda) . \tag{23}$$

However, the more interesting case of spectrally distinct sources is more complicated to solve. In fact, its solution would require N integral equations since there are N unknown spectral distributions $s_i(\lambda)$. The most straightforward way to obtain the required number of independent equations would be to use N different regions Ω_i, not necessarily disjoint, and solve the resulting linear system of equations

$$C_{\Omega_i}(h) = \tau_{\rm eff} \int d\lambda \, D_R(h, \lambda) \left[\sum_j A_{\Omega_i}(\lambda, \hat{q}_j) s_j(\lambda) \right] . \tag{24}$$

These equations are "coupled" to the extent that the $A_{\Omega_i}(\lambda, \hat{q}_j)$ for $i \neq j$ are nonzero, i.e., whether or not source j has a PSF contribution to region Ω_i (see Fig. 2). Such a
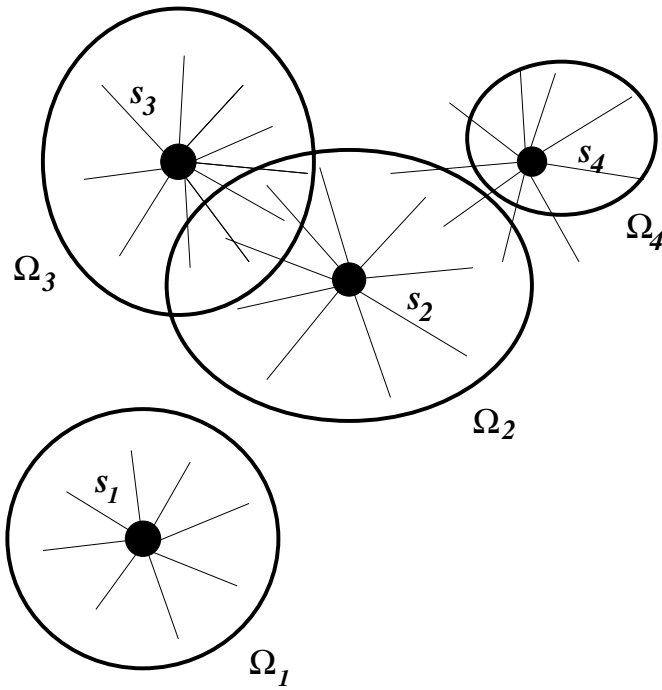
FIG. 2.—Figure illustrating the use of multiple regions in eq. (24) for a system of four point sources. In this figure, the thin radial lines emanating from each of the point sources represent the PSF (100% enclosed power) for the corresponding source, and the elliptical borders outline regions Ω_i over which counts are summed. The PSF from source s_1 does not contribute to regions Ω_2, Ω_3, and Ω_4; hence, it effectively decouples from the other sources. Similarly, s_4 may be treated by itself since none of the other sources contribute any counts to Ω_4. However, in dealing with s_2, the contributions from both s_3 and s_4 must be taken into account.

system of equations may be handled using *sherpa* (Doe et al. 1998), the *Chandra* Data System spectral analysis program.

3.3. *Extended Source with Uncorrelated Spatial and Spectral Distributions*

One of the simplest examples of an extended source is one in which the spatial and spectral distributions are uncorrelated. That is, the source distribution $S(\lambda, \hat{p})$ factors according to

$$S(\lambda, \hat{p}) = s(\lambda)\rho(\hat{p}) , \quad (25)$$

where $\rho(\hat{p})$ defines the spatial distribution, assumed to be properly normalized such that

$$1 = \int d\hat{p}\, \rho(\hat{p}) . \quad (26)$$

Combining this distribution with equation (16) yields

$$C_\Omega(h) = \tau_{\text{eff}} \int d\lambda\, D_R(h, \lambda)\left[\int d\hat{p}\, A_\Omega(\lambda, \hat{p})\rho(\hat{p})\right]s(\lambda) . \quad (27)$$

Suppose that the form of $\rho(\hat{p})$ is known. Then, the ARF could be combined with the known spatial distribution by defining

$$A_\Omega^{(\rho)}(\lambda) = \int d\hat{p}\, A_\Omega(\lambda, \hat{p})\rho(\hat{p}) , \quad (28)$$

which leads to the *xspec* style equation

$$C_\Omega(h) = \tau_{\text{eff}} \int d\lambda\, D_R(h, \lambda)A_\Omega^{(\rho)}(\lambda)s(\lambda) \quad (29)$$

for the unknown spectral function $s(\lambda)$. Of course, this methodology cannot be used if the spatial distribution $\rho(\hat{p})$ is not known. The more general problem is addressed below in § 4.

