Title
Backgrounds and the Big Bang: Some Extracts from Their History

Journal
Symposium - International Astronomical Union, 168(168)

ISSN
0074-1809

Author
Trimble, Virginia

Publication Date
1996

DOI
10.1017/s0074180900109891

BACKGROUNDS AND THE BIG BANG: SOME EXTRACTS FROM THEIR HISTORY

VIRGINIA TRIMBLE
Astronomy Department, University of Maryland,
College Park, MD 20742
& Department of Physics, University of California,
Irvine CA 92717, USA

ABSTRACT. The first astronomical background firmly established was that of cosmic rays. Photons at various wavelengths came later; and in some bands we have not year clearly peeled away all the sources to see a true background, if one even exists. All known backgrounds are astrophysically important and at least several cosmologically so. The path by which the standard hot big bang came to be generally regarded as standard is littered with the detritus of mistaken impressions, misunderstandings, and missed opportunities.

1. Introduction

The present symposium had its origins in two initially separate and rather different sorts of proposals. The first, put forward by M. Hanner and S. Bowyer, was for a discussion of all known diffuse backgrounds, many of which, like zodiacal light, have very little to do with cosmology. The other was for a meeting on the robustness of the standard hot big bang and alternatives to it, proposed by M. Kafatos and Y. Kondo. These entered into the mysterious interior of the IAU executive committee and came forth as a single symposium, bearing some resemblance to the mythical animals of Dr. Suess.

The remaining sections deal with (2) the discovery and significance of the backgrounds, (3) the standard HBB along the lines of "how we know that" it is standard, (4) items within that picture still to be determined, (5) the cosmic tolerance quotient for deviations from it, and (6) a compromise answer to whether the big bang is important.

2. Known and Expected Backgrounds

The table shows all of these I could think of, electromagnetic and other. Henry (1991) displays the photon backgrounds in an insightful graph. The cautious reader will note that the observed infrared consists of inter-stellar cirrus and thermal and scattered (zodiacal light) contributions from interplanetary dust, not emission from hypothetical primeval

M. Kafatos and Y. Kondo (eds.), Examining the Big Bang and Diffuse Background Radiations, 9–16.

galaxies, while detection and significance of the ultraviolet and EUV ones are in considerable disorder (hence the accompanying alternative review by Bowyer, 1991). There is no space to discuss all the wavebands, and I focus here on the gamma ray background (the ~ 100 MeV part of which, due to pion decay, was predicted by Hayakawa etc al. 1958) and the relict microwaves.

KNOWN AND EXPECTED ASTRONOMICAL BACKGROUNDS

TYPE	D/S	WHO/WHEN	Cosm Sign?
Particles (cosmic rays)	D	Hess, Bothe, Kohlhörster 1911-29	½
Neutrinos			
- 1.9 K	D	To be discovered	Yes
- high energy	S		½
Gravitational radiation			
- early universe	D	To be discovered	Yes
- core collapses etc.	S		½
Gamma rays			
~ 100 MeV (pions)	D	Kraushaar, Clark, Garmire 1961-68	No
~ 1 MeV	S?	Arnold et al. 1962	½
X-rays			
- from hot IGM	D	Known to be small	Yes
- from AGNs etc	S	Giacconi et al. 1962; Bowyer, Friedman 1963	½
Soft X-ray to EUV			
- ~0.1 keV (local ISM)	D	Yentis, Novick, Vanden Bout 1972. Many later confirmations	No
1216 - 2500 A			
- dust-scattered starlight	D	See reviews, Henry (1991), Bowyer (1991)	No
- extragalactic	S?		?
Optical			
- airglow + Zodiacal light	D	- Known to ancients; first photoelectric measurement van Rhijn 1921	No
- stars and galaxies	S	- Seares, van Rhijn et al. 1925; Roach 1960ff	½
- residual	D?	- To be discovered	?
Infrared			
- 10-25μ (warm interplan dust)	D	IRAS confirmation	N
- 60-100μ (cold interst. dust)	D	IRAS confirmation	N
- 1-200μ from galaxy formation	S	DIRBE - last word is not yet in	Y
Submillimeter excess (era of confusion)	X	Nagoya-Berkeley balloon to COBE	N!
Radio			
- mm to cm	D	Penzias, Wilson 1964-65 (Ohm 1961 and others)	Y
- cm to decameter (galactic synchrotron and thermal)	D	Jansky 1933; Reber 1940	N

Cosmic rays (indicated in the table as D, for truly diffuse, as opposed to S, for sum of courses) were initially thought to be very high energy photons. Their particulate nature was established from their extraordinary penetrating power, in a classic one-page paper by Bothe and Kohlhörster (1929). No significant knowledge of German is required to make sense of the change from "Gammastrahlung" in the first sentence to "Korpuskularstrahlen" in the last. The "Cosmological Significance" column for cosmic rays says neither Yes or No, but ½, meaning that not all sources (especially for the highest energies) have been identified, and there may or may not be information about the large scale structure and evolution of the universe to be learned from them.

