Great Experiments in Physics

6. Discovery of the Cosmic Microwave Background Radiation

Amit Roy

Amit Roy is at the Nuclear Science Centre, New Delhi, building a superconducting linac booster for the pelletron accelerator. He spent over two decades at the Tata Institute of Fundamental Research, Mumbai investigating nuclei using accelerators and probing symmetries in physics. His hobbies are books and music.

Previous articles in this series:

1. Discovery of Transistor Effect that changed the Communication World, *Resonance*, Vol.3, No. 9, 1998.
2. Tunnelling in Superconductors: The Josephson Effect. *Resonance*, Vol.3, No.11, 1998.
3. Measuring Diameters of Stars: The Hanbury Brown–Twiss Effect, *Resonance*, Vol.4, No.5, 1999.
4. Birth of Quantum Electronics – Masers, *Resonance*, Vol.4, No.9, 1999.
5. Birth of Quantum Electronics – Lasers, *Resonance*, Vol.4, No.10, 1999.

In 1965, Arno Penzias and Robert Wilson of Bell Telephone Laboratories in USA discovered weak microwave radiation coming from all directions which was soon realised to be a relic from a hot dense early phase of our Universe – the 'big bang'. This article focuses on the key experimental techniques and innovations which made this momentous – and accidental! – discovery possible.

Chance enters the realm of scientific discoveries frequently enough so that the word 'discovery' often carries the connotation of accident. Naturally the role of chance is different in different fields. The whole field of 'radio astronomy' arose by such an accident, when Karl Jansky of Bell Telephone Laboratories sought the cause of a particular kind of radio static and found that it consisted of noise being received from near the direction of the centre of our galaxy. From that beginning, radio astronomy has grown into a well-developed field and has been the means of a number of discoveries, some of which are not yet fully understood.

Among these discoveries, the one that has perhaps the greatest implication for cosmology, is the observation of a background microwave radiation, apparently universal in distribution and well represented as the spectrum of blackbody radiation at a temperature of 2.7K. Its discovery by Arno A Penzias and Robert W Wilson in 1965 was truly a serendipitous discovery without any experimental or theoretical clues known to them at that time. The cosmic microwave background radiation is considered to be a relic of the 'big-bang' at the beginning of the universe. Among the theories of development of the universe, one class is characterised by the assumption that the universe started in a spatially small, highly condensed state. The initial stages, at

Box 1. Big Bang and Steady State Theories

The theories of cosmology can be grouped into two categories, first being the evolutionary theories and the second, the steady state.

There have been many contributors to the development of evolutionary theories of the universe including Einstein, de Sitter, Friedman. The idea of the universe starting from a very small region with an explosive event is attributed to Belgian astrophysicist Abbe Lemaitre and was championed by George Gamow. The observed expansion of the universe follows from the initial Big-Bang with the local density constantly changing. The universe could be open or closed, i.e.; the expansion could go on forever or halt at some stage depending on the amount of matter in the universe. Initially in the extremely dense phase matter would be in the form of plasma and in equilibrium with radiation. At a stage in the expansion, when atoms formed, radiation decoupled from matter and further expansion brought the radiation to an equivalent temperature of 3 K.

In the steady-state theory, the universe had no beginning nor has any end. It is assumed to be uniform in space and also unchanging in time when viewed on a sufficiently large scale This idea was put forward by Herman Bondi, Thomas Gold and Fred Hoyle. Jayant Narlikar in our country has been a proponent of this school of thought and one of the architects of the quasi steady state theory. To take into account the expansion of the universe, the idea of continual creation of matter is proposed to keep the mean density of matter unchanged. The universe is not unchanging in detail. Individual galaxies age and move apart from each other due to the expansion and newly created matter form galaxies in the intervening spaces. The rate of this creation is, of course, very low, owing to tenuous distribution of matter in the universe and the slowness of expansion. To put numbers, in the whole volume of earth it would amount to the addition of a tiny particle of dust every million years. So direct experimental evidence of this creation would be hard to find. One attractive feature of this group of theories is that there is a clear-cut mathematical basis for them.

least, of the development from this state were so rapid as to be explosive, so that the theory is often referred to as the 'big-bang' theory. G Gamow and his colleagues, R A Alpher and R C Herman recognised in their pioneering work in 1946 that the early stages of such a universe would be dominated by blackbody radiation, and that the remnant of such radiation should still be present. They estimated that expansion of the universe would have reduced its temperature to a value in the neighbourhood of 5 K. The verification of the existence of such a remnant would thus give strong support to the big-bang theory.