3.4. *An ARF in the Presence of a Spatially Varying RMF*

It is important to note that the ARF, given in equation (17), is useful only when one can disregard spatial variations in the RMF. Unfortunately, this may not always be possible. For example, *Chandra* ACIS CCDs have a spatially varying response that must be properly taken into account. Provided that one wants to stay within the confines of the existing ARF + RMF paradigm, the only way to handle such cases properly is to filter the observed events over a region in detector coordinates where spatial variations in the RMF may be neglected. Mathematically, this procedure may be stated as follows. Let Γ denote the region on the detector where the RMF does not vary. Then define a filter $F^\Gamma(\sigma)$ on this region by

$$F^\Gamma(\sigma) = \begin{cases} 1 & \sigma \in \Gamma , \\ 0 & \text{otherwise} . \end{cases} \quad (30)$$

Multiplication of equation (11) by this filter, followed by aspect correction and summing over the sky region Ω, yields

$$C_\Omega^\Gamma(h) = \tau_{\text{eff}} \int d\lambda\, D_R^\Gamma(h, \lambda) \int d\hat{p}\, A_\Omega^\Gamma(\lambda, \hat{p})S(\lambda, \hat{p}) , \quad (31)$$

where

$$A_\Omega^\Gamma(\lambda, \hat{p}) = \frac{1}{\tau_{\text{eff}}} \int_\Omega d\hat{p}' \int_0^\tau dt\, T(t)F^\Gamma(\sigma(\hat{p}_t', t))$$
$$\times Q(\lambda, \sigma(\hat{p}_t', t))\mathscr{F}_A(\lambda, \hat{p}_t', \hat{p}_t)M(\lambda, \hat{p}_t) . \quad (32)$$

In these equations, $C_\Omega^\Gamma(h)$ is the expected number of counts in the sky region Ω that also fall within the detector region Γ, $A_\Omega^\Gamma(\lambda, \hat{p})$ is the ARF appropriate for this region of the detector, and $D_R^\Gamma(h, \lambda)$ is the region-dependent RMF.

4. DEFINITION OF THE EXPOSURE MAP

The ARF presented in the previous section is primarily of use for spectral analysis over small spatial regions, e.g., the analysis of point sources. Much of the spatial information useful for the treatment of extended sources was lost in the construction of the ARF by integrating the response over a region Ω of the sky. A related product, the exposure map, does not depend upon the integration over a sky region permitting it to be used for certain types of extended source analysis. The goal of this section is to define an exposure map and show how it may be used with extended sources.

Start by integrating equation (14) over an observation time τ to produce

$$C(h, \hat{p}') = \int_0^\tau dt\, S_A(h, \hat{p}', t)$$
$$= \int_0^\tau dt \int d\lambda\, T(\sigma(\hat{p}_t', t), t)D(\sigma(\hat{p}_t', t), h, \lambda)$$
$$\times \int d\hat{p}\, \mathscr{F}_A(\lambda, \hat{p}_t', \hat{p}_t)M(\lambda, \hat{p}_t)S(\lambda, \hat{p}) . \quad (33)$$

Here $C(h, \hat{p}')$ represents the expected total number of aspect-corrected counts with pulse height h attributed to the sky position $\hat{p}'$.

In general the PSF is small for rays not too far off the optical axis, although it can become quite large for far off-axis rays. Suppose that the pointing motion amplitude is small enough that the PSF may be regarded as a scalar under the motion, i.e.,

$$\mathscr{F}_A(\lambda, \hat{p}', \hat{p}) = \mathscr{F}_A(\lambda, \hat{p}'_t, \hat{p}_t) , \tag{34}$$

and then consider the integration over $\hat{p}$ in equation (33). For off-axis positions where the size of the PSF is small, for a fixed $\hat{p}'$, only a narrow range of $\hat{p}$ contributes to the integral. For the moment, assume that the mirror effective area does not vary much over this range. Then the approximation

$$\int d\hat{p}\, \mathscr{F}_A(\lambda, \hat{p}'_t, \hat{p}_t) M(\lambda, \hat{p}_t) S(\lambda, \hat{p})$$

$$\approx M(\lambda, \hat{p}'_t) \int d\hat{p}\, \mathscr{F}_A(\lambda, \hat{p}', \hat{p}) S(\lambda, \hat{p}) \tag{35}$$

can be used in equation (33) to yield

$$C(h, \hat{p}') = \int d\lambda \int_0^\tau dt\, T(\sigma(\hat{p}'_t, t), t)$$

$$\times D(\sigma(\hat{p}'_t, t), h, \lambda) M(\lambda, \hat{p}'_t) S_{\mathscr{F}}(\lambda, \hat{p}') , \tag{36}$$

where a PSF-smeared source $S_{\mathscr{F}}(\lambda, \hat{p}')$ has been defined by

$$S_{\mathscr{F}}(\lambda, \hat{p}') = \int d\hat{p}\, \mathscr{F}_A(\lambda, \hat{p}', \hat{p}) S(\lambda, \hat{p}) . \tag{37}$$

