THE COSMIC MICROWAVE BACKGROUND RADIATION

B. Winstein*
Center for Cosmological Physics
The University of Chicago
Chicago, Illinois, 60637

ABSTRACT

These lectures will attempt to convey the excitement and promise in studies of the microwave radiation left over from the early universe. They are aimed at an audience of experimental high energy physicists.

The first lecture will concentrate on how the radiation is characterized, how it and its anisotropies are generated, and the physics that it reveals.

The second lecture will concentrate on the techniques that are used in experiments, addressing the question of how such small signals can be detected so precisely in a regime of extremely low signal-to-noise.

The third and final lecture will present the main thrusts in this field as foreseen for the coming decade.

*Supported by the NSF

1 Introduction

By studying the Cosmic Microwave Background Radiation field, cosmologists are uncovering extraordinary information about the early universe. Cosmology is in a very exciting and data-driven stage right now and arguably studies of the CMB have been the most fruitful.

These lectures will attempt to convey the excitement and promise to high energy physics experimentalists. The emphasis will be on experiment: how measurements are made and by what techniques. We will also need to understand some of the basic concepts to fully appreciate the science. The presentation will be more pedagogical than inclusive, as appropriate for the classroom. In particular, I apologize in advance to my new colleagues for an incomplete reference list, hoping that they will remember when they were first learning the subject.

Several people have been of great help in the preparation of these lectures; their names are given in the acknowledgments.

2 Lecture 1

We begin with a discussion of the very early stages of this field. We discuss the origins of the radiation, its black-body nature, how its anisotropies are generated and characterized, and how the hypothesis of Inflation accounts for the observed regularities. We close this section with a summary of the physics addressed by the CMB, using the recent WMAP results as the best example.

2.1 The Early Years

As is well known, the radiation was discovered in 1965 by A.A. Penzias and R.W. Wilson. It was a serendipitous discovery: the scientists were not explicitly looking for extragalactic radiation, let alone radiation of such import, but rather were led to the unavoidable conclusion that there was an isotropic source of "noise" in their detectors coming from the cosmos.

Figure 1 shows a rough plot of the number of articles published per year since the initial discovery. Today there are approximately 150 experimentalists working in this area. WMAP is having a great influence on the field: most of the publications in 2003 treat WMAP data in one form or another. It is interesting

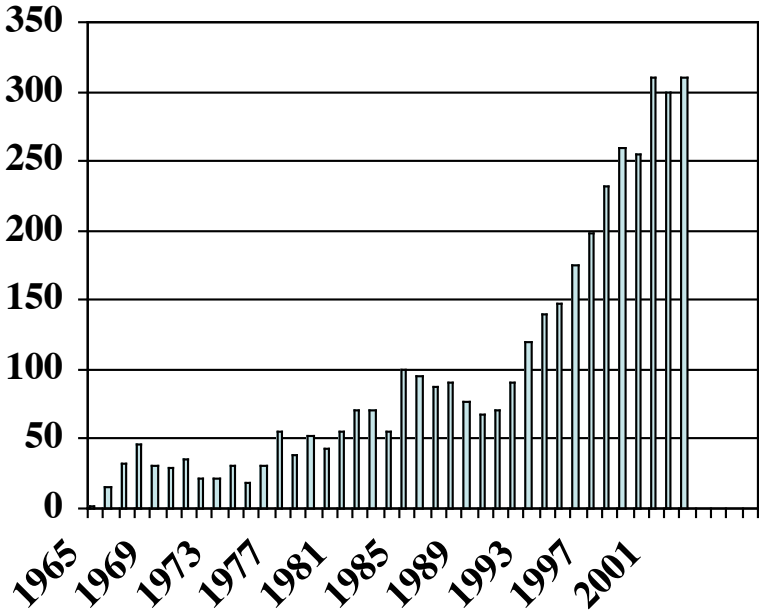

Figure 1: **Approximate Number of Papers on the CMB per Year**

to note that this publication profile is similar to that for papers published on CP Violation.

The first measurement was made at a wavelength of approximately 10 cm. The reported temperature was:

$$T = (3.5 \pm 1)K. \tag{1}$$

This measurement was way off the peak of the radiation which is more like 3 mm. This points out an important feature of the spectrum of black body radiation: if you know the flux at *any* known frequency, that determines the temperature of the radiation field. Figure 2 gives the flux of detected radiation (in watts per square meter per Hz per steradian) vs. wavelength for a variety of temperatures, nicely displaying the linear increase in detected radiation vs. temperature at low frequencies (the Raleigh-Jeans part of the spectrum) and the displacement of the peak to higher frequencies as the temperature increases.

Initially it was thought that the radiation might be reemission of star light by interstellar dust; this was ruled out by studies of the spectrum. Another possibility was that it was emission from radio galaxies but this hypothesis was discarded because it was known that there were not enough of these galaxies to reproduce the smoothness of the radiation.

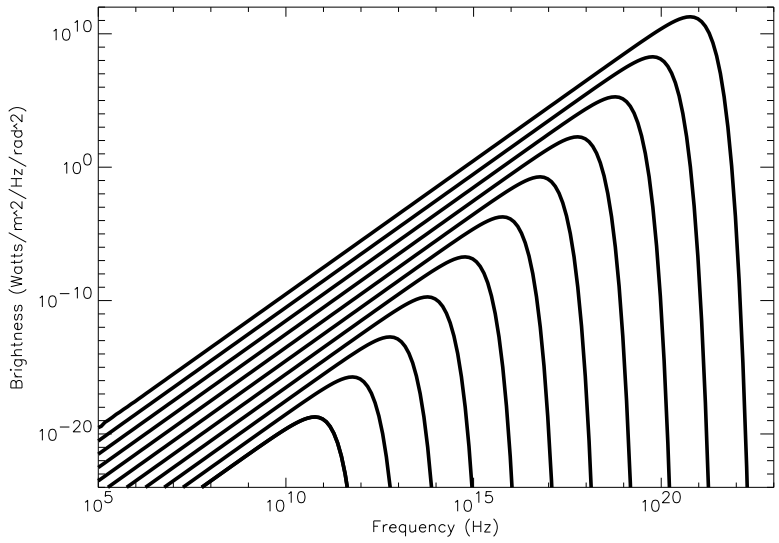

Figure 2: **The Black Body Spectrum.** Shown is the flux distribution for 1K, 10K, 100K, ... to $10^{10}K$. The Measured Flux at any temperature uniquely determines the temperature. (Plot courtesy of M. Hedman)

Key in the interpretation was whether the radiation indeed was isotropic and whether it really followed a blackbody spectrum. By 1967, Partridge and Wilkinson had shown over large regions of the sky, that:

$$\frac{\Delta T}{T} = (1-3) \times 10^{-3}. \tag{2}$$

In 1967 a Princeton group made the first measurement at a different frequency; this was crucial in verifying the black body nature of the radiation. Figure 3 shows the two measurements on top of the expected curve. Also shown are measurements of the galactic background at several different wavelengths, showing that it has a very different spectral index from that for a black body.

Another important "confirmation" came from a study of the pattern of absorption lines in interstellar CN molecules. It turns out that there are a pair of low-lying rotational states of this molecule. The first is excited with 2.64 mm radiation and the second can be reached from the first state with 1.32 mm radiation. From the observed relative population of these states, it was determined that the molecules were bathed in a sea of radiation with a temperature of

$$T = (3.05 \pm 0.35)K. \tag{3}$$

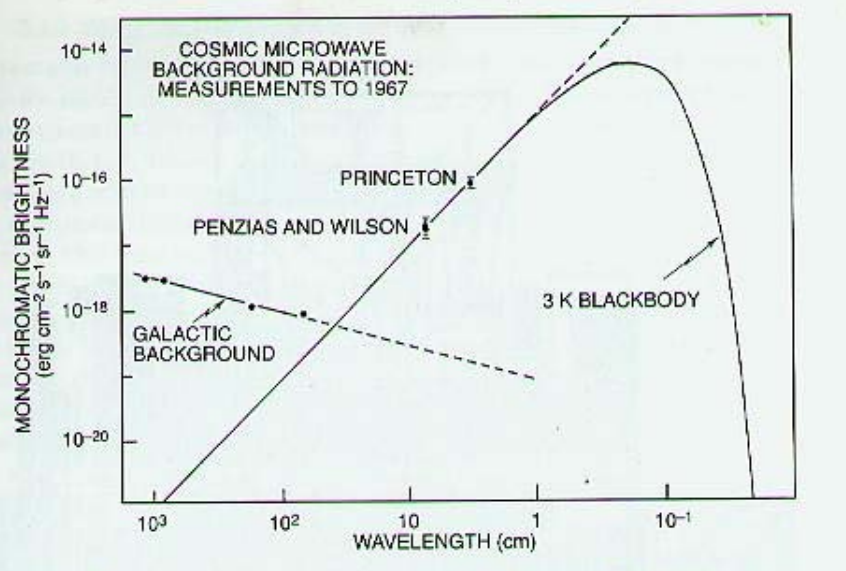

Figure 3: **The First Confirmation at 3 cm.** The flux from the original measurement together with the first confirmation at 3 cm are shown. Also shown are measurements of the galactic foreground (which was then called a "background"), showing its different (non-black body) spectral shape. (Figure from ref.[1])

This was a most remarkable result for at least two reasons: is showed for the first time that this new radiation extended beyond the solar system; and this discovery was made in 1941!

Today's spectral information on the CMB is shown in figure 4. The short wave-length measurements are dominated by the FIRAS instrument on the COBE satellite- the errors on these points are FAR smaller than the points themselves. Many other measurements at longer wavelengths contribute to making this the most accurately measured blackbody.

2.2 The Origin of the Radiation

The radiation we see today came from an era when the Universe was hot. When it had a temperature $T > 13$ev, protons, electrons and photons were the main constituents of a plasma in equilibrium; the expansion of the universe cooled the plasma to where hydrogen atoms formed (called the era of recombination, even though the atomic constituents had never been combined prior).

The parameters of this era are the following:

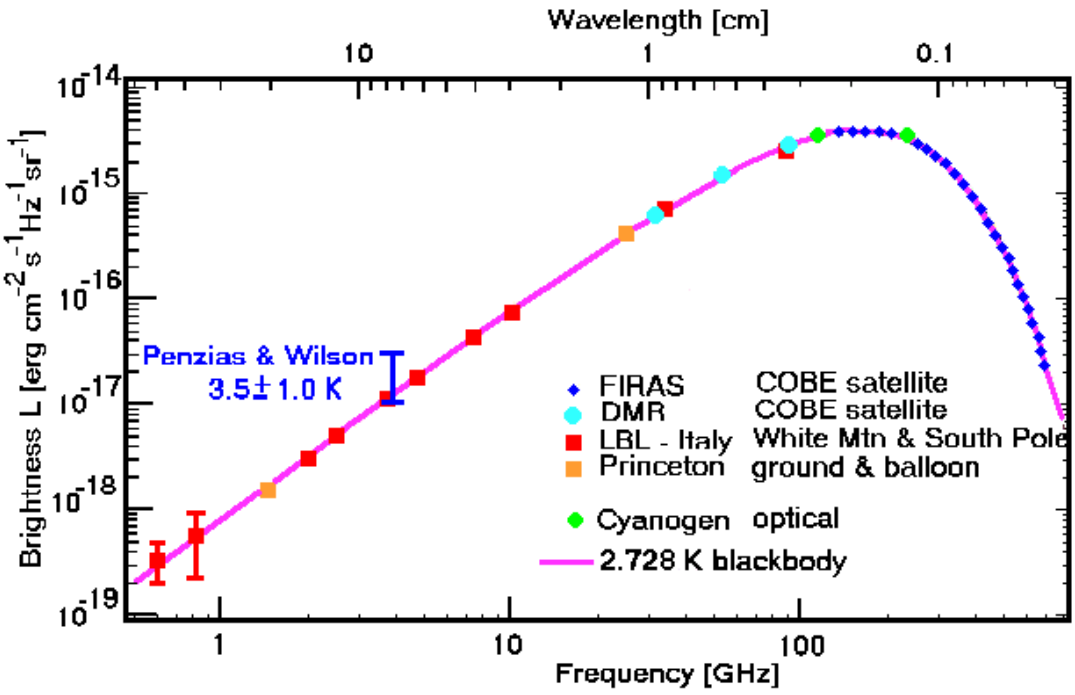

Figure 4: **The Planck Spectrum of the CMB**.