No extra-solar-system gamma rays were seen until the 1960s, and the numbers were initially small. Kraushaar and Clark (1962), for instance, speak of "the remaining 22 events, which came from a variety of directions in space...", accounting for the rumor that gamma ray astronomy is the field where one photon is a discovery, two is a spectrum, and three is the Rossi Prize. Jim Arnold, who found the 1 MeV flux in Ranger 3 data early in his career, recently retired as director of Calspace, indicating the length of a generation in the filed. Additional early history is described in Fazio's 91967) review article -- written before the first non-solar source had even been firmly established.

The early history of the microwave background is closely coupled with that of big bang nucleosynthesis. McKellar's (1941) interpretation of CN absorption lines in a spectrum taken by W.S. Adams is widely known. But the real stinger is Herzberg's (1950) discussion of the result in his classic volume, Spectra of Diatomic Molecules, "...a rotational temperature of 2.3° K follows, which has of course only a very restricted meaning." the real (and highly unrestricted!) meaning just eluded at least three pre-Penzias-and-Wilson measurers of microwave sky temperatures. Woltjer (elsewhere in this volume) alludes to a French measurement, and a symposium participant mentioned a Russian one that may be earliest of all. I note here the work of Ohm (1961), because its misinterpretation led Zeldovich (1962, 1963) and at least some of his colleagues to confine their attention to a cold big bang for some time.

The problem was at least partly a verbal one. The measured sky temperature was very closely what had been predicted, without any cosmological component. But as Ohm (1994) has explained, this is exactly as it should be -- the phrase "sky temperature" meant the radiation attributable to thermal emission from the earth's atmosphere (a cosecant θ component). the total coming into the whole system is the "system temperature", and, as Ohm's Table II shows, this was about 3.3° K larger than the sum of all the contributions he could think of. The excess was only about 1σ and was, correctly, reported as an upper limit.

Better calibration gave much higher significance to the "measurement of excess antenna temperature at 4080 Mc/S" by Penzias and Wilson (1965). A radio astronomer participant explained that it is not just a coincidence that their observing frequency was exactly 10 times the 408 Mc/s frequency of some Cambridge source surveys, but rather a result of the way radio frequencies are assigned for various purposes.

The "Alpha Beta Gamma" (Alpher et al. 1948) paper is a real one, though very short, published on April 1st. I own an authentic reprint, "deaccessioned" by Sir. Fred Hoyle in about 1970. This was part of an extensive office clearing effort and probably has no significance for the history of science.

The first relict radiation prediction (Alpher and Hermann, 1948) appeared later the same year. Like the Gamow et al. nucleosynthesis discussions, it presupposes a primordial soup of pure neutrons, as does their more detailed 1949 treatment of matter and radiation densities in the universe. The correct initial conditions, with protons and neutrons in thermal equilibrium, were first treated by Hayashi (1950). In this very little read (or cited) paper, the author expresses some hopes that the mix of protons, neutrons, deuterons, tritons, and helium nuclei may permit bridging the A = 5 and 8 gaps to form carbon and other heavy elements.

3. How the Standard Hot Big Bang Got That Way

The established facts of observational cosmology can still be counted with your shoes one. First, apparent brightness falls like $(\text{distance})^{-2}$, as much have been noticed by the first paleolithic tribe to carry their campfires from place to place. Second, light travels at a finite speed, first measured by the Dane Ole Roemer (while in Paris), from timing of eclipses of the moons of Jupiter. He also built the first transit circle instrument, in case anyone is interested.

Third comes the large wavelength shifts of the spiral nebulae, first recorded for M31 by Slipher in 1912. His spectrogram, which took two December nights to record, is reproduced in volume 2 of the classic Russell, Dugan and Stewart text and shows, rather dimly, the F and G bands and the calcium H and K absorption lines (and no emission features). Theory then intrudes, with Einstein and general relativity followed in the same year (1916) by Friedmann's solutions, describing a homogeneous, isotropic universe. It could expand, contract, or sit still (but only unstably). If GR is the right (classical) description of gravity -- we know it is wrong in quantum mechanical limits -- then theoretical freedom is much restricted. Our confidence in GR derives from solar system (weak field) tests, but also from strong field effects in binary pulsars, which GR describes better than the available alternatives, even ones motivated by the desire for a quantum theory and unification of all the forces (Taylor et al. 1992).