With the introduction of the exposure map $E(h, \lambda, \hat{p})$ defined by

$$E(h, \lambda, \hat{p}) = \frac{1}{\tau_{\mathrm{eff}}} \int_0^\tau dt\, T(\sigma(\hat{p}_t, t), t)$$

$$\times D(\sigma(\hat{p}_t, t), h, \lambda) M(\lambda, \hat{p}_t) , \tag{38}$$

equation (36) may be recast as

$$C(h, \hat{p}) = \tau_{\mathrm{eff}} \int d\lambda\, E(h, \lambda, \hat{p}) S_{\mathscr{F}}(\lambda, \hat{p}) . \tag{39}$$

It is important to understand that this is an integral equation describing the PSF-smeared source and not the true source. After "solving" this equation, one still has the task of removing the effects of the PSF to determine the true source. Nevertheless, equation (39) does have one very important feature not shared by the equations involving the ARF: the spatial distribution of the expected aspect-corrected counts, $C(h, \hat{p})$, is the same as the PSF-smeared source's spatial distribution.

A common use of the exposure map is to remove instrumental artifacts in images to obtain a better looking image. This is also known as "flux-correcting" the image. The method essentially assumes that the pulse-height resolution of the detector permits the separation of counts originating from photons from different energy bands, supplemented by the assumption that the source flux may be regarded as constant within a band (Snowden et al. 1994). To express this idea in quantitative terms, consider a range of Δh of pulse heights centered on some pulse height h and assume that the RMF $D_R(h, \lambda)$ is such that only those photons from the wavelength band $\Delta\lambda$ about λ can produce pulse heights in the specified range. Now sum equation (39) over this range and consider only the photons from the wavelength

band $\Delta\lambda$ to yield

$$\sum_{h \in \Delta h} C(h, \hat{p}) = \tau_{\mathrm{eff}} \int_{\lambda \in \Delta\lambda} d\lambda \sum_{h \in \Delta h} E(h, \lambda, \hat{p}) S_{\mathscr{F}}(\lambda, \hat{p}) , \tag{40}$$

which may be written in the more compact form

$$C(\Delta h, \hat{p}) = \tau_{\mathrm{eff}} \int_{\lambda \in \Delta\lambda} d\lambda\, E(\Delta h, \lambda, \hat{p}) S_{\mathscr{F}}(\lambda, \hat{p}) , \tag{41}$$

with the understanding that the pulse-height range Δh is summed over.

If the bandwidth $\Delta\lambda$ is such that $E(\Delta h, \lambda, \hat{p})$ does not vary much over the band, then it may be removed from the integrand to obtain

$$\int_{\lambda \in \Delta\lambda} d\lambda\, S_{\mathscr{F}}(\lambda, \hat{p}) \approx \frac{1}{\tau_{\mathrm{eff}}} \frac{C(\Delta h, \hat{p})}{E(\Delta h, \lambda, \hat{p})} . \tag{42}$$

This equation says that the integrated profile of the PSF-smeared source flux over the wavelength band may be obtained by dividing the exposure map into the counts image constructed from the appropriate pulse-height range.

The resolution of the source spectrum obtained by this technique is generally poorer than that of the detector's energy resolution because the wavelength band $\Delta\lambda$ must be large enough to cover all the wavelengths that could contribute to the range of pulse heights Δh. At the same time it must be small enough to ensure that $E(\Delta h, \lambda, \hat{p})$ may be treated as a constant in the wavelength band. It is possible that there may be bands in which these constraints are mutually exclusive. For this reason, the use of the exposure map is limited to situations in which spectral resolution is of secondary importance. For example, spatial resolution is much more important than spectral resolution when doing source detection. For such a situation, one would run the source detection algorithm on an image obtained by dividing the total counts image over an exposure map integrated over the bandpass of the telescope. Another application of the exposure map would be to use it to get a crude estimate of the true spectrum and to use that as the first approximation in some more refined technique.