- $T \approx 300K \approx 0.3$ ev
- $z = 1088 \pm 2$
- $\delta z = 195 \pm 2$
- $t = (379 \pm 8)KY$
- $T_{today} = 2.728K$

The photons that we see today were then emitted when the universe was some 1100 times smaller than today, from a "surface" with a thickness of about 20%, about 380 thousand years after the Big-Bang. The (impressively small) errors on the values are taken from the recent WMAP results.

Except for second order effects (to be discussed in the 3rd lecture), photons come to us from this surface of last scattering with no further interactions, save the red-shifting due to the expansion. As such they provide an important and unique view of the universe at a well defined time, indeed the earliest such time. It isn't immediately obvious how a view of the Universe, even a very precise view, when it had a temperature as less than 1 ev can yield unique information about its origins, or at least its properties when the particle energies were far greater.

Other "relics" come from substantially earlier eras. Neutrinos, because of their much lower cross-section, decoupled minutes after the Big-Bang but the prospects for their detection is slim. And the protons in our bones were first created at an even much earlier era but have had an extremely complicated history so that only their abundance hints at anything "primordial."

2.3 Why (and how good) a Black Body?

An important question is why the radiation we see today is so precisely thermal and how well we can limit non-thermal contributions.

Before (re)combination, Thomson scattering dominates the interactions between the electrons and photons. This (elastic) process will change photon directions but not the number of photons. The latter needs to change when there is a non-thermal distribution of injected energy but Thomson scattering can not be responsible for the restoration of equilibrium when, for example, e^+e^- annihilation occurred at $z \approx 10^9$.

The relevant processes that do change photon number are Bremstrahhlung and Compton scattering; these processes would take an essentially arbitrary spectrum at $z \approx 10^6$ and thermalize it. When there is not enough "time" for photon numbers to shift, the spectrum looks like a black-body but with a linear shift to higher frequencies. The fractional shift is parameterized by the variable y.

The photon number distribution for a pure blackbody (the Planck distribution) is governed by the factor:

$$\frac{1}{e^{\frac{h\nu}{kT}}-1} \tag{4}$$

whereas the number density for particles in "kinetic equilibrium", where there are no processes changing the number of photons (the Bose-Einstein distribution), is governed by the factor:

$$\frac{1}{e^{\frac{h\nu}{kT}+\mu}-1} \tag{5}$$

where μ is the chemical potential.

The spectrum shown in figure 4 fits a blackbody with residuals less than 50 ppm. The corresponding limits on the spectral distortions are:

$$y < 15 \times 10^{-6} \tag{6}$$

and

$$\mid \mu \mid < 9 \times 10^{-5}. \tag{7}$$

For an energy release ΔU with $10^5 < z < 3 \times 10^6$, a Planck Spectrum becomes Bose-Einstein and from the limit on the size of the chemical potential the limit[2] for energy injection during this era is:

$$\frac{\Delta U}{U} = 0.7\mu. \tag{8}$$

while an energy release with $z < 10^5$ causes a spectral shift and:

$$\frac{\Delta U}{U} = 4y. \tag{9}$$

Thus the purity of the black body spectrum lends great confidence that no unusual or unexpected energy injections, at the level of a few $\times 10^{-5}$ of the total energy in the photons occurred later than $z \approx 10^6$ or later than a few months after the Big-Bang.

2.4 Anisotropies

From the very beginning, cosmologists began searching for anisotropies in the background radiation: were the radiation pattern, and therefore the associated matter density, precisely uniform, then it would have been difficult to explain the origin of the obvious anisotropy in the matter distribution today.

Technically the first anisotropy was seen in the late 1970's, that arising from the motion of the sun through the microwave radiation field, creating a dipole of magnitude 10^{-3}. But the true (primordial) anisotropy was detected first by the COBE satellite in 1992. The detector had a "beam" of 7 degrees so the anisotropies could only be detected on scales larger than this. Regions this far separated were never in causal contact with each other, at least in the standard big-bang model so the detected anisotropies, of the order of 10^{-5} or $30\mu K$, need to be considered as "initial" conditions.

Figure 5 shows the path of the radiation since decoupling, indicating that the causal horizon is of order 2^0.

Jumping to the present, the recent map of the radiation field by the WMAP satellite is shown in Figure 6. This is a beautiful achievement, the precise whole-sky determination of the temperature with a spatial resolution of about 0.3^0.

2.4.1 Characterization of the Temperature Field

Cosmologists have developed techniques for characterizing the temperature field which we will now describe. First, let's denote the temperature T at the surface

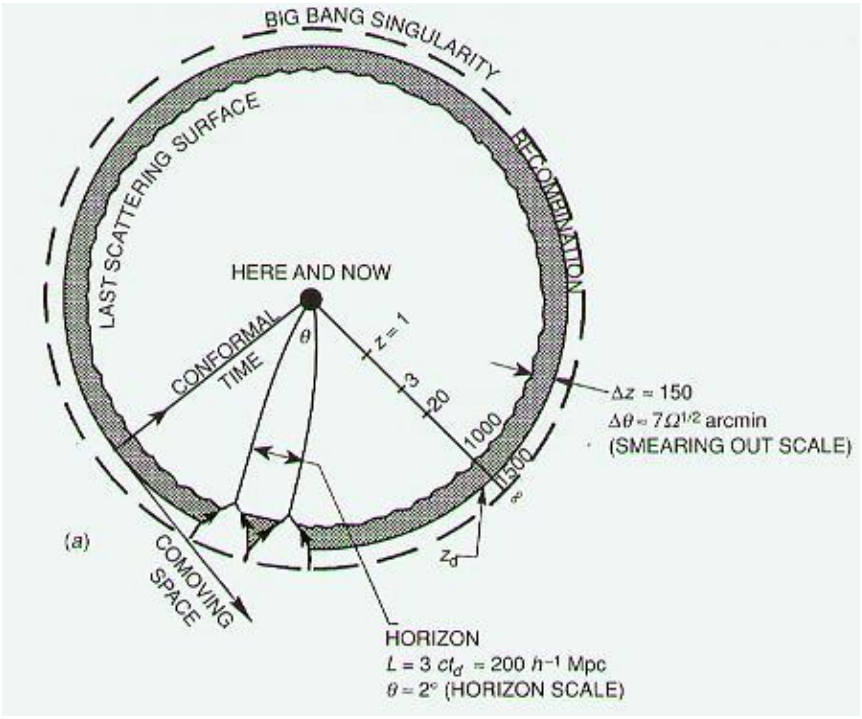

Figure 5: **Schematic of the path of the CMB.** In the Big-Bang (pre-inflationary) model, patches of the CMB sky observed today that are more than about 2^0 apart were never in causal contact.

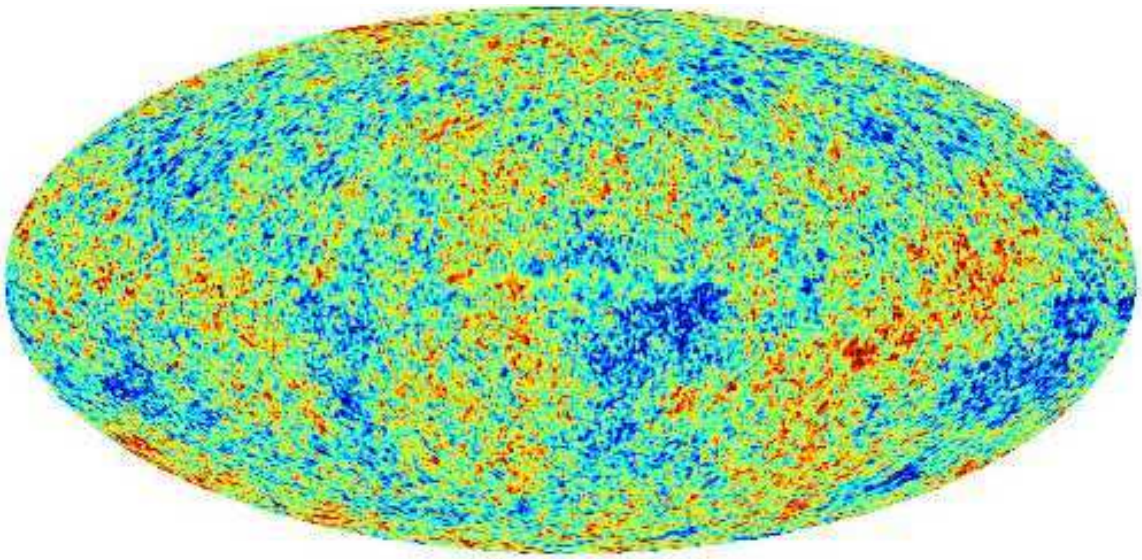

Figure 6: **The WMAP Map of the CMB Sky.** The temperature scale spans $\pm 200 \mu K$.

of last scattering in direction $\hat{n}$ by its deviation from the average:

$$\Theta(\hat{n}) = \frac{\Delta T}{T}. \tag{10}$$

Next, we can consider the multi-pole decomposition of this temperature field:

$$\Theta_{lm} = \int \Theta(\hat{n}) Y^*_{lm}(\hat{n}) d(\hat{n}). \tag{11}$$

Finally we can define the "power spectrum" of the radiation field, which represents the strength of the radiation field vs. multi-pole. If the sky temperature field arises from Gaussian random fluctuations (as the inflationary paradigm predicts), then:

$$\langle \Theta^*_{lm} \Theta_{lm} \rangle = \delta_{ll'} \delta_{mm'} C_l. \tag{12}$$

Under these conditions, the best estimate of C_l is the square of the width of the power in the $(2l+1)$ samples of this multi-pole of the temperature field.

Because there are only the $(2l+1)$ modes with which to detect the power at multi-pole l, there is a fundamental limit beyond which the power cannot be better determined; this is known as the cosmic variance:

$$\frac{\Delta C_l}{C_l} = \sqrt{\frac{2}{2l+1}} \tag{13}$$

For historical reasons, the quantity that is usually plotted is:

$$\Delta^2_T \equiv \frac{l(l+1)}{2\pi} C_l T^2_{cmb} \tag{14}$$

This is the variance (or power) per logarithmic interval in l; this quantity is expected to be (nearly) uniform in inflationary models ("scale invariant" fluctuations) and would go like l^2 if the "noise" on the sky were pure "white".

The power spectrum derived from the map in figure 6 is then shown in figure 7. WMAP data extends up to an l of about 700; two other recent results with smaller beams and therefore higher l reach are also shown in the figure.

In the figure the region below $l \approx 50$ is that of "initial conditions" which we have already discussed; at higher l values there are "acoustic oscillations". We will now discuss this mechanism which is responsible for the large enhancements over the basic level of fluctuations in the figure.