Penzias and Wilson were involved in measurements in radio astronomy with the 20 foot diameter horn reflector antenna at

Bell Telephone Laboratories. Their aim was to measure the absolute radiation intensity from our Milky Way Galaxy at high latitudes i.e., in the galactic halo region. Existing low frequency measurements in 1963 indicated that the brightness temperature of the halo would be less than 0.1 K at 7 cm wavelength. Thus a background measurement at 7 cm was expected to produce a null result.

Every material body at a finite temperature emits radiation. In the situation where the radiation reaches full equilibrium with matter, its spectrum is characteristic of the temperature of the body. But a radio telescope pointed at the sky receives radiation not only from space, but also from other sources including the ground, the earth's atmosphere and the components of the telescope itself. An antenna collects radiation from a desired direction to the exclusion of other directions. So by first pointing the antenna to the direction of a source of interest and then to a background region nearby, the contribution from the source can be identified. The difference of the two contributions would subtract out the local noise. The 20' horn antenna was built by A B Crawford in 1960 to be used with an ultra low-noise communications receiver for signals bounced from the Echo satellite. *Figure* 1 shows a picture of this antenna. The basis of the measurement is the fact that any electrical circuit element at a temperature above absolute zero

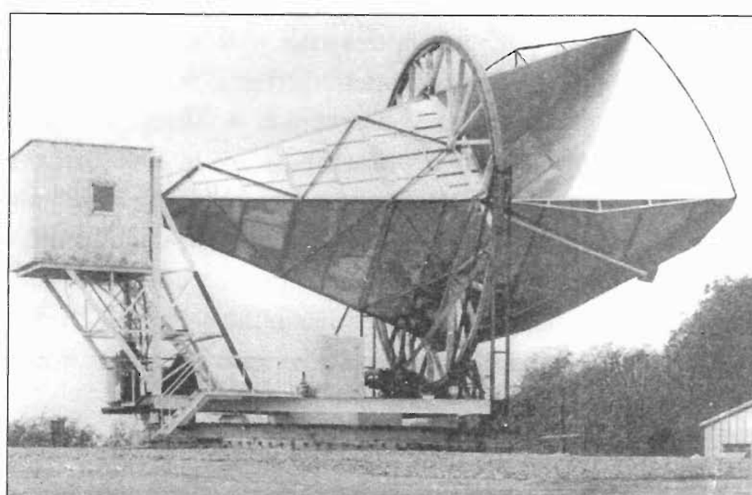

Figure 1. A picture of the 20' horn antenna used by Penzias and Wilson for their discovery. (Reproduced from Bell Systems Technical Journal, Vol 40, p.1097, 1961 with permission from Lucent Laboratories.)

Box . 2. Blackbody Spectrum

A blackbody can be defined as one, which absorbs 100% of the radiation falling on it. Substances such as lampblack (or soot) come very close to being perfect blackbodies, but not quite 100% absorption. To get a close description of a blackbody we can imagine a box with a small aperture. If the dimension of this aperture were made sufficiently small, the box would represent a perfect blackbody. Any radiation entering the small aperture would have a chance of emerging only after repeated reflections. At each reflection, part of the energy is absorbed and for very large number of reflections the radiation is unable to emerge from the box, i.e. is completely absorbed. The emission from a blackbody can be studied with the help of such a box and increasing the temperature of the box. The spectrum of emissions from such a body can be described by a Planckian distribution. This distribution can be written as,

$$I(f) = (2\pi h f^3/c^2)/(\exp(hf/kT)-1),$$

where f is the frequency, h is Planck's constant, k is Boltzmann's constant, c is the velocity of light and T is the temperature. Such a distribution for a temperature of 2.7 K is shown in the *Figure* below.