Before leaving this section, it is important to point out that equation (39) is valid only as long as the size of the PSF is small enough that any variation in the effective area over the PSF can be neglected. If this is not the case, then it is impossible to give a definition of an exposure map that has the simple relationship between the observed counts image and the PSF-smeared source as described by this equation. By implication it follows that such an exposure map cannot be used for flux correction via the simple division of equation (42). However, if the detector response $D(\sigma, h, \lambda)$ is uniform, then it is possible to commute the response with $\mathscr{F}(\lambda, \hat{p}', \hat{p})$ in equation (33) to produce

$$C(h, \hat{p}') = \tau_{\mathrm{eff}} \int d\hat{p} \int d\lambda\, \mathscr{F}_A(\lambda, \hat{p}', \hat{p}) E(h, \lambda, \hat{p}) S(\lambda, \hat{p}) . \tag{43}$$

This means that one must first deconvolve the effects of the PSF before correcting with the exposure map. The feasibility of this will depend upon the energy dependence of the PSF where one may have to perform the deconvolution in specific energy bands. This prescription resembles the one advocated by White & Buote (2000) for the analysis of *ASCA* data. Alternatively, Ikebe (1995) has argued that one start essentially from equation (33) and employ a "forward-folding" method to estimate the source distribution $S(\lambda, \hat{p})$.

5. DERIVATION OF THE GRATING ARF

When diffraction gratings are added, the subsystem factorization, equation (5), must be modified to

$$R(h, \sigma, \lambda, \hat{p}, t) = D(h, \sigma, \lambda) \int d\hat{p}' \, \delta(\sigma - \sigma_G(\hat{p}', t))$$

$$\times \left[\sum_m g_m(\lambda) \mathscr{F}_m(\lambda, \hat{p}', \hat{p}) \right] M(\lambda, \hat{p}) , \quad (44)$$

where m represents the diffraction order and $\delta(\sigma - \sigma_G(\hat{p}', t))$ symbolizes the coordinate transformation of a diffracted ray at the grating node with direction $\hat{p}'$ to the detector coordinate σ. The actual diffraction into the mth order is represented by the term $g_m(\lambda)\mathscr{F}_m(\lambda, \hat{p}', \hat{p})$, which gives the probability for a ray with direction $\hat{p}$ and wavelength λ to diffract into the mth order with direction $\hat{p}'$. The function $g_m(\lambda)$ is the mth-order grating efficiency, and by definition the redistribution function satisfies the normalization condition

$$1 = \int d\hat{p}' \, \mathscr{F}_m(\lambda, \hat{p}', \hat{p}) . \quad (45)$$

Despite the similarity in form of $\mathscr{F}_m(\lambda, \hat{p}', \hat{p})$ to the imaging PSF $\mathscr{F}(\lambda, \hat{p}', \hat{p})$, it is important to appreciate one very important difference between these two functions. The imaging PSF is sharply peaked about the set of directions $\hat{p}'$ near $\hat{p}$. However, $\mathscr{F}_m(\lambda, \hat{p}', \hat{p})$ is sharply peaked about a set of directions $\hat{p}'$ that vary linearly with wavelength λ according to the diffraction equation

$$(\hat{p}' - \hat{p}) \times \hat{n} = \frac{m\lambda}{d} \hat{l} , \quad (46)$$

where d is the grating period, the vector $\hat{n}$ is normal to the plane of the grating, and $\hat{l}$ is in the direction of the grating bars (see Fig. 3).

It follows trivially from equation (44) that the expected count rate into mth order is given by

$$S_D^{(m)}(h, \sigma, t) = T(\sigma, t) \int d\lambda \left[D(h, \sigma, \lambda) g_m(\lambda) \right.$$

$$\times \int d\hat{p}' \, \delta(\sigma - \sigma_G(\hat{p}', t))$$

$$\left. \times \int d\hat{p} \, \mathscr{F}_m(\lambda, \hat{p}', \hat{p}_t) M(\lambda, \hat{p}_t) S(\lambda, \hat{p}) \right] . \quad (47)$$

As in the imaging case, an aspect-corrected count rate may be defined by

$$S_A^{(m)}(h, \hat{p}, t) = J(\sigma(\hat{p}_t, t), \hat{p}_t, t) S_D^{(m)}(h, \sigma(\hat{p}_t, t), t) , \quad (48)$$

with the result

$$S_A^{(m)}(h, \hat{p}', t) = T(\sigma_G(\hat{p}'_t, t), t)$$

$$\times \int d\lambda \, D(\sigma_G(\hat{p}'_t, t), h, \lambda) g_m(\lambda)$$

$$\times \int d\hat{p} \, \mathscr{F}_m(\lambda, \hat{p}'_t, \hat{p}_t) M(\lambda, \hat{p}_t) S(\lambda, \hat{p}) . \quad (49)$$

This equation may be simplified for the special case of a point source located at $\hat{q}$ with a spectrum $s(\lambda)$, i.e.,