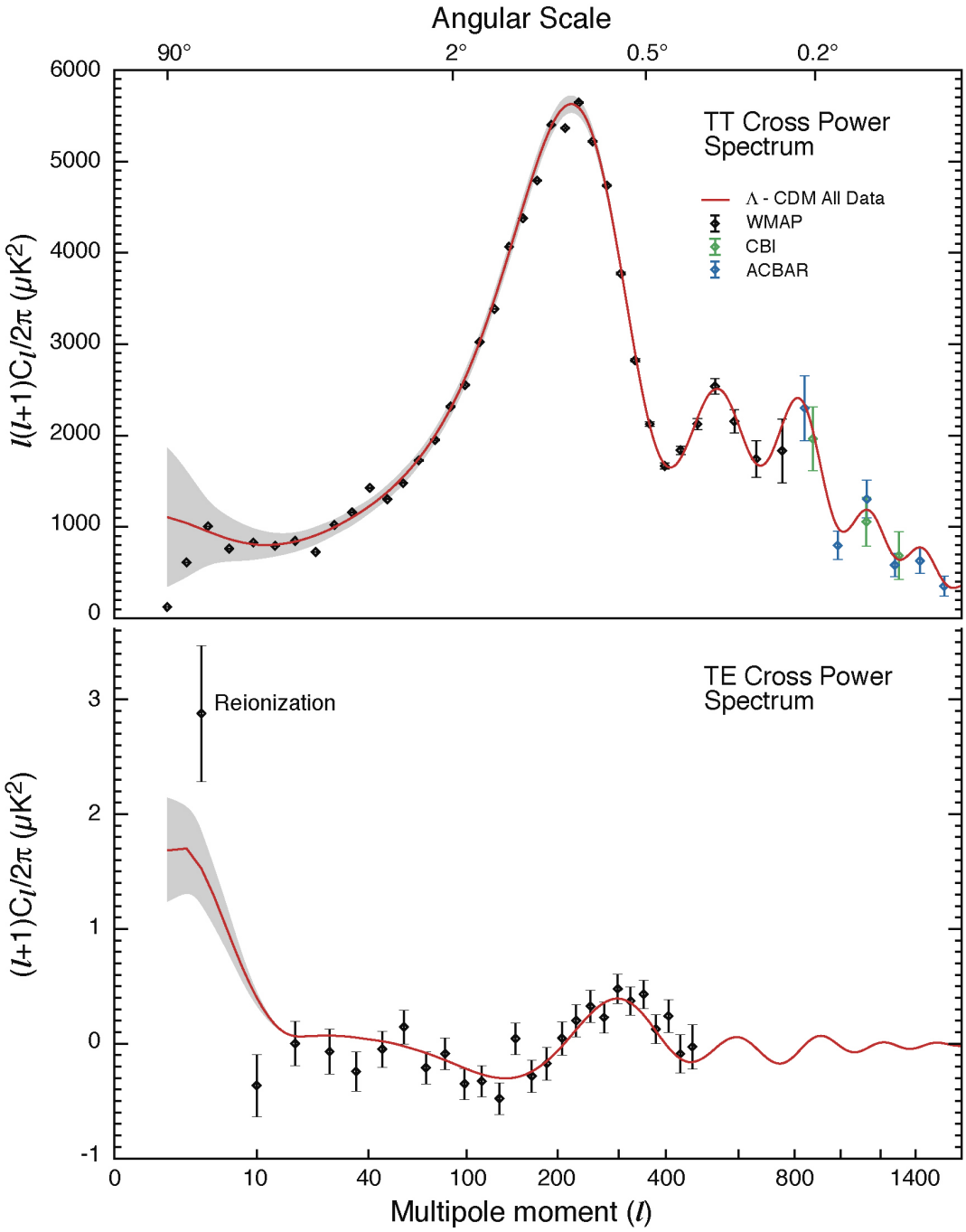

Figure 7: **The Multi-pole Power Spectrum Derived from WMAP Data**. The upper plot shows the power spectrum of the temperature field while the lower one shows that for the correlated temperature-polarization fields. High-l points from the CBI and ACBAR experiments are also shown together with the best model fit to the WMAP data. (Polarization will be treated in lecture 3.)

2.5 Acoustic Oscillations

The standard picture of the early universe is that the "initial conditions" were a very smooth and flat universe with very small initial inhomogeneities. These are small over-densities in dark matter, protons, electrons, neutrinos, and photons where each component has exactly the same distribution. Such conditions are called "adiabatic."

The regions with more density can attract further mass and grow but only when enough time has passed that the surrounding region becomes aware of the over-density, i.e., in the language of the cosmologists, when the surrounding region comes into or crosses the horizon.

As already mentioned, the horizon at decoupling corresponds today to about 2^0 on the sky. Regions smaller than this would have had time to compress while those larger would not. For sufficiently small regions, the photon pressure that builds up with compression would have reversed the compression and the region would then expand. This sets up an oscillation process: the relativistic fluid of photons is coupled to the electrons via Thomson scattering, and the protons follow the electrons to keep a charge balance. At decoupling, one is presented with a snap-shot of the state of the fluid at that time. (A excellent pedagogical description of the oscillations can be found at.[3])

At the time of decoupling, perturbations of particular sizes may have undergone: one compression; one compression and one rarefaction; one compression, one rarefaction, and one compression; etc. These particular regions will exhibit extrema in the power-spectrum.

To put this in a slightly formal context, consider a standing wave on the sky with a frequency ω and wave number k. The wave displacement w can then be written:

$$w(x,t) \propto \sin(kx)\cos(\omega t) \tag{15}$$

The displacement will be <u>maximal</u> when $t = 0, \frac{\lambda}{2v_{sound}}, \frac{\lambda}{v_{sound}}, \ldots$. There will then be a series of frozen peaks (and valleys) at t=t_{rec} which will depend on the velocity of sound in the medium.

To help explain these ideas, I've borrowed a few frames from an animation created by Wayne Hu. Figure 8 shows a density fluctuation on the sky from a single mode in k-space and how it is "detected" with time. The wave-length of this mode can be taken to be roughly the sound horizon at decoupling.

The figure shows the "particle horizon" just after decoupling. This represents the farthest distance one could in principle "see" or detect anything; it is effectively the age of the universe times the speed of light. No one living at the center of the figure could by any means have knowledge of anything outside this region.

Of course, just after decoupling, one could "see" a far shorter distance: light can only then propagate freely so one would not have received any photons from any appreciable distance.

The subsequent frames show the pattern "detected" at later times: at first one sees a dipole, this when photons can travel the scale of the excited mode; later one would see a quadrupole, then an octupole, etc., until the present time when that single mode in density inhomogeneities creates very high multi-poles in the temperature anisotropy.

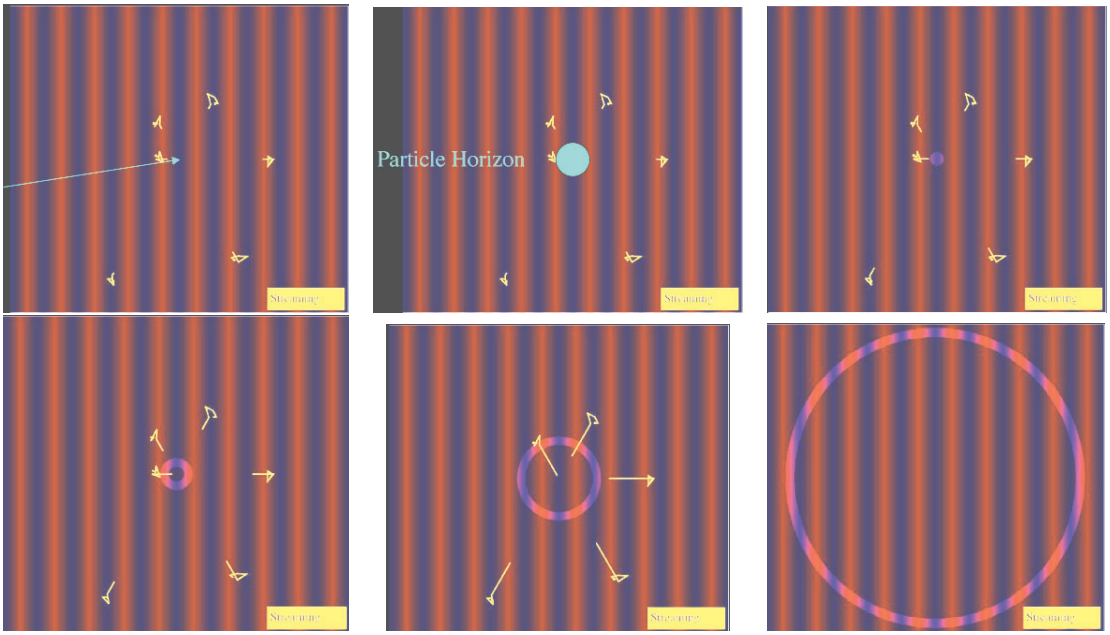

Figure 8: **The Signature of One "Frozen" Mode after Decoupling.** These illustrations (from W. Hu) show, left to right, top to bottom: one mode in temperature just after decoupling with the observer where the arrow is pointing; the observer's particle horizon at this time, when only the monopole can be detected; some time later the quadrupole is first seen; even later the quadrupole is well detected; much later the octupole is detected; and "today" a very high, well aligned multi-pole, from just this single mode in k-space, is detected.

2.6 The Idea of Inflation

The boldness in the idea of inflation can be expressed as follows: at the time shown in figure 8 top-middle, no one could have any knowledge of fluctuations outside the particle horizon. However, a smart theorist at the time could predict that the observed dipole pattern would become progressively more complex as time evolves, knowing about the existence of peaks and troughs of this mode outside the particle horizon.

Our universe could have begun as in figure 8: only one highly oriented mode as its initial condition. This would have, at least, led to a very different WMAP result! But it seems that the initial conditions were rather all k-modes equally excited in all directions, where "equally" means drawn from the same Gaussian random distribution.

Inflation is a mechanism whereby these fluctuations are created without violating causality, behaving as effectively initial conditions. There seems not to be a better candidate mechanism to explain the observed regularities. Nevertheless, Wolfgang Pauli's famous statement about the neutrino comes to mind:

> I have done a terrible thing: I have postulated a particle that cannot be detected!"

Sometimes it seems that Inflation is an idea that cannot be tested, or incisively tested. Of course Pauli's neutrino hypothesis *did* become testable and it tested positive, and similarly there is hope that the idea of inflation can reach the same footing. As a note of caution, however, let's remember that we have not (yet) seen *any* scalar field in nature. In the third lecture, we'll discuss what has been claimed as the smoking-gun test of Inflation- the detection of gravity waves in the CMB- but even so, will we ever say with certainty that the Universe grew by 10^{63} in volume in just 10^{-35} seconds?

2.7 What we learn from the WMAP Power Spectrum

The WMAP team has used its data to report on the following features of the Universe:

- Its Geometry
- The spectral "tilt" to the power spectrum
- Its Matter Content

- Its Baryon Content

This information is determined from the structure in the peaks observed and here we will try to give a feeling for how this comes about. For this discussion we are guided by reference.[4]

2.7.1 The Geometry of the Universe

As we have discussed, the angular scale of the peaks Θ_A is set by the (co-moving) sound horizon at decoupling, r_S, divided by the relevant distance to the surface of last scattering, d_A:

$$\Theta_A = \frac{r_S(z_{dec})}{d_A(z_{dec})} \tag{16}$$

The quantity d_A, the "angular diameter distance" is that distance that properly takes into account the expansion history of the Universe so that when it is multiplied by an observed angle, the result is the feature size. In a non-expanding universe, this would simply be the physical distance. The expression depends on the content of the Universe. For a **flat** Universe, we have:

$$d_A = \int_0^{z_{dec}} \frac{H_0^{-1} dz}{\sqrt{\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_\Lambda}} \tag{17}$$

In this expression, H_0 is today's value of the Hubble constant and $\Omega_{r,m,\Lambda}$ is the fraction of critical density in radiation, matter, and cosmological constant, respectively. It is easy to see how the dilution of the different components with red-shift z enters the expression for the angular diameter distance.

