21. 'EVOLUTIONARY' THEORIES OF THE X-RAY BACKGROUND

M. J. REES*

Harvard College Observatory, Cambridge, Mass. 02138, U.S.A.

Abstract. It is unclear whether evolutionary effects must be invoked to account for the intensity of the X-ray background, nor whether the background could all be due to unresolved X-ray sources of the types already known. Intensity fluctuations on small angular scales are probably on the threshold of detectability. Future studies of these fluctuations – with improved sensitivity – should help to elucidate the nature of the background. The interpretation of such studies is discussed.

At the time of the 1969 IAU Symposium on X-ray astronomy, the existence of an apparently isotropic background was already well established. Its spectrum was known to follow a power law, at least over the range 1–20 keV; and the large-scale isotropy was established to a precision of around 10%. Because no class of extragalactic object seemed capable of accounting for the strength of the background (unless drastic evolutionary effects were invoked) some astrophysicists devoted much attention and ingenuity to devising emission mechanisms which might operate uniformly throughout intergalactic space, or at least in very extended regions such as clusters of galaxies. Examples of such 'diffuse' mechanisms are inverse Compton scattering of microwave background photons, and bremsstrahlung (thermal or non-thermal). The inverse Compton process, in particular, yielded a reasonable fit to the *shape* of the spectrum, but it was hard to reproduce the *strength* of the observed background without relegating its production to early epochs ($z \approx 2-5$), when more energy might have been available, and/or the emission mechanism more efficient. To account for the *hard* x-ray and γ-ray background, even larger redshifts – z up to ~ 100 in some theories – have been invoked. Though these 'evolutionary' models are not inherently implausible, nor even completely *ad hoc*, they are very speculative and difficult to test observationally. But they certainly make the X-ray background seem even more interesting to cosmologically-inclined theorists, because of the possibility that it may tell us something about even remoter epochs than the most distant known discrete sources, and about the thermal history and density of intergalactic matter. These ideas were reviewed by Setti and Rees (1970) at the Rome meeting (see also the article by Silk, 1970), and there do not seem to have been any really significant developments along these lines in the subsequent three years.

On the observational front, however, progress has been much more substantial. Two particular developments are especially important for increasing our understanding of the X-ray background: the number of identified extragalactic sources has risen from ~ 1 to ~ 10, which permits us to make a somewhat less conjectural estimate of how much of the background could come from different categories of

* On leave from Institute of Theoretical Astronomy, Madingley Road, Cambridge, England.

Bradt and Giacconi (eds.), X- and Gamma-Ray Astronomy, 250–257.

discrete sources; also, the limits on possible small-scale anisotropies in the background are now more stringent. We may expect rapid accumulation of data in both these areas to continue. It therefore seems worthwhile to devote the rest of this paper primarily to considering what such information tells us – or may soon tell us – about the nature of the background.

If the mean emissivity at frequency ν, per unit comoving volume, is $\mathscr{E}_z[\nu]$ (the suffix z indicating that $\mathscr{E}$ may vary with epoch) then the mean energy flux per unit frequency received by a detector with effective solid angle Ω is

$$I(\nu) = \Omega \int_0^{r_{max}} \frac{\mathscr{E}_z[\nu(1+z)]}{1+z} \, \mathrm{d}r. \tag{1}$$

r is the coordinate distance, defined so that $\mathrm{d}r = c\mathrm{d}t(1+z)$ (t being the epoch corresponding to redshift z), and r_{max} corresponds to $z=\infty$. When $\mathscr{E}$ is independent of z, and the X-rays are emitted with a power law spectrum $\mathscr{E}[\nu] \propto \nu^{-\alpha}$, (1) can be readily integrated. For an Einstein-de Sitter universe one obtains

$$I(\nu) = \frac{3}{3+2\alpha} \tau_H \mathscr{E}[\nu] \, \Omega, \tag{2}$$

where τ_H is the inverse Hubble constant. This means that, for $\alpha \approx 1$, we would get a correct estimate of the background – in the non-evolutionary case – by considering a static Euclidean universe of radius $\sim\frac{3}{5}$ the Hubble radius. The precise value of this factor depends on the cosmological model; but it is always around $\frac{1}{2}$ except in Lemaitre models with a long 'coasting phase', where it may be much larger. Another consequence of (2) is that about half of the integrated background comes from $z \gtrsim 0.3$.