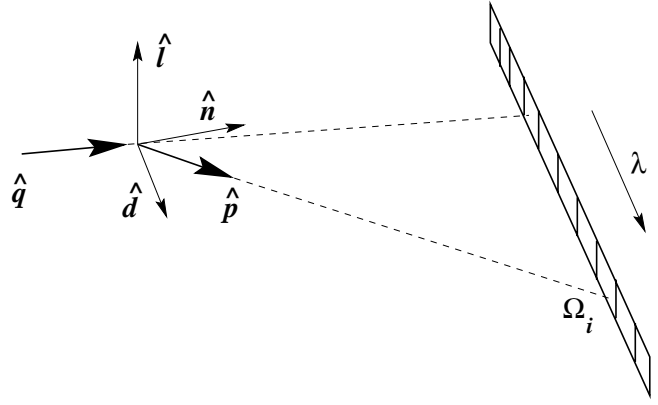$$S(\lambda, \hat{p}') = s(\lambda)\delta(\hat{p}' - \hat{q}) . \quad (50)$$

FIG. 3.—Figure illustrating how the grating RMF, $G_{\Omega_i}^{(m)}(\lambda, \hat{q})$, represents a redistribution from wavelength λ to regions Ω_i. In this figure, the triad of unit vectors $(\hat{n}, \hat{l}, \hat{d})$ specifies an orthonormal coordinate basis centered upon the diffraction grating, with $\hat{n}$ normal to the surface of the grating and $\hat{l}$ in the direction of the grating bars. For a perfect grating, a photon with direction $\hat{q}$ and wavelength λ will diffract into the direction $\hat{p}$ in accordance with eq. (54) and a specified diffraction order m. For a realistic grating, the diffraction process must be described by a probability distribution $\mathscr{F}_m(\lambda, \hat{p}, \hat{q})$ that is sharply peaked around the set of values that satisfy eq. (54). The mth-order grating RMF, $G_{\Omega_i}^{(m)}(\lambda, \hat{q})$, defined by eq. (56), represents the redistribution probability for an incoming photon with direction $\hat{q}$ and wavelength λ to diffract into a region Ω_i. In other words, for a specified zeroth-order direction $\hat{q}$, the grating RMF may be regarded as a mapping from wavelength λ to region Ω_i in much the same way as the detector RMF represents a mapping from wavelength λ to pulse height h.

In addition, assume that the telescope pointing motion amplitudes are small enough that the grating redistribution function behaves like a scalar under the motion as in equation (34). Then, the integral over $\hat{p}'$ may be readily performed to yield

$$S_A^{(m)}(h, \hat{p}, t) = T(\sigma_G(\hat{p}_t, t), t) \int d\lambda \, D(\sigma_G(\hat{p}_t, t), h, \lambda)$$

$$\times g_m(\lambda)\mathscr{F}_m(\lambda, \hat{p}, \hat{q}) M(\lambda, \hat{q}_t) s(\lambda) . \quad (51)$$

Integration of this equation over a set of regions Ω_i and time τ yields for the total expected number of counts with pulse height h in the regions

$$C_{\Omega_i}^{(m)}(h) = \int d\lambda \int_{\Omega_i} d\hat{p} \left[g_m(\lambda) \int_0^\tau dt \, T(\sigma_G(\hat{p}_t, t), t) \right.$$

$$\left. \times D(\sigma_G(\hat{p}_t, t), h, \lambda) M(\lambda, \hat{q}_t) \right]$$

$$\times \mathscr{F}_m(\lambda, \hat{p}, \hat{q}) s(\lambda) . \quad (52)$$

For fixed λ and $\hat{q}$, the sharp-peaked nature of the grating redistribution function $\mathscr{F}_m(\lambda, \hat{p}, \hat{q})$ implies that only a very narrow set of directions $\hat{p}$ will contribute to the term in square brackets. Moreover, any telescope pointing motion is expected to smooth out any nonuniformities in the detector QE appearing in this term such that one can evaluate it using the value of $\hat{p}$ determined by the grating equation. In other words, the term in square brackets can be replaced by a function $A_m(h, \lambda)$ defined by

$$A_m(h, \lambda) = g_m(\lambda) \frac{1}{\tau_{\text{eff}}} \int_0^\tau dt \, T(\sigma_G(\hat{p}_t, t), t)$$

$$\times D(\sigma_G(\hat{p}_t, t), h, \lambda) M(\lambda, \hat{q}_t) , \quad (53)$$

where $\hat{p}$ satisfies

$$(\hat{p} - \hat{q}) \times \hat{n} = \frac{m\lambda}{d}\,\hat{l}\,. \qquad (54)$$

Hence, the number of counts in the mth order with pulse height h is expected to be

$$C_{\Omega_i}^{(m)}(h) = \tau_{\text{eff}} \int d\lambda\, G_{\Omega_i}^{(m)}(\lambda,\,\hat{q})A_m(h,\,\lambda)s(\lambda)\,, \qquad (55)$$

where

$$G_{\Omega_i}^{(m)}(\lambda,\,\hat{q}) = \int_{\Omega_i} d\hat{p}\,\mathscr{F}_m(\lambda,\,\hat{p},\,\hat{q}) \qquad (56)$$

and τ_{eff} is given by equation (18).