WMAP measures

$$\Theta_A = (0.601 \pm 0.005)^0. \tag{18}$$

Completely within the WMAP data, there is a "geometrical degeneracy" between Ω_k, a contribution to the critical density from the curvature of space, and Ω_m. Taking a very weak prior of $h > 0.5$, they then determine that $\Omega_k = 0.03 \pm 0.03$, i.e. no evidence for curvature. They then assume $\Omega_k = 0$ for subsequent analysis.

2.7.2 Fitting for Spectral Tilt, Matter and Baryon content

It is easy, in principle, to see how one determines the spectral tilt: if one knew all the other parameters, then the tilt would simply be the slope of the power spectrum after the removal of the contributions from the other parameters. However, as should be clear, there is a coupling between all of the parameters.

Let us first look at the ratio of the 2nd to the 1st peak, denoted (by the WMAP team) by H_2^{TT}. The value reported is $H_2^{TT} = 0.426 \pm 0.015$ and its uncertainty, in terms of that on the parameters of interest, is given by:

$$\Delta H_2^{TT} = 0.88\Delta n_S - 0.67\frac{\Delta\omega_b}{\omega_b} + 0.39\frac{\Delta\omega_m}{\omega_m} \tag{19}$$

Here the ω parameters are the physical densities.

It is often said that the ratio of the first two peaks gives the baryon density. This is of course true if one assumes that there is no tilt and one takes ω_m from elsewhere. The baryons can be considered as extra loading to the oscillations; if their density increases, the second peak is less intense. But it is clear that the (cold dark) matter density also contributes as does the spectral tilt.

These latter two parameters are determined by looking at the ratio of the 3rd to the first peak ($H_3^{TT} = 0.42 \pm 0.08$), for which data outside of WMAP is used; and the ratio of the 1st to the second peak in the polarization-temperature cross correlation power spectrum ($H_3^{TE} = 0.33 \pm 0.10$). *

The uncertainties in these quantities, in terms of the uncertainties in the parameters, are:

$$\Delta H_3^{TT} = 1.28\Delta n_S - 0.39\frac{\Delta\omega_b}{\omega_b} + 0.46\frac{\Delta\omega_m}{\omega_m} \tag{20}$$

and

$$\Delta H_3^{TE} = -0.66\Delta n_S + 0.095\frac{\Delta\omega_b}{\omega_b} + 0.45\frac{\Delta\omega_m}{\omega_m} \tag{21}$$

The baryon density has little influence on the ratio of the TE peaks but otherwise it is clear that the three parameters are quite important in the three measured peak ratios. A simultaneous fit to the three results in the following values:

We have also included the overall amplitude of the power spectrum and the optical depth, determined from the discovery of the era of re-ionization which will be mentioned in Lecture 3.

2.8 Summary: Temperature Anisotropies

The WMAP data set has taught us much about the Universe. While the overall conclusions from the analysis of the peak structure are not dramatically different from those drawn from a collection of earlier ground- and balloon-based experiments, WMAP has put the reigning cosmological model on much stronger footing.

*We will discuss polarization measurements in lecture 3.

Table 1: Cosmological Parameters from WMAP

Baryon Density	$\Omega_b h^2$	0.024 ± 0.001
Matter Density	$\Omega_m h^2$	0.14 ± 0.02
Hubble Constant	h	0.72 ± 0.05
Amplitude	A	0.9 ± 0.1
Optical Depth	τ	$0.166^{+0.076}_{-0.071}$
Spectral Index	n_S	0.99 ± 0.04

Prior to WMAP, only two experiments claimed systematic errors on the spectral power less than 10%; WMAPs errors are less than 0.5%. The overall amplitude is strongly affected by the reionization, changing previous values by of order 30%. And with full-sky coverage, WMAP could determine the power spectrum in individual l-bins with negligible correlations.

The χ^2 for the grand fit is 1431 for 1342 degrees of freedom. The fluctuations by all tests appear to be derived from Gaussian distributions. Yet the probability to exceed this value is less than 5% so, as a result, some thoughts have arisen as to possible effects that, if present, would provide a better fit. Allowing the tilt parameter n_S to "run" does reduce the χ^2. Perhaps more significantly, WMAP confirms the COBE observation of unexpectedly low power in the lowest multipoles. The WMAP team claims this is an effect way beyond statistics but there is a lively literature on the subject. Figure 9 shows the lowest 3 multi-poles where not only is it clear that the quadrupole has little power, it appears to be "aligned" with the octupole! But the situation is murky in that the quadrupole lines up reasonably well with the galaxy itself and there is concern that the cut on the galaxy then reduced the inherent quadrupole power.

We will now leave our survey of present knowledge of the temperature anisotropies to discuss the techniques used to make these measurements.

3 Lecture 2

In this lecture, we will concentrate on the means for detection of the microwave radiation. This will include a discussion of the basic elements of radio telescopes, the different types of receivers, sources of noise, the factors governing the sensitiv-

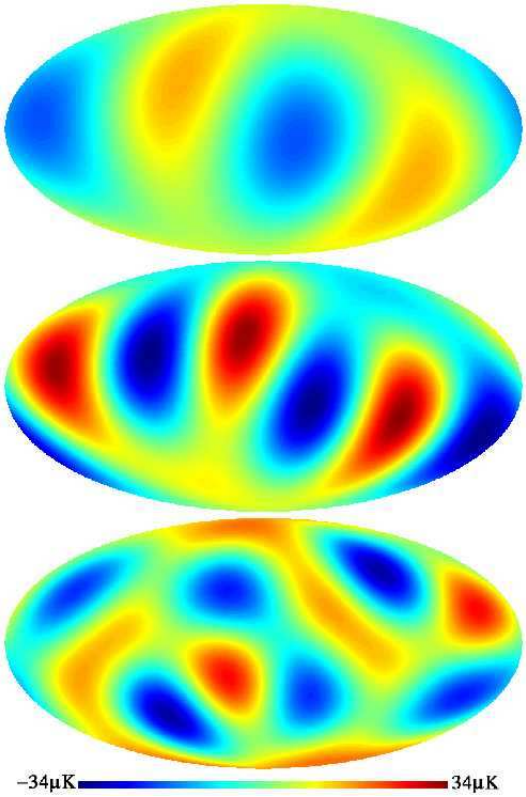

Figure 9: **WMAP, multi-poles l=2,3,4, from top to bottom.**

ities of receivers, observing strategies and trade-offs, and the analysis chain from raw data to power spectra.

3.1 Radio Telescopes

Consider first an experiment that observes one point on the sky: what is needed for this? Figure 10 shows R. H. Dicke and colleagues calibrating an early radio telescope.

What Dicke used (see Figure 11) was:

- An Antenna, with a particular acceptance pattern on the sky (its "beam")
- An amplifier, which must be low-noise and high bandwidth
- A filter, which selects the frequency band of interest
- a power meter, which squares the electric field, and
- an (essentially DC) amplifier

Figure 10: **R. H. Dicke in the mid 1940s.** Here Dicke is holding an absorber in front of his radio telescope.[1]

(In this section, we are describing what are called coherent detectors, those that amplify the electric field. These are the type, for example, used in WMAP. In the next lecture, we will describe bolometric detectors which immediately turn the radiation into heat.)

The depiction of an antenna "beam" on the sky is shown in figure 12. It is useful to invoke the reciprocity of Maxwell's equations: the acceptance pattern of an antenna on the sky is precisely that of its radiated power. Because the microwave radiation has wave-lengths λ in the mm-cm range, ray optics does not suffice. A typical pattern is a nearly Gaussian beam with a nominal beam solid angle Ω and a collecting area A. There are inevitably side lobes- the typical diffraction pattern- to the main beam, some near to the main beam and others far from it. The intensity in the far side-lobes must be carefully controlled as often it is aimed at the ground which is at $\approx 300K$ which could completely dwarf the anisotropies in the $3K$ "signal".

Because we are in the diffraction limit, the more collimated we wish to make the antenna pattern- to study fine features of the microwave radiation- the larger we must make its collecting area. The relevant relation is: $A\Omega = \lambda^2$.

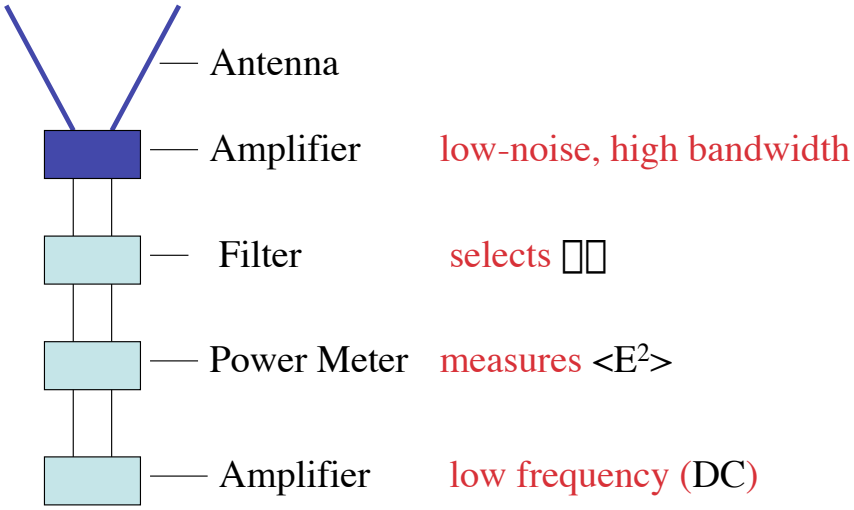

Figure 11: Simplified Radiometer Diagram

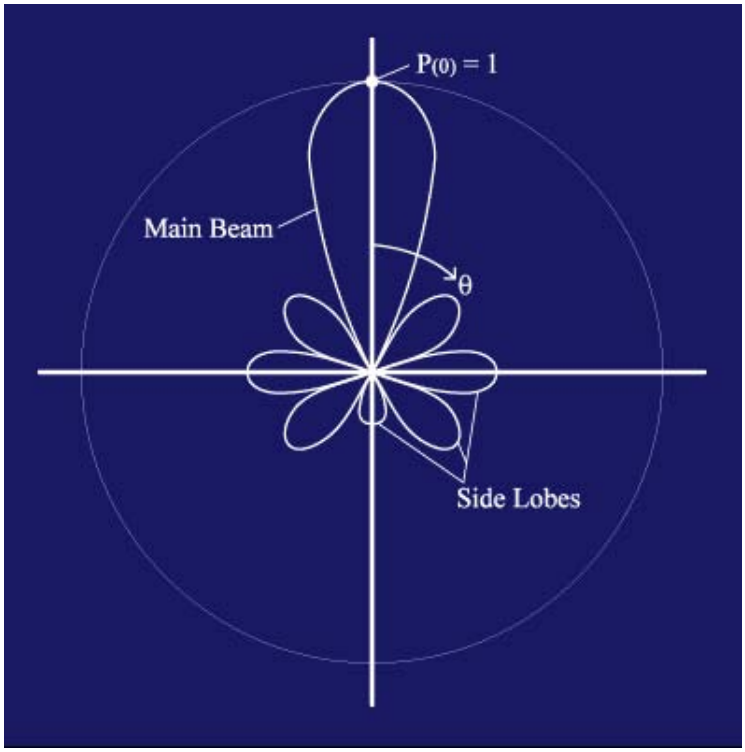

Figure 12: **Illustration of a Microwave Beam Pattern.**

3.1.1 What is the CMB flux?

The Planck black-body spectrum is given by:

$$B_\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{kT}}-1} W/m^2/str/Hz \tag{22}$$

Consider a simple horn antenna, not unlike that in figure 10. Using the following parameters:

- central frequency $\nu_0 = 90$ GHz
- bandwidth $\Delta\nu = 10$ GHz
- beam $FWHM = 8^0$
- beam area = 8 cm^2 ,

we find that the CMB flux on the horn is just 2.5×10^{-13} Watts.