The higher the temperature, the more distinct the maximum of the radiation curve, and the wavelength at the maximum of the curve is given by the Wien displacement law, $l_m = 2.886 \times 10^{-3}/T$, where l_m is in metres and T is in degrees Kelvin. Another feature of the radiation curve is that the area under the curve is proportional to the fourth power of the temperature as given by the Stefan–Boltzmann law.

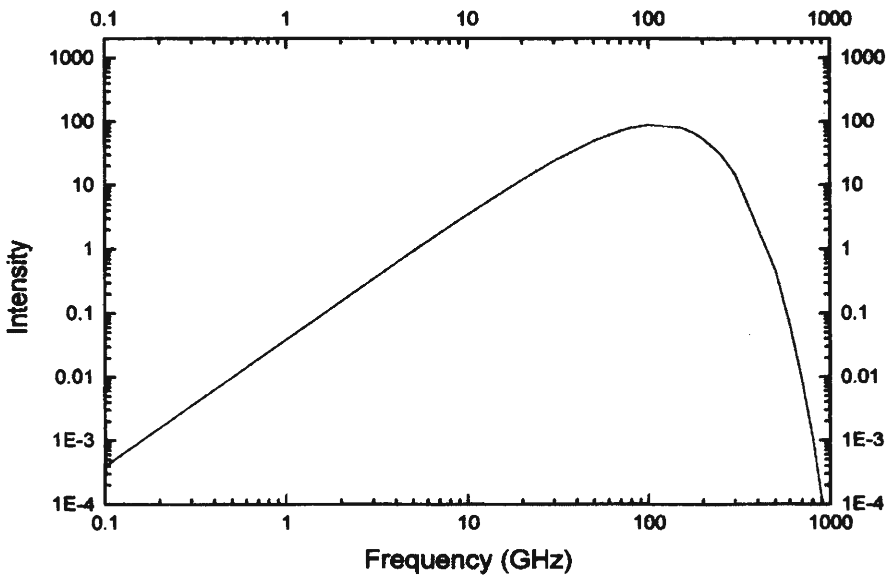

Figure. A typical blackbody spectrum represented by the Planck distribution formula corresponding to a temperature of 2.7 K.

Box 3. Wave Guides and Horn Antenna

We know that for domestic power lines, wires are used to transmit electrical power. This method works for electromagnetic waves at low frequencies, e.g., at 50-60 Hz. As we go to higher frequencies this method doesn't work well because the circuits would radiate energy in all directions. Upto frequencies of several megahertz (10^6 Hz) i.e., for radio waves coaxial lines are used, and in the gigahertz frequencies (10^9 Hz) or microwave region hollow conductors are used. These are called waveguides, which allow waves above a certain cut-off frequency to be transmitted. The geometrical shape of the waveguide determines this cut-off frequency. A good transmission line would let the power go through without any appreciable loss on the way. Power levels are usually expressed in their ratios using the logarithmic unit decibel (abbreviated dB). If the powers being compared are P_1 and P_2, then in decibels,$dB = 10\log_{10}(P_2/P_1)$. The sign associated with the number of decibels indicates which power is greater. A negative sign would mean that P_2 is less than P_1. The logarithmic ratio permits representation of an enormous range of power levels conveniently without using very large numbers.

To transmit power in the radio frequencies and higher, the transmission lines are connected to an antenna, which then radiates the energy into desired directions in space. The same antenna could be used also to receive signals. Antennae come in many shapes and sizes depending on the frequency of the wave and also the application. For radio and television, we are familiar with the loop and array antenna. For higher frequencies, dish and horn antennae are used. If a waveguide has an open end, power will be radiated from it. Since the dimensions of the aperture or the cross section of the waveguide are usually comparable to the wavelength, the radiation pattern is quite broad and non-directional. The aperture may be increased by flaring out the waveguide to a horn shape

generates a noise as a consequence of thermal motions (principally of conduction electrons). The noise power per unit bandwidth, i.e., per unit frequency range, is directly proportional to the temperature. The coefficient of proportionality depends in a known way on the electrical properties of the element. This relationship can be used backwards to assign an effective temperature to any noisy element, regardless of whether or not the noise is of thermal origin.