For reasons that will soon become clear, $A_m(h,\,\lambda)$ is called the grating ARF, and $G_{\Omega_i}^{(m)}(\lambda,\,\hat{q})$ is called the grating RMF. To see this, consider the meaning of equation (56). For fixed λ, equation (56) represents a redistribution from wavelength λ to the region Ω_i, which may be regarded as the ith bin in $\hat{p}$-space (see Fig. 3). This is the analog of the imaging RMF that describes a redistribution from λ to a bin in pulse-height space. With this interpretation, equation (55) is formally identical to equation (20), provided that one identifies $A_m(h,\,\lambda)$ with the ARF. Hence, any techniques that are applicable to equation (20) may be readily applied to the solution of equation (55).

Although $A_m(h,\,\lambda)$ depends upon the pulse height, in practice events will be filtered upon the pulse height in order to perform order separation, provided that the intrinsic energy resolution of the detector is adequate. For detectors with poor energy resolution, some other means of identifying mth-order events will have to be used. In any case, $A_m(h,\,\lambda)$ will most likely be summed over the range of pulse heights appropriate to mth-order events. In fact, the pulse-height range will generally vary with the wavelength such that the quantity

$$A_m(\lambda) = \sum_{h = h_0(\lambda)}^{h_1(\lambda)} A_m(h,\,\lambda) \qquad (57)$$

will actually be what is used in practice. For this reason, it is preferable to define the summed quantity $A_m(\lambda)$ as the grating ARF.

6. CONCLUSION

In this paper, explicit expressions for the imaging ARF, grating ARF, and the exposure map were given in terms of the underlying instrumental responses that are consistent with the current use of the objects. These quantities were obtained from first principles by relating the expected detector count rate to an incident photon source flux via the overall telescope response function suitably factored into individual instrumental responses.

One of the complications in the derivation of these quantities concerned the proper treatment of time-varying effects due to telescope pointing motion, e.g., dither. At the same time, the assumed presence of motion about some nominal pointing allowed some important factorizations to take place that otherwise would have been suspect in regions containing detector boundaries or bad pixels. For this reason, purposely dithering an observation is recommended, provided, of course, that one can reconstruct the aspect history with sufficient accuracy.

An added benefit of the first principles approach taken here is that it allows one to consider problems that cannot readily be handled by conventional means through the use of an ARF and an RMF. For example, as shown in § 3.3, an ARF may be applied to the analysis of an extended source provided one knows a priori that the source flux distribution factors into a known spatial component and an unknown spectral component. It is easy to find sources where such a factorization is not permissible; the supernova remnant, Cassiopeia A, is one. Another problem that does not appear to be treatable through standard techniques is the analysis of an extended source in the presence of a diffraction grating. The grating ARF defined in § 5 was derived assuming a point-source distribution. The basic problem with the analysis of an extended source is that, unlike a point source, there is no unique zeroth-order position that one could use in the grating equation. By judicious filtering in pulse-height space, one may find regions where there is enough of a pointlike behavior to permit the grating ARF to be used. However, how to handle a generic extended source in the presence of a diffraction grating is still an open question. It is hoped that the mathematical formulation of the extended source problem as given in §§ 3 and 5 will lead to better insights into these problems and ultimately to their solution.

This work also highlights some important practical considerations that should be taken into account in the design of astronomical data analysis software systems. For instance, to allow for the possibility of spatial variation in the underlying detector redistribution function, the software component responsible for the filtering of events should allow the user easily to filter *simultaneously* on both sky coordinates and detector coordinates. In addition, both filters would need to be passed to the program that generates the ARF. Finally, spectral fitting programs should be enhanced to facilitate the analysis of blended sources by handling the coupled integral equations in equation (24).

APPENDIX A

NUMERICAL CONSIDERATIONS

In this appendix, some "approximations" used for the practical computation of the ARF and grating ARF are discussed. In fact, these approximations are actually employed by the *Chandra* exposure map code suite for the generations of exposure

maps and ARFs. Since the code is freely available,[3] the actual implementation details will not be discussed here. The reader should also note that some of these approximations may only be valid for the *Chandra* telescope, which dithers, and for other missions one may have to resort to the full definitions given in the main body of the text.