The actual existence of galaxies outside the Milky Way was established only in 1924, when Hubble identified Cepheid variables in M31, in some sense our fourth important fact. Meanwhile, Slipher and then Hubble and Humason were busily adding to the body of measured wavelength shifts. Lundmark, Witz, Stromberg, and Robertson were among those early tempted to plot the shifts vs. distances to the galaxies, in about 1925. With no allowance for Malmquist bias and angular diameter as the common distance indicator, the resulting functional form was typically a quadratic, as predicted for an empty, de Sitters, universe. At least one contemporary writer on the subject continues to find his quadratic.

The fifth and most important pre-war fact of observational cosmology is the linearity of the redshift-distance relation, put forward in 1929 by Hubble. His original drawing, from the Proceedings of the National Academy of Sciences is widely reproduced in astronomy history books. The velocity units were, accidentally, km rather than km/sec, and the full range is only 1000. With maximum galaxy distances of 2 Mpc, Hubble arrived at $H = 536 \pm 25$ km/s/Mpc. His fractional error bars are about the same size as those reported by van den Bergh and Tammann (elsewhere in this volume).

Hubble himself oscillated between universal expansion and tired light explanations of his correlation. In principle, the latter can be ruled out observationally. Tired light predicts that surface brightnesses will scale as $(1+z)^{-1}$ vs. $(1+z)^{-4}$ for expansion, and that time scales of variable phenomena will not be time dilated as they would be by true cosmic expansion. The surface brightness test (recently pursued by Sandage) remains mired in observational difficulties. Two (and only two) type Ia supernova at $z = 0.3$ and 0.5 seem to show time dilation, and it has arguably also been seen in gamma ray busters (but other explanations of time scale vs. flux correlations are possible). It is, however, fair to say that there is no conventional physical mechanism for tired light: the Feynman graphs for it sum to zero.

If (and perhaps only if) you accept both GR and expansion, then the universe has a hot, dense state in its past. The time scale for "past" in independently established by considerations of stellar evolution and nucleocosmochronology. It has been known at least crudely since 1905, when Rutherford showed that some earth rocks with radioactive content are at least 10^9 years old.

This hot, dense stage provides, of course, a simple explanation for both the 2.7 K background radiation and for the abundance of helium (etc.) in unevolved objects. If you try to produce the helium in galaxies over the characteristic 10^{10} year lifetime to their stars, you end up with a ratio of luminosity to baryon mass of about 10 in solar units, far larger than in the galaxies we see. The problem with accounting for the radiation in a non-evolving universe is not so much the energy required as the observed isotropy and black body spectrum. If your try to achieve these with discrete sources and reprocessing, your universe is likely to be opaque to microwaves at redshifts where we see sources unabsorbed. A separate new constraint comes from recent observations that the small fluctuations of background intensity in the sky also have a black body character. The importance of this in ruling out absorption and reemission or wavelength-dependent scattering (vs. electron scattering in the conventional hot big bang) as the mechanism for thermalization and isotropization is emphasized by Rees (elsewhere in this volume).

4. To Be Determined.

The hot big bang may be standard, but its parameters clearly are not. Factors of two or more still surround Hubble's H, q (deceleration), Ω and Ω_b (total and baryonic density), χ (cosmological constant), the age of the universe, and its radius of curvature (or k). That we have been asking for 40-some years has become something of an embarrassment to the astronomical community and is unfortunately sometimes used to support claims that cosmologists haven't a clue what they are talking about. Obviously most Symposium 169 participants disagree!

I would like, however, to put in a plea here for both observers and theorists to be clear about which parameter they are trying to measure. Geometry of the universe, but not about deceleration, unless you have assumed a value of χ, and conversely.

Other items not yet established include, notoriously, the nature of (at least much of) the ubiquitous dark matter and the extent and topology of the largest-scale structures and deviations from Hubble flow in the present universe and how those have changed with redshift.

5. The Limits of Cosmic Tolerance

How much deviation from the standard hot big bang can you take without getting nervous? I list these in order of increasing nervousness on my part, while reminding you that inflation (hence $\Omega = 1$, the Harrison-Zeldovich spectrum, and so forth) is not really part of the standard model (according to Rees), or at any rate not yet (according to Turner).

Deviations that would be interesting but not threatening include $\Omega \neq 1$ and $\chi \neq 0$. Fluctuations (either in the background radiation or assumed in the initial density field to make galaxies) that are non-Gaussian, not Harrison-Zeldovich, not adiabatic, etc. are also OK. Seeding by strings for galaxy information, tensor (gravitational radiation) contributions to irregularities in the background radiation, and inhomogeneous nucleosynthesis are even expected in some versions of the "standard" and are similarly non-distressing.

We have gradually got used to larger and larger cluster/void structures at least up to 100 h^{-1} Mpc and to streaming and coherent peculiar velocities over similar length scales up to 1000 km/sec or so. Varying levels of skepticism seem to have greeted Brent Tully's super-duper clusters of 10^{18} M_θ, the Broadhurst et al. indications of regularity at large scales, and the Lauer and Postman report of very extensive streaming. Much of the skepticism is probably justified by the limited data; some is probably nervousness, expressed in the usual way of conservative scientists. I would not myself be horrified to learn that part of all of our local 3K dipole is really anisotropic of the universe (but others would be).