The whole system consisted of three components, viz., antenna, radiometer and reference source at a fixed temperature. Great care was taken to match the impedance[1] of these three components and to minimise the losses due to leakage in the joints.

[1] This technical term implies that signals are not reflected at the junctions of these elements.

The antenna consisted of a large expanding waveguide or horn, with an off-axis section parabolic reflector at the end as shown in

Figure 1. The focus of the paraboloid is located at the apex of the horn, so that a plane wave travelling along the axis of the paraboloid is focused into the receiver at the apex of the horn. The location of the receiver at the horn apex eliminates the loss and noise contribution of a connecting line. Its design emphasises the rejection of radiation from the ground. It is easy to see from the figure that in the configuration the receiver is well shielded from the ground by the horn. This feature ensures that when the antenna is pointed at the sky, very little noise power is received from ground or other sides. Since the design is based on geometrical optics and has no frequency sensitive component, it is extremely broadband. It was also not polarisation sensitive and thus had same efficiency for any linearly or circularly polarised radiation.

A radiometer is a device for measuring the intensity of radiation. A microwave radiometer consists of a filter to select a desired band of frequencies followed by a detector, which produces an output voltage proportional to its input power. Since practical detectors are not sensitive enough for the low power levels received by the radio telescope, amplification is normally used ahead of the detector to increase the signal level. Fluctuations in the power level of the noise in the first stage of the amplifier and the transmission line limit the sensitivity of a radiometer. Hence, a very low-noise amplifier was needed for the first stage. Shortly after the discovery of the maser amplifier by Townes and coworkers at Bell Laboratories (see *Resonance* articles on Masers and Lasers, Vol.4, No.9, p.6; No.10, p.8, 1999), H E D Scovil and his associates built the then world's lowest noise amplifier using a ruby maser. These amplifiers were cooled to 4.2 K by liquid helium and thus contributed only small noise to the system. This maser amplifier was incorporated in the radiometer tremendously increasing its sensitivity. Since astronomical radio sources produce random thermal noise very much like that from a hot resistor, the power received by the radiometers is quoted in terms of the equivalent temperature of a resistor which would deliver the same power to the radiometer. The equivalent noise

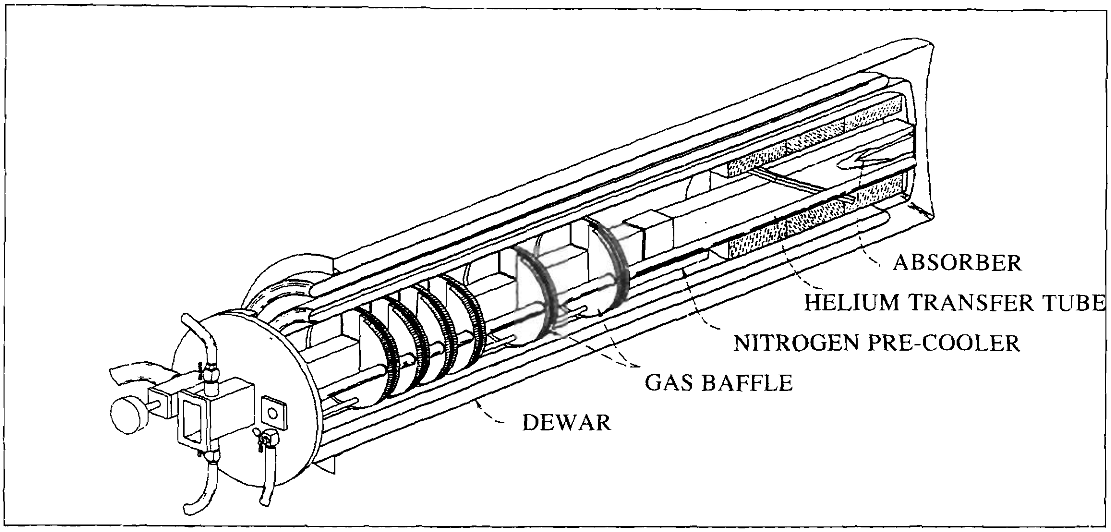

Figure 2. A schematic view of the reference source of Penzias. (Reproduced from Review of Scientific Instruments, Vol.36, p.68,1965 with permission from Lucent Laboratories and American Institute of Physics)

temperature is proportional to the received power (except in case of very short waves). The maser itself had a noise temperature of about 3.5 K.