A1. PERFORMING THE TIME INTEGRATIONS VIA AN ASPECT HISTOGRAM

The integrals over the observation time appearing in the equations for the ARF and the grating ARF can be quite computationally expensive, especially for long observation times. The general form of these integrals is given by

$$I = \int_0^\tau dt\, T(t)F(\alpha(t)) \,, \tag{A1}$$

where $\alpha(t)$ is an N-dimensional time-dependent vector that characterizes the dither of the telescope and the relative motion of its subsystems. For example, N is 3 if there is no internal movement, and the dither is characterized by the roll, pitch, and yaw of the telescope.

By multiplying the preceding equation by the identity

$$1 = \int d\alpha\, \delta(\alpha - \alpha(t)) \,, \tag{A2}$$

it trivially follows that

$$I = \int d\alpha\, \mathscr{H}(\alpha)F(\alpha) \,, \tag{A3}$$

where

$$\mathscr{H}(\alpha) = \int_0^\tau dt\, T(t)\delta(\alpha - \alpha(t)) \,. \tag{A4}$$

The quantity $\mathscr{H}(\alpha)d\alpha$ has a very simple interpretation. It represents the total amount of time, weighted by $T(t)$, that the point spent in the volume element $\Delta\alpha$ at α. Now, if the telescope dithers around some nominal pointing, and if the time-dependent internal motions due to, e.g., thermal expansion are small, then the point $\alpha(t)$ will be confined to some small volume in the N-dimensional space. This means that $\mathscr{H}(\alpha)$ will be nonzero only in that small volume and zero everywhere else. So, to compute the time integration over long observation times for the case of small dither amplitudes, it is often more efficient to compute the value of $\mathscr{H}(\alpha)$ and use it to evaluate equation (A3). In practice, the portion of the N-dimensional space where $\mathscr{H}(\alpha)$ is nonzero is subdivided into small volume elements $\Delta\alpha$. Then the discretized quantity $\mathscr{H}_{\alpha;\Delta\alpha} = \mathscr{H}(\alpha)\Delta\alpha$ is computed and used in a discretized version of equation (A3), i.e.,

$$I \approx \sum_\alpha \mathscr{H}_{\alpha,\Delta\alpha} F(\alpha) \,. \tag{A5}$$

For reasons that should be apparent, $\mathscr{H}_{\alpha;\Delta\alpha}$ is called the aspect histogram.

The major advantage of this approach for the case of small dither amplitudes is that there are likely to be many fewer terms to sum in equation (A5) than if a straightforward discretization were used to perform the time integration in equation (A1). Moreover, there are efficient algorithms based upon 2^N-ary trees for computing the aspect histogram. For example, the code for both the *Chandra* and *ROSAT* missions uses an octtree for $N = 3$. (For *Chandra*, the value of N used is 3 rather than 6 through the use of "effective" offsets.)

A2. COMPUTATION OF THE IMAGING ARF

The ARF is a complicated function requiring complete knowledge of the detector QE, mirror effective area, aspect solution, and the PSF. To compute it directly from equation (17) or from equation (32) in the case of a spatially varying RMF, one would need to carry out an integration over time as well as a two-dimensional integration over the sky region and do this for every point in the sky. Clearly, this is not practical, and in view of the fact that there will be uncertainties in the instrumental responses at this level of detail, such a calculation is unwarranted. Instead, one can make several simplifying approximations that permit the ARF to be computed in an economic manner.

As written, equation (17) is valid for any motion of the spacecraft, including slew. However, here it shall be assumed that one is dithering about some mean pointing and that the scale of the dither is small enough that any variations in the PSF and the mirror effective area on this scale can be neglected. Therefore, equation (17) will be approximated by

$$A_\Omega(\lambda, \hat{p}) \approx \frac{1}{\tau_{\text{eff}}} \int_\Omega d\hat{p}'\, \mathscr{F}_A(\lambda, \langle \hat{p}' \rangle, \langle \hat{p} \rangle)M(\lambda, \langle \hat{p} \rangle) \int_0^\tau dt\, T(t)Q(\lambda, \sigma(\hat{p}'_t, t)) \,, \tag{A6}$$

---

[3] See http://chandra.harvard.edu for more information.

where $\langle \hat{p} \rangle$ represents the time average of $\hat{p}_t$, i.e.,

$$\langle \hat{p} \rangle = \frac{1}{\tau_{\text{eff}}} \int_0^\tau dt \, T(t) \mathscr{R}(t) \cdot \hat{p} \, . \tag{A7}$$