If the X-rays are coming from discrete sources, then we can write

$$\mathscr{E}_z[\nu] = \int_L \varrho_z(L[\nu]) \, L[\nu] \, \mathrm{d}L, \tag{3}$$

$\varrho(L)\mathrm{d}L$ being the coordinate density of sources whose luminosity lies between L and $L+\mathrm{d}L$. Setti and Woltjer (this volume, p. 208) have used (2) and (3) to estimate the contribution to the background from different types of extragalactic sources. Despite the limited observational material and the consequent difficulty of making reliable calculations, it cannot be excluded that rich clusters of galaxies, and Seyfert galaxies, may each collectively contribute up to 10% of the total background, even if there is no evolution with z. The likely contribution from quasars is more uncertain still, but could well be *even more* if the evolution inferred from the steep number-magnitude relation is allowed for.

The recent data thus leave open the possibility that *all* the background could come from discrete sources. A special mechanism for diffuse X-ray emission may

https://doi.org/10.1017/S0074180900100439 Published online by Cambridge University Press

therefore be quite superfluous. Even if the bulk of the X-ray background *does* have a genuinely diffuse origin, the contribution from sources is almost certainly not negligible. But there is no reason why the X-rays from sources and the diffuse X-rays (if any) should have the same spectrum – completely different radiation processes may be involved. This means that we cannot ascertain what the spectrum of the hypothetical diffuse component is; and this uncertainty is unlikely to be resolved before the luminosity function and evolutionary behavior of extragalactic discrete sources has been established. Until that time, it is perhaps premature to dwell on the virtues and defects of rival 'evolutionary' theories.

Even if the background were all due to (say) the inverse Compton effect, we might expect some spatial irregularities in the electron distribution, which would give rise to inequalities between the intensities measured in different areas of sky. But *some* of the background is due to discrete sources, and this component will obviously produce a certain amount of 'graininess' if the background is studied with high angular resolution.

Provided that there is some length scale $l \ll c\tau_H$ such that the universe is statistically homogeneous if averages are taken over regions larger than l – and this is an article of faith among most cosmologists, to which the high overall isotropy of the microwave background, radio surveys, and the X-ray background lend strong support – the spatial structure of the X-ray emissivity can be adequately described by a 3-dimensional auto-correlation function. Knowledge of this function, and its z-dependence, allows us to calculate the 2-dimensional auto-correlation function for the brightness of the X-ray sky. A calculation of this kind has been given by Wolfe and Burbidge (1970) who considered whether the anisotropy limits found by Schwartz (1970) were compatible with a model where the X-ray emissivity displayed the same spatial autocorrelation as galaxies. This and related questions have also been discussed by Silk (1970), Schwartz *et al.* (1971), Fabian (1972), Webster (1972) and Craven and Sciama (1972). The essential features of Wolfe and Burbidge's work – for instance, the way the fluctuations depend on detector beam area and on the 'clumpiness' of the emission – can be clarified by a more simple-minded analysis.

Suppose that the background intensity is measured, in a particular bandwidth, in n non-overlapping areas of sky. For simplicity, we consider a detector with circular beam and uniform sensitivity over the whole beam area, but the arguments can be straightforwardly extended to detectors with more general and more realistic properties.

Consider first a model in which the background is entirely due to *randomly distributed point sources* with a single luminosity L, and there are no evolutionary effects. I shall follow the treatment given by Craven and Sciama (1972). The expected mean-square deviation $\overline{\delta I^2}$ from the expected flux (1) is given by

$$\overline{\delta I^2} = \Omega \int_0^{r_{max}} \varrho r^2 (L[(1+z)\nu])^2 \left\{ \frac{1+z}{r^2(1+z)^2} \right\}^2 \mathrm{d}r. \tag{4}$$

https://doi.org/10.1017/S0074180900100439 Published online by Cambridge University Press