The universe has non-zero baryon number (or we would not be here to talk about it). Standard big bang nucleosynthesis sets the lepton number equal to baryon number (or to zero, depending on exactly what you do with your neutrinos, and the difference does not matter). No current observations, however, exclude an unbalanced density of photons. Large lepton number in this sense makes a major difference to the amounts of H^2, He^4, Li^7 etc. coming out of the hot big bang, and this particular non-standard case may deserve investigation with a state-of-art nucleosynthesis code.

Red (or blue) shifts due neither to Hubble expansion, to ordinary physical motion, nor to strong gravitational fields cross the border into the intolerable for most cosmologist. I can imagine incorporating them into some rational model only if they carried a clear signature (apart from the red/blue shift itself). Non-constant / might be one thing to look for.

The standard model breaks down ever more completely if the early universe was cold (though dense), redshifts are quantized or not linear in distance, or there was no dense phase at any temperature.

6. Is the Big Bang Important?

Ask a silly question (my mother used to say) and you get a silly answer. Thus an analogy may be useful. If you believe (as many apparently do) that the universe was created in 4004 BC or thereabouts (12 noon on 29 October is an optional refinement), then you can have a successful career in medicine and physiology, laboratory physics or chemistry, engineering or mathematics. But you had better stay away from astronomy, the geosciences, evolutionary and ecological biology.

By the same token, if we are all wrong and there was never a hot dense stage (or, alternatively, if there was, but you don't believe it), then some parts of astronomy are still perfectly OK. You can work on the solar system, cosmic rays sources and acceleration (though watch out for the highest energy ones not confined by galactic magnetic fields), formation, structure and evolution of stars (at least within the Local Group), and physics and chemistry of the interstellar medium. The possible range of initial helium abundances begins to produce difficulties in studying stellar populations in other galaxies, even nearby. And if redshift is not an accurate guide to distance (hence luminosity and lookback time) then all bets are off for any investigation of galaxy (normal or active) formation and evolution, and even the interpretation of colors in terms of stellar populations at large redshift.

It is perhaps appropriate to end this section with quotes from two senior astronomers who have questioned the correctness of the standard hot big bang. Sir Fred Hoyle is supposed to have said (in a cosmological context), "I can see no reason to disbelieve something just because it is impossible." (to which Oort is said to have replied, "I can think of no better reason."). And I have heard Thomas Gold say (in another context), "If we are all going the same direction, it must be forward."

Acknowledgements

Over the years, I have been privileged to hear astronomers from Arp to Zwicky discuss these issues. Some of them (starting with G.R. Burbidge) would undoubtedly want to go on record as saying that they do not at all agree with how I have interpreted their remarks. A special thanks this time to Ed Ohm for copies of his classic papers and explanations of how the data should be thought about. Since maternal great-grandmother's father was an Ohm, I have hopes that we might be related.

References

Alpher, R.A., H. Bethe & G. Gamow 1948. Phys. Rev. 73, 803
Alpher, R.A. & R. Herman 1948. Nature 162, 775 1949. Phys. Rev. 75, 1092
Arnold, J.R. et al. 1962. JGR 67, 4878
Bothe, W. & W. Kohlhörster 1929. Phys. Zeit. 30, 516
Bowyer, S. 1991. ARA&A 29, 59
Fazio, G.G. 1967. ARA&A, 481
Hayakawa, S. et al. 1958. Prog. Theor. Phys. Sup. 6,1
Hayashi, C. 1950. Prog. Theor. Phys. Sup. 6, 1
Henry, R.C. 1991. ARA&A 29, 89
Herzberg, G. 1950. Spectra of Diatomic Molecules (Van Nostrand) p. 496
Jansky, K. 1933. Nature 132, 66
Kraushaar, W.L. & G.W. Clark 1962. Phys. Rev. Lett. 8, 106
McKellar, A. 1941. Publ. DAO 7, 251
Ohm, E.A. 1961. Bell System Tech. J. 60, 1068 1994. Personal communication
Penzias, A.A. & R.E. Wilson 1965. ApJ 142, 420
Reber, G. 1940. ApJ 91, 621; Proce. IRE 28, 681
Seares, F.H., P.J. van Rhign, Mary C. Joyner & Myrtle L. Richmond 1925. ApJ 62,320
Taylor, J.H. et al. 1992. Nature 355, 125
Yentis, D.K., R. Novick & P. Vanden Bout 1972. ApJ 177, 365 & 375
Zeldovich, Ya. B. 1962. JETP 43, 1561 1963. Usphekii 6, 475