To give a feeling for actual signal levels in an experiment, we note the following:

- typical gain of the front-end (low noise) amplifier is 10^6
- the transmission factor in the filter is of order 0.4
- the power meter typically gives, for 1μ Watt input, 1 milli-Volt output
- typical gain of the final amplifier is 100

so this leads to a roughly 10mV signal for the 3^0 incident radiation.

3.2 Heterodyne Receivers

With *coherent* receivers, one can "mix down" the radio frequency to an intermediate frequency (IF). For example, a band from 84 to 100 GHZ, when multiplied by an 82 GHz local oscillator (LO), is reduced to an IF band from 2-18 GHz. At these frequencies, signals can be transfered in coax (rather than wave guide) and important manipulations can be performed as we will see. In addition, amplifiers are less noisy and less expensive the lower the frequency.

The next few pictures are intended to give the reader an idea of what real radio receivers and telescopes look like. I use pictures and examples from the CAPMAP[5] experiment.

Figure 13 shows some of the microwave "'plumbing". Shown are MMICs, which provide multistage RF amplification. As in photo-multipliers, the first stage is the most critical with respect to added noise. Also shown are the RF filters and

"mixers" in which the RF and LO signals are mixed down to IF frequencies. Other components of this (polarization) receiver will be explained in the 3rd lecture.

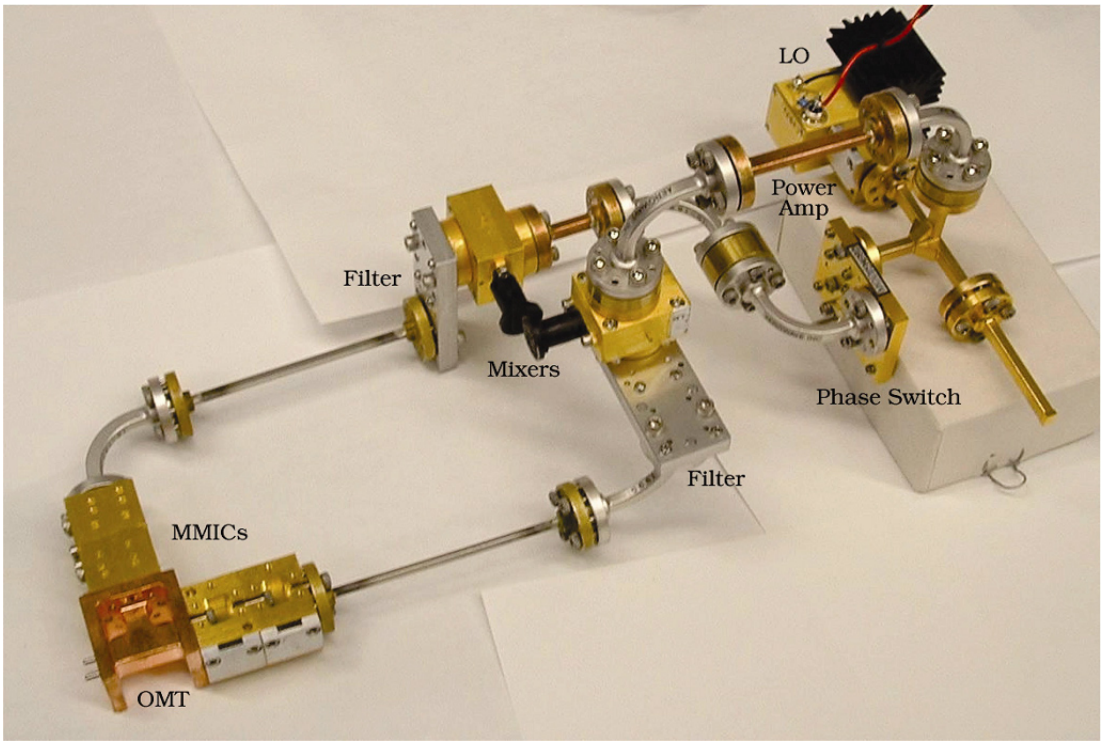

Figure 13: **A CAPMAP radiometer.** This radiometer detects polarization; polarization will be treated in Lecture 3

Figure 14 shows power detectors and the final stage of amplification.

Figure 15 shows four receivers in a cryogenic dewar. Inside the metal cans are the horns; on top of the cans are lenses to match the output of the horns to the radio telescope, shown in figure 16.

3.3 System Temperature

The sensitivity of a radio receiver depends on what is called its System Temperature. It turns out it is possible to characterize the noise in a radio receiver as an effective temperature due to the power coming from it even in the absence of a true signal. The sensitivity of the receiver is better if the system temperature is made as small as possible. Here we will describe the 3 major contributions to the noise (or system temperature) of ground-based radio receivers.

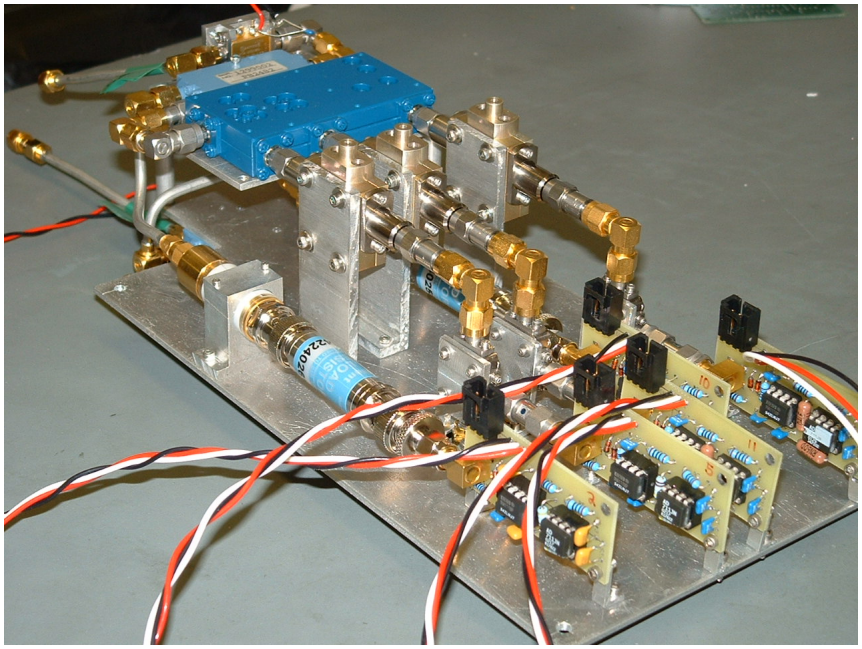

Figure 14: **A CAPMAP IF stage.** The module has two coaxial inputs, carrying 2-18 GHz, which are split and correlated to provide 5 outputs.

First we must mention that the language we are using was originally introduced by Bob Dicke in a seminal paper in 1946. Its first page is reproduced in figure 17. By considering an antenna immersed in a cavity whose walls are at temperature T, he related the power transmitted down the antenna with the Johnson (thermal) noise in a resistor to which the antenna is connected. Arguing from thermodynamic equilibrium, he noted that the power levels going each way had to be equal when the resistor was held at what he defined as the "antenna temperature". In this way, one can treat a source on the sky and the thermal noise of an amplifier with the same language. In this paper, we will see that he also derives the fundamental expression for the sensitivity of a radio receiver.

3.3.1 Atmospheric Noise

On the ground, one is looking through an (hopefully stable) atmosphere. The atmosphere will both absorb incident radiation and emit its own radiation.

These are connected by Kirchoff's Law. Let's consider a detector looking at a source with a temperature T_S through a cloud with temperature T_C. (See figure

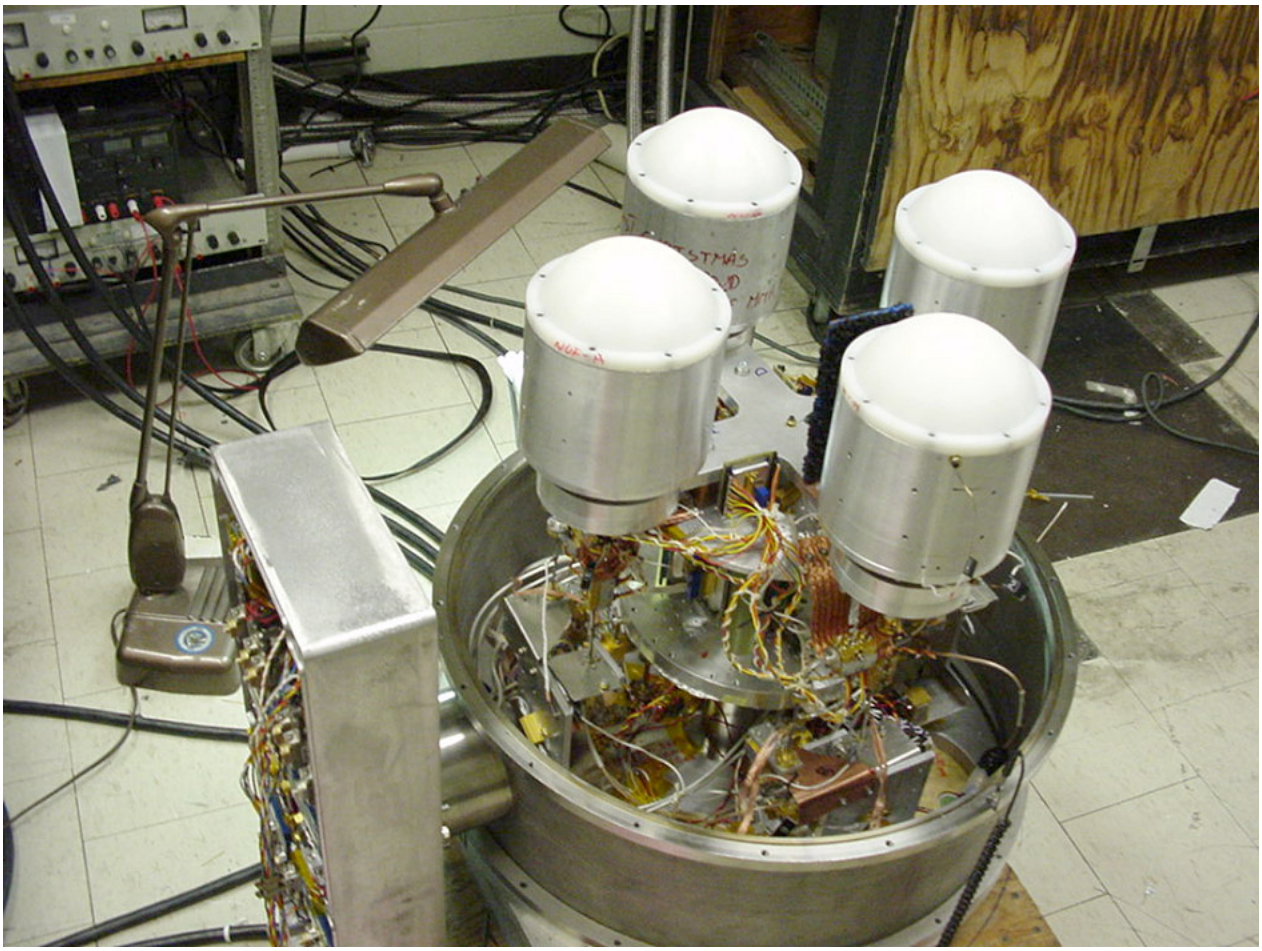

Figure 15: **A CAPMAP Dewar.** Four horns are each covered by a lens to match the radiation pattern onto the telescope. The box containing the IF modules can be seen on the left.