The most important feature of (4) is that *the main contribution to* $\overline{\delta I^2}$ *comes from small* r. In one respect this simplifies things, because it means that the fluctuations can be adequately discussed in a Euclidean approximation. (In contrast to the *mean* integrated background which, by the familiar Olbers argument, comes mainly from 'cosmological distances'.) We observe, however, that the integral (4) actually diverges as $r \rightarrow 0$! This is because the contribution of the nearest source to the mean-square fluctuation depends on the inverse fourth power of its distance, whereas the probability that there *is* one source within a given distance varies only as (distance)$^{+3}$. This divergence is symptomatic of the highly skew and non-Gaussian character of the probability distributions involved, for which rms methods of analysis are not really appropriate. In practice, however, we are concerned merely with one realization of the ensemble of possible source distributions, in which the closest source will be at some definite distance $r_{\min}$. It is therefore legitimate to take $r_{\min}$, instead of zero, as the lower limit of integration. The median distance of the closest source expected within a solid angle $n\Omega$ is $(3 \log_e 2/n\Omega\varrho)^{1/3}$. Taking this value for $r_{\min}$ yields

$$\overline{\delta I^2} \simeq (\varrho\Omega)^{4/3} (L[\nu])^2 (n/3 \log_e 2)^{1/3}; \tag{5}$$

and so, defining δI as $(\overline{\delta I^2})^{1/2}$,

$$\frac{\delta I}{I} \simeq \frac{3 + 2\alpha}{3\varrho^{1/3}\tau_{\mathrm{H}}} \Omega^{-1/3} \left(\frac{n}{3 \log_e 2} \right)^{1/6}. \tag{6}$$

The appearance in this expression of n (the number of areas surveyed) perhaps at first sight seems rather surprising. It reflects the fact that there will generally be one area in which the nearest source is $n^{1/3}$ times closer than the average, and it is this area alone which makes the main contribution to $\overline{\delta I^2}$. This result again indicates that the single quantity $\delta I/I$ does not characterise the fluctuations adequately – however many areas are surveyed, its numerical value remains sensitive to how close we happen to be to the nearest source which is included.

If n is *not* very large, (6) reduces, very roughly, to

$$\delta I/I \simeq R/(c\tau_{\mathrm{H}}) \Omega^{-1/3} \tag{7}$$

when R denotes the mean spacing between neighboring sources. The fractional fluctuations from one beam area to another are thus roughly given by the inverse *cube* root of the number of sources within the beam. Or, in other words, $\delta I/I$ is approximately the ratio of the distance of the nearest source in the beam to the Hubble radius (i.e. the fractional contribution this source makes to I) and it is the variations in the actual distance of this nearest source that account for about half of the total $\delta I/I$.

If the sources span a range of luminosities, then (6) and (7) are modified only by an extra factor $(\overline{(L[\nu])^2})^{1/2}/\overline{L[\nu]}$ on the right hand side. The higher-luminosity sources obviously contribute proportionately more to δI than to I.

The above analysis is easily generalised to the case when the 'sources' are extended, with (say) a diameter d. This is relevant if we wish to test whether the X-ray background could be due to many galaxies in each cluster (the individual galaxies not being resolvable) or to intracluster gas. A source closer than $d/\Omega^{1/2}$ would subtend a a solid angle larger than the beam. Its contribution to I cannot exceed a definite finite value $\sim L\Omega/d^2$, however close it is, and this removes the divergence in (4). If Ω is so large that even the nearest source is smaller than the beam (as is marginally true for the experiments so far carried out, if clusters are the sources) then (6) and (7) still hold. But if Ω were smaller, so that the nearest sources *did* fill the beam, then one finds that

$$\delta I/I \simeq (3 + 2\alpha)/\tau_H (\varrho d)^{-1/2} \Omega^{-1/4}. \tag{8}$$

$\delta I/I$ then becomes $\propto \Omega^{-1/4}$ instead of $\propto \Omega^{-1/3}$. The major contribution to $\delta I/I$ is due to sources at distances $\sim d/\Omega^{1/2}$. If Ω is so small, or d so large, that even a source at a distance $\sim c\tau_H$ fills the beam, then the fluctuation amplitude is almost independent of Ω. This means that the largest value attained by $\delta I/I$ is

$$\sim (d^2 c\tau_H/R^3)^{-1/2}; \tag{9}$$

the quantity in parentheses being the mean number of sources intercepted by a line of sight extending out to the Hubble radius.

It is obviously inaccurate to approximate clusters as uniform spheres. It might be more realistic to assume that the emissivity falls off in a Gaussian fashion away from the cluster center. However the crude results given above are good enough to reproduce, within a factor two, the form of the autocorrelation function for the X-ray sky brightness calculated by Wolfe and Burbidge (1970). Note that we can always obtain an *upper limit* to $\delta I/I$ by using the point source approximation: if the sources are extended the fluctuations will always tend to be smeared out.