The reference noise source consisted of a 1.22 m piece of 90 percent copper-brass wave guide connecting a carefully matched microwave absorber in liquid helium to a room temperature flange at the top as shown schematically in *Figure* 2. The bottom section of the waveguide is filled with liquid helium and a mylar window at a 30° degree angle keeps the liquid out of the rest of the waveguide. The design of the cryostat with carefully designed radiation shields and gas cooled baffles, is such that a small amount of liquid helium (about 20 litres) could provide 20 hours of operation. The temperature of the waveguide was monitored with a series of diode thermometers and the contribution of each section of the waveguide to the equivalent temperature of the reference source was calculated. When cooled, the calculated total temperature of the reference source was 5 K. Robert Wilson built a waveguide switch, schematically shown in *Figure* 3, with which either the antenna or the reference load could be connected to the maser amplifier by a simple mechanical rotation. This allowed accurate comparison of the equivalent temperature of the antenna to that of the reference load.

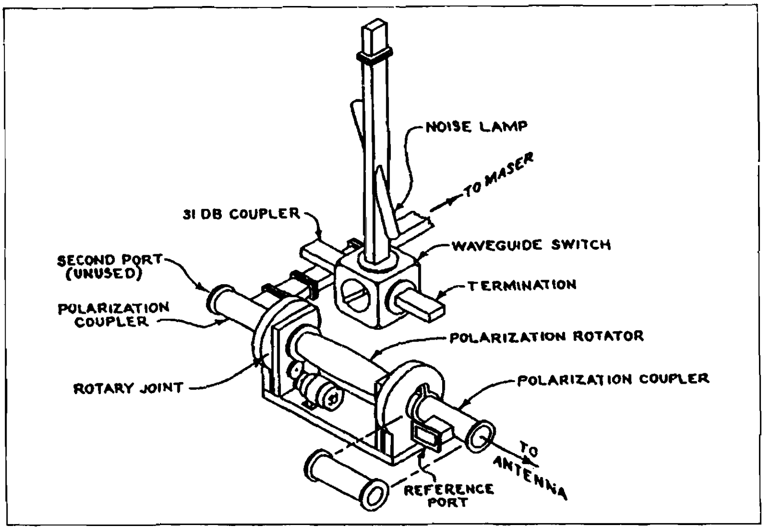

Figure 3. A schematic view of the radiometer and the switch used by Penzias and Wilson. (Reproduced from Astrophysical Journal, Vol. 142, p.1149,1965 with permission from Lucent Laboratories and American Institute of Physics)

The horn antenna was operated at a frequency of 4.08 GHz. Signals leaving the maser amplifier needed to be further amplified before detection. The remainder of the radiometer consisted of a down converter to 70 MHz followed by intermediate frequency (IF). amplifiers, a precision variable attenuator and a diode detector. The output of the diode detector was amplified and went to a chart recorder. The antenna, radiometer and the reference terminator[2] were matched so well that the reflection losses were less than 55 dB [3]. Thus the errors in the measurement of the effective temperature due to impedance mismatch could be neglected. The estimated error in the measured value of the total antenna temperature was 0.3 K and came largely from uncertainty in the absolute calibration of the reference termination.

[2] This term is used to mean the circuit element placed at the end of a waveguide or transmission line.

[3] This means reflected power is smaller than incident by a factor of $10^{5.5}$.

The measurements gave a total antenna temperature for the antenna pointed at the zenith, of 6.7 K. Three contributions to be subtracted from this value were that due to (a) atmospheric absorption, (b) ohmic losses in the antenna, and (c) back-lobe response. The contribution due to atmospheric absorption was obtained by recording the variation in the antenna temperature with elevation angle and employing the secant law. It is simply that the effect of the atmosphere is proportional to the path length through the atmosphere the radiation has to travel. This varies as the secant of the elevation angle and yielded a result of (2.3 ± 0.3) K. This result was in good agreement with published values for atmospheric contribution. The contribution from ohmic losses was calculated to be (0.8 ± 0.4) K from standard waveguide theory. The possibility of losses in the antenna horn due to imperfections in its seams was eliminated by attaching aluminium tape along the joints and observing no change in the antenna temperature. A thorough cleaning[4] of the antenna did not significantly alter the value of the recorded temperature. The back-lobe response to ground radiation was checked with a small transmitter on the ground and was estimated to be (0.1 ± 0.1) K.