18) Then the temperature that the detector sees, T_D, is given by:

$$T_D = T_S e^{-\tau} + T_C(1 - e^{-\tau}) \tag{23}$$

where τ is the opacity of the cloud.

The table shows some limits to this expression.

In CAPMAP, observing from New Jersey, clouds typically have a temperature of about 250K and an absorption of about 20%. Thus they attenuate the CMB signal by 20% and at the same time add a 45K atmospheric noise term. The table shows that the expression for T_D makes sense in the two limits: if the optical depth of the cloud is zero, the detector sees the source temperature directly while

Figure 16: **The Bell Labs Crawford Hill Radio Telescope Used by CAPMAP**

THE REVIEW OF SCIENTIFIC INSTRUMENTS VOLUME 17, NUMBER 7 JULY, 1946

The Measurement of Thermal Radiation at Microwave Frequencies

R. H. DICKE*
*Radiation Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts***
(Received April 15, 1946)

The connection between Johnson noise and blackbody radiation is discussed, using a simple thermodynamic model. A microwave radiometer is described together with its theory of operation. The experimentally measured root mean square fluctuation of the output meter of a microwave radiometer (0.4°C) compares favorably with a theoretical value of 0.46°C. With an r-f band width of 16 mc/sec., the 0.4°C corresponds to a minimum detectable power of 10^{-16} watt. The method of calibrating using a variable temperature resistive load is described.

INTRODUCTION

SINCE radio waves may be considered infrared radiation of long wave-length, a hot body would be expected to radiate microwave energy thermally. In order to be a good radiator of microwaves, a body must be a good absorber and the best thermal radiator is the "blackbody."

Although their discoveries were historically unconnected, there is a very close connection between "Johnson noise" of resistors and thermal radiation. The thermal fluctuations of electrons in a resistor set up voltages across the resistor. These "noise voltages" are of such a magnitude that a noise power per unit frequency of kT can be drawn from the resistor. k is Boltzmann's constant; T is the absolute temperature of the resistor.

The connection between thermal radiation and Johnson noise can best be shown by considering the system of Fig. 1.

An antenna is connected to a transmission line which is in turn terminated by a resistor. The radiation impedance of the antenna is assumed to be equal to the characteristic impedance of the coaxial line, i.e., the antenna is "matched" to the line. Also the resistor is assumed to "match" the line. When a transmission line is terminated by a "matched" load, the running waves in the transmission line incident on this load are completely absorbed without reflection. The antenna is completely surrounded by black

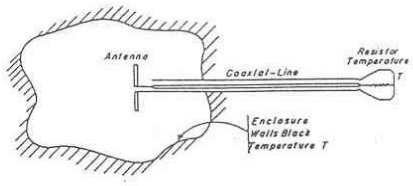

FIG. 1. Antenna system in black enclosure.

* Now at Palmer Physical Laboratory, Princeton University, Princeton, New Jersey.
** This paper is based on work done for the Office of Scientific Research and Development under contract OEMsr-262 with the Massachusetts Institute of Technology.

Figure 17: **The Cover Page of Dicke's Original Paper which Launched this Field**

if it is very large, the detector sees only the cloud temperature.

The atmosphere is not uniformly transparent in the microwave region. Figure 19 shows the absorption of the atmosphere vs. frequency; 90 GHz is a favored region. The absorption at that frequency is largely governed by water vapor.

3.3.2 Amplifier Noise

For an *ideal* amplifier, its power generated **with no input** depends only on its (physical) temperature T and the frequency ν:

$$p = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1} d\nu \Rightarrow kT d\nu \tag{24}$$

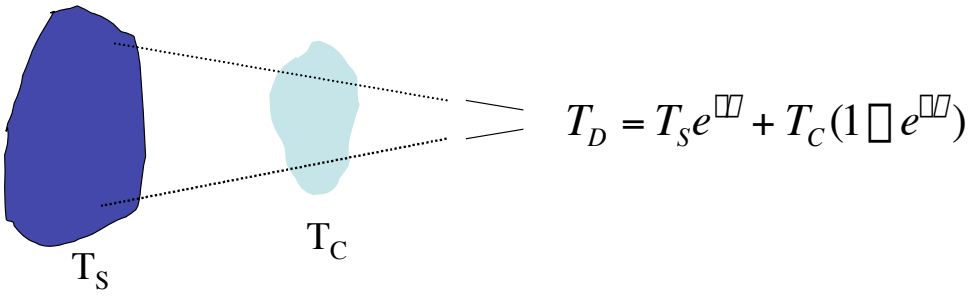

Figure 18: **Looking at a Source (S) Through a Cloud (C).** T_D is the detected temperature and τ the opacity of the cloud.

Table 2: How the Detected Signal Depends Upon the Optical Depth of a Cloud

Optical Depth τ	Detector Signal T_D
0	T_S
∞	T_C
0.2	45K ($T_C = 250$K

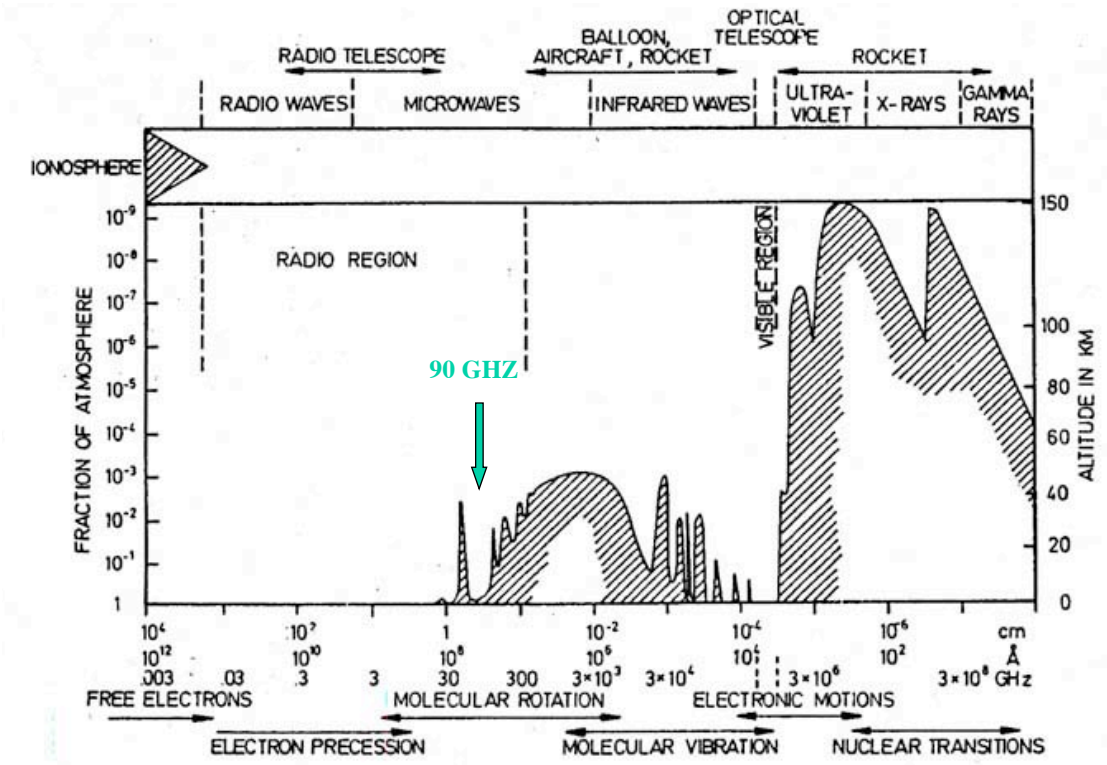

Figure 19: **The Optical Transparency of the Atmosphere.** Plotted[7] vs. frequency is the atmospheric height where the radiation is attenuated by 0.5. The 90 GHz region is one where CAPMAP and many other experiments can operate from the ground.

or kT power per unit band-width, when $h\nu \ll kT$. The state-of-the-art for 90 GHz amplifiers, which operate at $T_{physical}$=10K, is $T_{noise} = 45$K.

3.3.3 Quantum Noise

It is worth mentioning an irreducible noise level to such coherent amplifiers, the so called "Quantum Limit".

Here we sketch an argument for the existence of a quantum limit but we warn the reader that the effect is subtle; for a very complete description of all the subtleties, see ref.[6]

We are all familiar with the Energy-time uncertainty relation (consider an ideal wave packet):

$$\Delta E \Delta t = \frac{h}{4\pi} \tag{25}$$

Now lets relate ΔE, the energy uncertainty, to the uncertainty in the number of photons:

$$\Delta E = h\nu \Delta n \tag{26}$$

and the time uncertainty to that in the phase of the wave:

$$2\pi\nu\Delta t = \Delta\Phi. \tag{27}$$

This then gives a fundamental relation between the uncertainty in the phase of a wave and in the number of quanta in the wave:

$$\Delta\Phi\Delta n = \frac{1}{2}. \tag{28}$$

Considering such an ideal packet at the output of a phase preserving amplifier, one is led to a contradiction when referred back to the input of the amplifier. The only resolution to this problem is that the amplifier adds ("quantum") noise, at a level kT= $h\nu$ = 5.5K @ 100GHz.

At 90 GHz, this is still well below the state of the art but at lower frequencies, coherent amplifiers are fast approaching the quantum limit. This quantum noise is absent in totally absorbing detectors (bolometers) and this is party why it is thought that for the ultimate precision, bolometric detectors will likely be the best choice.

3.4 Receiver Sensitivity

For the situation we have been analyzing, the components of the signal, which add linearly, are:

- 3K from the CMB
- 45K from atmospheric noise
- 45K from amplifier noise

leading to a total system temperature of about 100K.

Now how do we calculate the sensitivity of a receiver with this temperature?

By sensitivity, we mean the uncertainty in the measurement of the temperature of a point on the sky. The result is:

$$\Delta T = \frac{T_{sys}}{\sqrt{\Delta\nu \times t_{obs}}} \approx 0.001K\sqrt{\text{sec}} \text{ (for CAPMAP)} \tag{29}$$

where $\Delta\nu$ is the receiver bandwidth and t_{obs} is the observing time.

We will derive this expression but first let's note that temperature sensitivity improves linearly with the system temperature but only as the square root of the bandwidth and the observing time. In a particle search, sensitivity will improve with the square root of time in a noise dominated regime, which is of course where we are in cmb detection. We emphasize that in both cases, it is very important to know (independently) the level of the noise.

We see that, with the system temperature and bandwidth of CAPMAP, each radiometer can measure a temperature to 1mK in one second. That figure is a benchmark for cmb radiometers: the WMAP radiometers also had approximately this sensitivity.

3.4.1 Dicke's Derivation of Radiometer Sensitivity

Dicke's original derivation used an argument that will be familiar to particle physicists: it is based on counting statistics.

We can view the signal coming out of a receiver as a pulse train. Our sensitivity will then be (inversely) proportional to the square root of the number of detected pulses.

Our ability to resolve pulses close in time depends on the frequency response of our full system. The greater the bandwidth, the closer we can resolve consecutive pulses. If $\Delta\nu$ is the bandwidth of the receiver, then we can see pulses separated by a time given by $\frac{1}{\Delta\nu}$.

Suppose that our system temperature is 100K and that we want to be sensitive to a temperature of $1\mu K$; since this is 10^{-8} of the system temperature, we would need 10^{16} counts to see such a temperature shift (at the 1σ level).

With a bandwidth of 10 GHz, we accumulate 10^{16} counts in 10^6 seconds, or about 12 days.

With this simple derivation, one sees clearly how the sensitivity is proportional to the system temperature and inversely proportional to the square root of both the observing time and the bandwidth. The great challenge, which we address later in this lecture, is to keep systematics associated with amplifier drifts, ground offsets, and atmospheric changes under control during such a large integration time.