So far, none of the attempts to measure $\delta I/I$ has detected *any* fluctuations exceeding those attributable to the finite number of photons counted, and therefore only upper limits are available. Comparing these limits with (6), one finds that, in a non-evolutionary model, the number of sources out to the Hubble radius must be at least 10^7–10^8. This is just compatible with a mean source separation $R \approx 10$ Mpc, and thus with an origin in clusters of galaxies. But an origin exclusively in rich clusters already seems ruled out. Wolfe and Burbidge were led to an over-strong conclusion owing to an error later pointed out by Webster (1972).

Another type of model which can be tested against observational limits on $\delta I/I$ is what we might call an 'extreme evolutionary model', in which the background is attributed to sources which are *all* at cosmological distances and all have the same *apparent* brightness. The expected fluctuations then depend on the inverse *square* root (instead of the inverse cube root) of the number of sources in the beam. The present observations are thus compatible with extreme evolutionary models even if there are only 10^5–10^6 effective sources in the sky. Models in which the background comes from quasars or strong radio sources with large redshifts are thus still in the field.

If we wish to discuss evolutionary models in greater generality, it is better to approach the problem in a rather different way. This alternative viewpoint will also clarify both the limitations and the potential usefulness of studies of fluctuations in the background. Even if nothing is known about the luminosity function and z-dependence of the sources, one can in principle determine their number-versus-intensity relation – i.e. the function $N(S)$, which denotes the number of sources per steradian whose intensity exceeds S – and, if the sources are randomly distributed, $N(S)$ contains all the information relevant to the magnitude and probability distribution of $\delta I/I$. We have

$$I = \Omega \int_{\infty}^{0} S\,(\mathrm{d}N/\mathrm{d}S)\,\mathrm{d}S \tag{10}$$

and

$$\overline{\delta I^2} = \Omega \int_{\infty}^{0} S^2\,(\mathrm{d}N/\mathrm{d}S)\,\mathrm{d}S\,. \tag{11}$$

In the particular case when $N(S) \propto S^{-3/2}$ (with a low-S cut-off at the intensity appropriate to a source at the Hubble radius), these two equations are precisely equivalent to (2) and (4). But $\overline{\delta I^2}$ could obviously be evaluated for *any* $N(S)$. In particular, one may consider the case $N(S) \propto S^{-\beta}$ $(\beta > 1)$, with a truncation at some intensity $S = S_{\mathrm{min}}$ to ensure convergence of (10). If $\beta > 2$, the situation resembles the 'extreme evolutionary model' already mentioned: the fluctuations (as well as, of course, the integrated flux) are dominated by sources with $S \approx S_{\mathrm{min}}$. For $1 < \beta < 2$, however, the fluctuations are due mainly to the high-S sources. One can derive a generalised version of (6) in which the dependence of $\delta I/I$ on the beam area and the number of regions surveyed is $\Omega^{1/\beta - 1}\, n^{1/\beta - 1/2}$.

Substantially more would be learnt about the nature of the background if one could determine not merely $\overline{\delta I^2}$, but the whole probability distribution $p(\delta I)$ of δI. This might be feasible if high sensitivity measurements could be made in a large number of areas of sky. The form of $p(\delta I)$ in point source models can be calculated in terms of $N(S)$. For $N(S) \propto S^{-\beta}$ $(1 < \beta < 2)$ it is skew, with a sharp cut-off on the negative side, but a long tail with the approximate form $p(\delta I) \propto (\delta I)^{-(\beta+1)}$ on the positive side, which reflects the probability of there being one exceptionally bright source within the beam. One could calculate the modifications in $p(\delta I)$ arising from the finite extent of the sources. Thus in principle the fluctuations contain information on the linear dimensions, correlation length, *and* evolutionary behavior of the contributors to the background. This procedure is essentially the same as the so-called $P(\mathrm{D})$ technique of Scheuer (1957) and Hewish (1961), whereby the form of $N(S)$ could be inferred for radio sources below the confusion limit by analysing the 'noise' from a radio interferometer.