Similarly, define

$$\langle Q(\lambda, \hat{p}') \rangle = \frac{1}{\tau_{\text{eff}}} \int_0^\tau dt \, T(t) Q(\lambda, \sigma(\hat{p}'_t, t)) \tag{A8}$$

to be the time-averaged value of the QE. Then one can write

$$A_\Omega(\lambda, \hat{p}) \approx M(\lambda, \langle \hat{p} \rangle) \int_\Omega d\hat{p}' \, \mathscr{F}_A(\lambda, \langle \hat{p}' \rangle, \langle \hat{p} \rangle) \langle Q(\lambda, \hat{p}') \rangle \, . \tag{A9}$$

The time-averaging over the dither motion has the effect of smoothing out any large variations in the QE over the region. In fact, this is the primary purpose of the dither. Now since $\langle Q(\lambda, \hat{p}') \rangle$ can be assumed to vary slowly over the region, and since $\mathscr{F}_A(\lambda, \hat{p}', \hat{p})$ is expected to rapidly go to zero as $\hat{p}'$ moves away from $\hat{p}$, $\langle Q(\lambda, \hat{p}') \rangle$ can be replaced by its average over the region and removed from the integrand. This leads to the result

$$A_\Omega(\lambda, \hat{p}) \approx f_\Omega(\lambda, \hat{p}) M(\lambda, \langle \hat{p} \rangle) \langle Q(\lambda) \rangle_\Omega \, , \tag{A10}$$

where

$$\langle Q(\lambda) \rangle = \frac{1}{\Omega} \int_\Omega d\hat{p}' \, \langle Q(\lambda, \hat{p}') \rangle \tag{A11}$$

is the average of $\langle Q(\lambda, \hat{p}) \rangle$ over the region and the PSF *fraction* in the region is given by

$$f_\Omega(\lambda, \hat{p}) = \int_\Omega d\hat{p}' \, \mathscr{F}_A(\lambda, \langle \hat{p}' \rangle, \langle \hat{p} \rangle)$$

$$\approx \int_\Omega d\hat{p}' \, \mathscr{F}_A(\lambda, \hat{p}', \hat{p}) \, . \tag{A12}$$

A3. COMPUTATION OF THE GRATING ARF

The grating ARF is defined by equation (57), rewritten here as

$$A_m(\lambda) = \left[\sum_{h=h_0(\lambda)}^{h_1(\lambda)} D_R(h, \lambda) \right] g_m(\lambda) \frac{1}{\tau_{\text{eff}}} \int_0^\tau dt \, T(t) Q(\sigma_G(\hat{p}_t, t), t) M(\lambda, \hat{q}_t) \, , \tag{A13}$$

where equation (8) has been used with the assumption that the RMF does not vary spatially. For a spatially varying RMF an additional spatial filter would need to be applied as was done for the imaging ARF to derive equation (32). In the above equation, $\hat{p}$ depends upon the source position and the wavelength according to

$$(\hat{p} - \hat{q}) \times \hat{n} = \frac{m\lambda}{d} \hat{l} \, . \tag{A14}$$

If the amplitude of the dither is small on the scale of the variations in the mirror effective area, then $M(\lambda, \hat{q})$ may be replaced by $M(\lambda, \langle \hat{q} \rangle)$ and removed from the integrand. Hence, the grating ARF may be approximated by

$$A_m(\lambda) \approx g_m(\lambda) \langle Q(\lambda, \hat{p}) \rangle M(\lambda, \langle \hat{q} \rangle) \sum_{h=h_0(\lambda)}^{h_1(\lambda)} D_R(h, \lambda) \, , \tag{A15}$$

where

$$\langle Q(\lambda, \hat{p}) \rangle = \frac{1}{\tau_{\text{eff}}} \int_0^\tau dt \, T(t) Q(\sigma_G(\hat{p}_t, t), t) \, . \tag{A16}$$

REFERENCES

Arnaud, K. A. 1996, in ASP Conf. Ser. 101, Astronomical Data Analysis Software and Systems V, ed. G. Jacoby & J. Barnes (San Francisco: ASP), 17
Chandra X-Ray Center, *Chandra* Project Science (MSFC), and *Chandra* IPI teams 2000, The *Chandra* Proposers' Observatory Guide, Version 2.0
Doe, S., Ljungberg, M., Siemiginowska, A., & Joye, W. 1998, in ASP Conf. Ser. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 157

George, I. M., Arnaud, K. A., Pence, B., & Ruamsuwan, L. 2000, Calibration Memo CAL/GEN/92-002
Ikebe, Y. 1995, Ph.D. thesis, Univ. Tokyo
Kahn, S. M., & Blissett, R. J. 1980, ApJ, 238, 417
Snowden, S. L., McCammon, D., Burrows, D. N., & Mendenhall, J. A. 1994, ApJ, 424, 714
Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., & O'Dell, S. L. 2000, Proc. SPIE, 4012, 2
White, D. A., & Buote, D. A. 2000, MNRAS, 312, 649