3.4.2 POOR Signal to Noise!

To emphasize the noise regime in CMB experiments, consider CAPMAP. Take the radiometer sensitivity to be 1mk in 1 second. CAPMAP samples at 100 Hz and uses a 24-bit ADC over a 4-Volt range. At 100 Hz, the sample fluctuations will then be 10mK; and using a gain figure of approximately 10 mV per K, we see that the temperature corresponding to 1-bit is about $100\mu K$. We don't get 24 noise-free bits but even so, this is about 100 times the signal we are trying to detect! Thus the noise dithers the signal so that it becomes detectable. The author "discovered" this feature when trying to check his code by turning off the noise.

3.5 Calibration

It is important to be able to calibrate a radiometer in the lab before deployment and in the field. One can use black bodies of differing temperatures to check the gain, linearity, and sensitivity of the entire system. In the field, astronomical bodies such as the moon, a few of the planets, and known radio sources are frequently observed (e.g. daily) to provide relative, and sometimes even absolute, measurements of the system response. Figure 20 shows reconstructed temperatures of Jupiter with the four initial CAPMAP radiometers. The planets also serve as a good means to determine and check the pointing of the telescope as well as the shape of the beams on the sky.

3.6 Astrophysical Foregrounds

Of course measuring a sky temperature does not guarantee a pure CMB detection: one has to guard against various astrophysical "foregrounds". These include galactic emission due to synchrotron radiation of electrons in galactic magnetic fields, bremstrahhlung radiation of electrons off of atomic nuclei, and emission from hot dust. These foregrounds can be either minimized or measured using radiometers with multiple frequencies: none of these foregrounds has a black body spectrum. Also one can choose regions of the sky for observing which previous surveys indicate are relatively clean of such sources.

There are also extra-galactic sources that one must pay attention to. These include: known (and unknown) radio galaxies; hot gas in galaxy clusters that

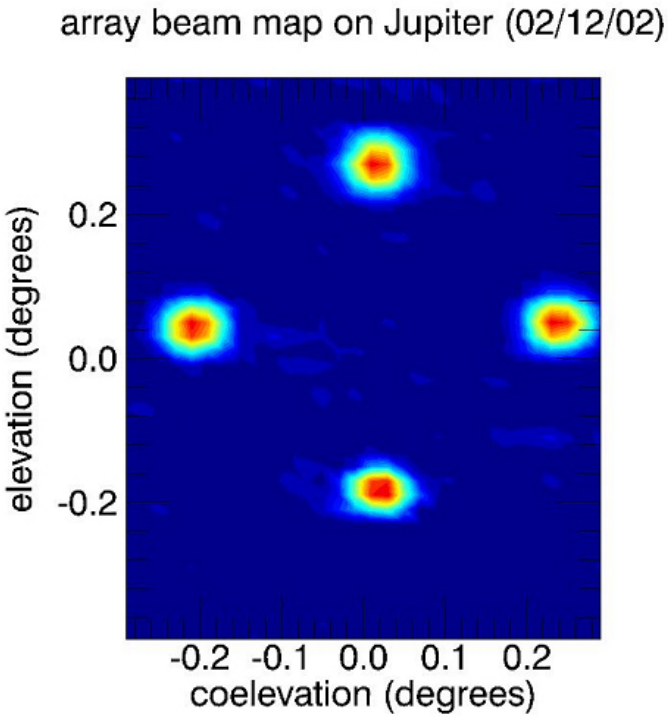

Figure 20: **Jupiter as Observed by Four CAPMAP Radiometers.**

interacts with the cmb, effectively boosting its temperature; and the gravitational lensing of the cmb due to gravitational potential of the intervening matter. Existing catalogues, good angular resolution, and statistical tests are the techniques used to control the first two of these effects; the third, gravitational lensing, is the more subtle and interesting foreground and we will discuss this effect in the 3rd lecture.

3.7 Instrumental Effects

We list here the most important instrumental challenges in achieving sensitivities at the parts-per-billion level of the system temperature.

- Amplifier Drifts: These must be controlled to a very high level. We will describe the technique of "Dicke switching", universally employed in all cmb experiments
- Mechanical pickup: Something always moves in cmb experiments: the tele-

scope moves; and mechanical refrigerators are commonly used. Electrical signals can be modulated due to microphonics.

- Optics/ground pickup: We have mentioned that it is important to shield the radiometers from the 300K ground. Even if the shielding is effective at the level of 10^{-5}, so that the contribution from the ground is just 3 milli- K, this ground pickup will certainly vary with time (as the ground temperature varies) and, more importantly, as the telescope moves since the side-lobes will be illuminating different patches of the ground.
- Thermal regulation: Since amplifier gains vary with temperature, it is important that they be kept at constant temperature.
- Electrical Grounding: Significant attention needs to be paid to grounding issues: one needs very secure connections and one needs to guard against the inevitable level shifts when, for example, heaters controlling the amplifier temperatures, which draw significant currents, are cycled.

3.8 Dicke Switching

We will go into this very clever technique in some detail as it is probably the most important of all those used in cmb experiments.

We want to consider the effect of a gain drift, ΔG in determining the temperature at one spot on the sky, T_1. The power detected, W, will be given by:

$$W = k(T_1 + T_{sys})G\Delta\nu. \tag{30}$$

Now the change in power, ΔW, when the gain drifts is give by:

$$W + \Delta W = k(T_1 + T_{sys})(G + \Delta G) \tag{31}$$

which then gives the power change due to a gain drift of:

$$\Delta W_{\Delta G} = \Delta G \Delta\nu k(T_1 + T_{sys}) \tag{32}$$

Now let's compare this with the power change due to an increment ΔT in the signal:

$$W + \Delta W = k(T_1 + \Delta T + T_{sys})G\Delta\nu \tag{33}$$

so that the change in power due to a signal increment becomes:

$$\Delta W_{\Delta T} = G\Delta\nu k\Delta T. \tag{34}$$

Equating these two power changes, we find:

$$\frac{\Delta T}{T_{sys}} = \frac{\Delta G}{G}, \tag{35}$$

i.e. a temperature sensitivity that is proportional to the (fractional) gain drift.

Now let's consider a differencing experiment, where one measures power at two spots:

$$W_1 = k(T_1 + T_{sys})G\Delta\nu. \tag{36}$$

and

$$W_2 = k(T_2 + T_{sys})G\Delta\nu. \tag{37}$$

The difference in power levels is then:

$$W \equiv W_1 - W_2 = k(T_1 - T_2)G\Delta\nu \tag{38}$$

The change in (the difference in) power due to a gain drift is:

$$W + \Delta W_{\Delta G} = k(T_1 - T_2)(G + \Delta G)\Delta\nu \tag{39}$$

while that from a signal change is:

$$W + \Delta W_{\Delta T} = k(T_1 - T_2 + \Delta T)G\Delta\nu. \tag{40}$$

Then the sensitivity to a change in temperature becomes:

$$\frac{\Delta T}{T_{sys}} = \frac{T_1 - T_2}{T_{sys}}\frac{\Delta G}{G}. \tag{41}$$

Now, the temperature sensitivity is proportional to the temperature *difference* between the two spots. A gain change can modulate this difference but, unlike in the previous non-differencing case, it cannot fake a temperature difference. This is the key benefit of Dicke switching.

One can see the effects of such switching in the following figures. Figure 21 shows just 8 seconds of data (from CAPMAP) for two channels, one switched and the other not. The samples are recorded every 0.01s; one can see that the switched channel looks much more "white" with an rms on the order expected: $1\mathrm{mK}\sqrt{\mathrm{sec}}$ while the other displayed channel shows drifts even over this very short period of order 0.5 K. In this case the drifts are a combination from both the atmosphere and the amplifiers.

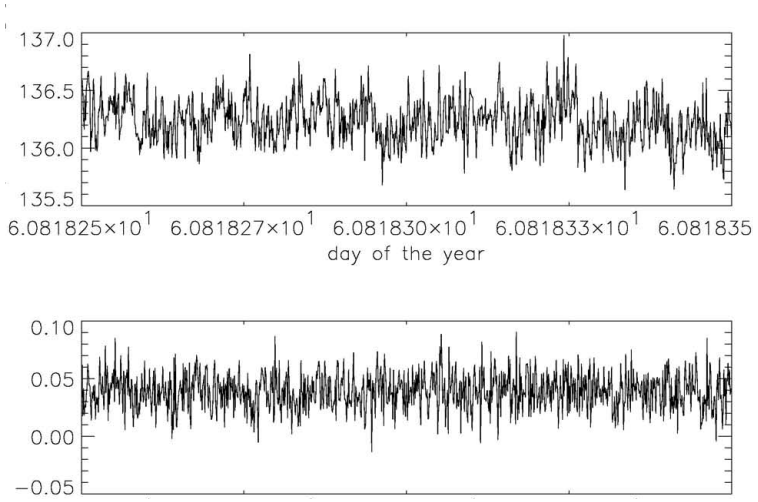

Figure 21: **Eight Seconds of CAPMAP Data.** The signal is sampled at 100 Hz. The upper plot is an "unswitched" channel in which non-random behavior is evident; the lower plot, which is much more Gaussian, is effectively a "switched" output (the switching will be explained in lecture 3).

Figure 22 shows the power spectrum of two similar channels but this time on time scales from 100 seconds to a day. The unswitched channel shows a rising noise spectrum as the frequency decreases, a "1/f" spectrum typical of gain, and in this case, atmospheric drifting, and that for a switched channel which even on scales up to a day is consistent with being purely white. (Historically "Dicke Switching" meant suitably rapidly switching one's beam between two sources or "loads." Another form of "Dicke Switching" is to look at the same point on the sky and **electronically** rapidly modulate the signal and look for the residual that remains after demodulation. The example shown in figure 22 is a combination of both types of switching.)

3.9 "Dicke Switching" in Particle Physics

Similar techniques have been used in particle physics. At SLAC, the experiment[8] measuring parity violation frequently (and randomly) switches electron helicities. At Fermilab, the KTeV experiment had two kaon beams side by side and its most important physics came from ratios of rates from the two beams, which also were switched periodically. And historically, the Fermi-Marshall experiment that