[4] In one of their reports, the authors referred to 'white dielectric material' which had to be removed. This was their humorous allusion to pigeon droppings!!

From the combination of the above-observed values, Penzias and Wilson found that an antenna temperature of (3.5 ± 1.0) K remained unaccounted for. They continued their radio intensity measurements for over a year, living with this excess antenna temperature. That the value remained constant over this time helped ruling out any unknown source in the solar system. The radiation was also found to be isotropic and unpolarised.

Penzias and Wilson in the meantime became aware of the calculations of P J E Peebles in R H Dicke's group in Princeton University on radiation in the universe. They had considered an extremely hot condensed phase of the universe in their calculations and had the idea that if the radiation from this hot phase were large enough, it would be observable. The calculation of Peebles showed that the universe should be filled with a relic blackbody radiation of a temperature of about 10 K. The measured results of Penzias and Wilson were quite close to the expected value for the relic radiation. The results were published as two consecutive letters in *Astrophysical Journal* in Vol.142, 1965. None of these papers, however, referred to the work of Alpher, Herman and Gamow.

For nearly a decade it was uncertain whether the radiation was truly a blackbody remnant. Several other measurements by other groups in this period at wavelengths between 0.33 and 73.5 cm were consistent with a blackbody spectrum at a temperature of about 2.8 K. However, the characteristic feature of a blackbody spectrum is that it passes through a maximum, which for a temperature of 2.7 K would occur at a wavelength of about 0.11 cm given by Wien's displacement law. The problem was that measurements at shorter wavelengths could not be made from the ground because of atmospheric absorption. Finally in April 1975, a measurement of the entire range from 0.33 to 0.025 cm made with a balloon borne detector, by David Woody and collaborators of the University of California and Lawrence Berkeley laboratory was published. Their result published in the *Physical Review Letters*, showed clearly the presence of the maximum in the spectrum, in spite of the sizeable uncertainty in the measurement. Penzias and Wilson were awarded the Nobel Prize in 1978 for their discovery. Several measurements in the past three decades including the recent ones made using the Cosmic Background Explorer (COBE) satellite have not only confirmed the cosmic microwave background radiation, but also made precise determination of the characteristic temperature, with a value of (2.73 ± 0.01) K. The spectrum obtained is extremely well represented by a blackbody distribution over more than three decades in frequency.

Penzias and Wilson found the radiation to be isotropic and unpolarised to the level of 10%. Current observations show that the radiation is unpolarised at the 10^{-5} level but has a dipole anisotropy at the 10^{-3} level, with smaller scale anisotropies at the 10^{-5} level. The measurement of these small effects is extremely important in checking the various models for the development of the universe, especially in the very early stages of evolution in the 'big-bang' scenario. There already exist two approved satellite missions, the NASA Microwave Anisotropy probe scheduled for launch in 2000 and the ESA Planck Surveyor in about 2004. These missions are expected to provide more precise measurements of these anisotropies and checks on the cosmological models of the Universe. These later developments are very important and exciting but fall outside the scope of this article. They would surely be dealt with in *Resonance* in the future.

Suggested Reading

[1] A A Penzias and R T Wilson, *Astrophysical Journal*, Vol. 142, p. 419, 1965.

[2] R H Dicke, P J E.Peebles, P G Roll and D T Wilkinson, *Astrophysical Journal*, Vol. 000, p.414, 1965.

[3] R A Alpher and R Herman, *Physics Today*, Part 1, Vol.41, No.8, p.24, August 1988.

[4] D P Woody, J C Mather, N S Nishioka and P L Richards, *Physical Review Letters*, Vol. 34, p.1036, 1975.

Address for Correspondence
Amit Roy
Nuclear Science Centre
Aruna Asaf Ali Marg
P.O. Box 10502
New Delhi 110 067, India.
Email: roy@nsc.ernet.in