A sufficiently large positive value of δI would in practice be attributed to a resolved source. This suggests that a more realistic (and more precisely measurable) value

https://doi.org/10.1017/S0074180900100439 Published online by Cambridge University Press

of $\delta I/I$ could be defined if $p(\delta I)$ were truncated at a certain level, any larger-amplitude fluctuation being deemed a 'detected source'. It is then interesting to consider the following problem. Suppose one wishes to know whether a particular class of objects – e.g. quasars or clusters – can account for the X-ray background. One could *either* check whether the closest members of the population would be strong enough sources to have shown up in surveys, *or* calculate whether the number of objects in the class is large enough that the small scale fluctuations in the background would be undetectable. Which of these tests is the most sensitive, bearing in mind that similar detectors, observing for similar lengths of time, would be utilised in each case? Craven and Sciama (1972) have discussed this question. The precise answer depends on the characteristics of the detector, and on the criteria that must be fulfilled before a source is said to have been detected, but in non-evolutionary models – perhaps not surprisingly – it turns out that the two procedures are more or less equally sensitive. When evolutionary effects are important, however, the balance tilts in favor of the fluctuation technique. This is because the power-law tail of the function $p(\delta I)$ steepens with increasing β (and in fact disappears for $\beta > 2$), making it less likely that individual sources will be detected.

If we compare recent developments in X-ray astronomy with the early history of *radio* astronomy, it is striking that, whereas investigations of the background have been regarded as an important part of X-ray astronomy, the radio background never achieved such a major role. This is partly because the isotropic radio background is swamped by emission from the Galactic Disk, but it also indicates a genuine difference in the type of extragalactic source which dominates in the respective frequency bands: it was clear from the very early days of radio astronomy that some discrete sources were so powerful that they could be readily detected at cosmological distances (and there is no doubt that much of the isotropic radio background is attributable to such sources); most of the extragalactic X-ray sources so far detected are, however, at distances $\ll (c\tau_H)$, and this automatically implies that only a minority of the X-ray photons reaching us come from resolved sources.

Another way of investigating the background and elucidating the nature of the emission mechanism is by studying the spectrum. It is obvious that we would not expect to see any lines (except perhaps in a Lemaître universe), because of the range of redshifts which would contribute. In principle it might be possible to detect 'edges' in the spectrum, and to infer the evolutionary properties of the emission from the shape of the spectrum at wavelengths longward of the edge, but in practice such observations would always be very ambiguous (see, for example, Tinsley, 1972). Note also that it would be quite possible for bremsstrahlung from gas with a range of temperatures to yield an integrated spectrum that mimics a power law. Thus the basic question of whether the emission is thermal or non-thermal remains open still.

Improved study of the small-scale isotropy seem much more likely to yield useful clues as to the nature of the background. The review I have given here is obviously over-simplified, being intended merely to illustrate the nature of the problem. In practice, there are probably several quite different processes, each making a significant contribution to the total background. Since the different contributions need not have the same spectrum, one would ideally like to have measurements at different wavelengths of the same regions of sky. (Wavelength-dependent absorption might also be important, especially if the background comes from large redshifts). Another complication is that the fluctuations could be dominated by a component which makes a relatively minor contribution to the total intensity, if that component has a larger correlation-length than the others. A further possible cause of confusion is that our own Galaxy may contribute significantly to the fluctuations (especially on larger angular scales), even though the overall isotropy tells us that it cannot contribute more than a few per cent of the total background.

Acknowledgements

I am grateful for helpful discussions with many colleagues, and would especially like to thank P. Craven, A. C. Fabian, J. E. Felten, H. Gursky, C. Hazard, D. W. Sciama, and A. M. Wolfe.

References

Craven, P. and Sciama, D. W.: 1972, in preparation.
Fabian, A. C.: 1972, *Nature Phys. Sci.* **237**, 19.
Hewish, A.: 1961, *Monthly Notices Roy. Astron. Soc.* **123**, 167.
Scheuer, P. A. G.: 1957, *Proc. Cam. Phil. Soc.* **53**, 764.
Schwartz, D. A.: 1970, *Astrophys. J.* **162**, 439.
Schwartz, D. A., Foldt, E. A., Holt, S. S., Serlemitsos, P. J., and Bleach, R. D.: 1971, *Nature Phys. Sci.* **233**, 110.
Setti, G. and Rees, M. J.: 1970, in L. Gratton (ed.), 'Non-Solar X- and γ-ray Astronomy', *IAU Symp.* **37**, 352.
Silk, J.: 1970, *Space. Sci. Rev.* **11**, 671.
Tinsley, B. M.: 1972, *Astrophys. Letters* **10**, 31.
Webster, A. S.: 1972, *Nature* **238**, 20.
Wolfe, A. M. and Burbidge, G. R.: 1970, *Nature* **228**, 1170.