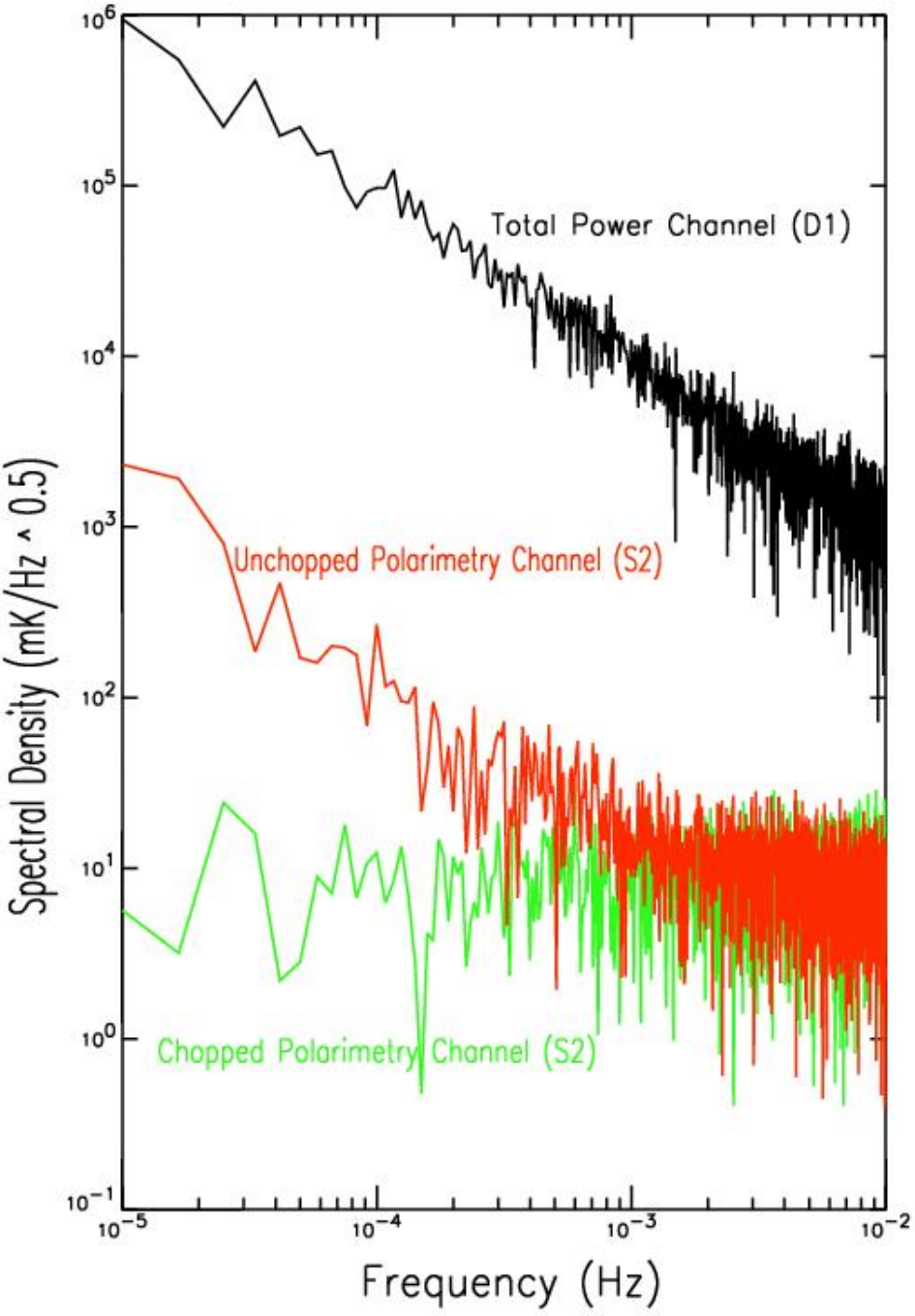

Figure 22: **Noise Power spectra from PIQUE, the predecessor to CAPMAP.** The upper (black) curve is for an "unswitched" output while the lower (green) curve is for the "switched" output. The effects of amplifier (and atmospheric) drifts on scales from a few minutes to a day are shown: the noise on the "switched" output is white while that on the "unswitched" one has a dominant 1/f behavior. Figure from the Ph.D. Thesis of M. Hedman.

measured neutron-electron scattering for the first time, used a very similar idea to Dicke's. It is doubtful that Dicke and Fermi communicated about this: the idea has been "discovered" independently by different scientists.

3.10 The experimental sensitivity to the power spectrum

Equation 13 gave the error in a measurement of C_l from "cosmic variance", i.e. at a given l there are only $(2l+1)$ "modes" with which to sample and determine the variance.

That expression is valid in the regime where one is measuring the entire sky and where the instrument noise is negligible. It also neglected the effect of finite angular resolution. In this section we will get a feel for how these effects compromise the sensitivity to power spectra.

3.10.1 Detector Noise

When detector noise is included, the expression for the error on a given C_l becomes:

$$\frac{\Delta C_l}{C_l} = \sqrt{\frac{2}{2l+1}} \left(1 + \frac{4\pi w^{-1}}{C_l}\right) \tag{42}$$

where w^{-1} is the total experimental "weight"; it is the grand error one would obtain were all the experimental sensitivity concentrated in one measurement. For example, a radiometer with the bench-mark sensitivity of $1\text{mk}\sqrt{sec}$ operating for 10^6 sec would have $w^{-1} = (1\mu K)^2$.

For the WMAP results (see figure 6), $w^{-1} = 0.041\mu K^2$. This "noise" term does not dominate the cosmic variance term until $l \geq 600$.

3.10.2 Observing over a finite patch of sky

It is often advantageous to concentrate one's observing into a finite patch of sky. Here one gives up on the smallest multi-poles but does better for higher multipoles. The expression for the error on a given C_l becomes:

$$\frac{\Delta C_l}{C_l} = \sqrt{\frac{2}{2l+1}} \left(\frac{1}{\sqrt{f_{sky}}} + \frac{4\pi w^{-1}}{C_l}\sqrt{f_{sky}}\right) \tag{43}$$

where f_{sky} is the fraction of the sky under observation.

The smaller the region scanned, the detector noise term decreases while the cosmic variance term increases: the effective number of modes is compromised.

3.10.3 Finite Beam Resolution

Finally we consider the effect of finite beam resolution. Typically the Gaussian approximation is used, at least in the design phase of an experiment. If the telescope beam has resolution σ_b, then the expression for the error on a given C_l becomes:

$$\frac{\Delta C_l}{C_l} = \sqrt{\frac{2}{2l+1}} \left(\frac{1}{\sqrt{f_{sky}}} + \frac{4\pi w^{-1}}{C_l} \sqrt{f_{sky}} e^{l^2 \sigma_b^2} \right) \tag{44}$$

It is clear that finite beam resolution degrades the sensitivity progressively worse at the higher l values. Figure 23 shows a convolution of a power spectrum with the beam shape for various angular resolutions. The WMAP beam of about 0.2 degrees limits sensitivity beyond l of about 600 while smaller beams are needed to go much higher.

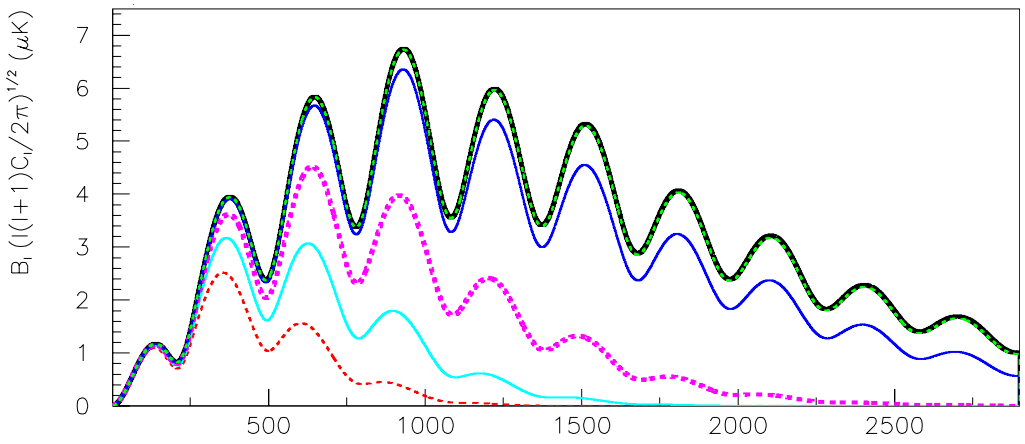

Figure 23: **A Power Spectrum Convoluted with a Finite Beam.** The upper curve is the actual power spectrum while the lower ones show the multipole distribution of the recovered power as it is affected by the finite beam size. The second (blue) curve shows the effects of the CAPMAP beam of 0.05 degrees while the 4th (magenta) shows that for WMAP with a beam of 0.24 degrees.

3.11 Choosing an Observing Stragtegy

Part of the art cmb science is choosing an observing strategy. Given an l-range on which one wants to concentrate, and a detector sensitivity, one can choose the

size of the sky patch, using equation 44, for optimum sensitivity. The next issue is how one best covers the patch.

In spite of any internal instrumental switching that takes place, there are still "offsets" that have to be dealt with. Another way to say this is that it is next to impossible to measure the sky temperature at a point sufficiently accurately to probe the fundamental anisotropies. The FIRAS instrument measured the blackbody temperature to of order 10^{-3}K, a fantastic measurement but nowhere near accurately enough to see even the largest of the anisotropies which are at the 100μK level.

That is why one must measure differences in temperature between different points on the sky. It is important that radiometers revisit the same points on the sky frequently and with different time scales; in this way (hopefully) slow drifts in response can be tracked and removed, with some usually small loss in sensitivity.

Figure 24 shows the CAPMAP observing strategy. A small patch of sky centered on the North Celestial Pole is observed. A beam of about 0.06 degrees is swept across the patch at a frequency of about once per 8 seconds, data being recorded at 100 Hz. Thus many swings past the same pixels are made, this while the sky is slowly rotating under the beam. In this way one can cover the entire patch in 12 hours and when this is done for a few months, many consistency checks can be made to assure that one can co-add all the data to give the final sensitivity.

3.12 Data Processing

Here we give the key steps in the processing of the data. Again we'll use CAPMAP as an example but all of steps are generic.

- Calibrate the Radiometers: this means turning measured voltages into degrees K, using primarily various known sources observed throughout the season.
- Data Cleaning: inevitably there will be bad data in the sample, coming either from bad weather, occlusions of the beam from, for example, birds or airplanes, or detector malfunctions. As in HEP experiments, it is helpful if the selection of good data periods can be done "blind", i.e. by looking at channels not directly used for the result or looking at the numerous monitors of, for example, component temperatures or currents.

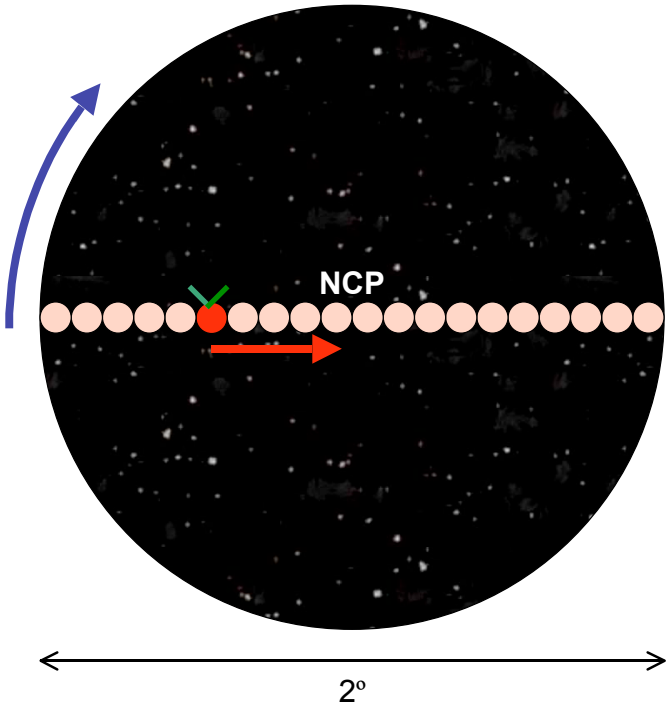

Figure 24: **How CAPMAP Scans the Sky.** The region covered is roughly 1 degree radius around the North Celestial Pole. One beam is shown in red; it scans back and forth across the diameter of the region, completing a cycle every 8 seconds, this while the sky rotates underneath the scan. (The radiometers measure the power **difference** in the two electric field directions shown by the green lines.) Figure courtesy of Peter Hamlington.

- Offset removal: in almost all CMB experiments, it is necessary to remove the mean signal seen for each scan. In the CAPMAP case, this means the mean of the 20 pixels across the NCP every 8 seconds. This does not result in too great a loss in sensitivity- one is removing just one degree of freedom out of 20. But there is also a residual slope across the scans, one that arises because the beam side-lobes hit the ground and that contribution makes a small but changing slope across the scan, hence the loss of another degree of freedom. Finally, over the whole season, there are residual offsets that survive both of these "subtractions". One must strive to make sure these are as low and as stable as possible. The analysis is done so that the results are independent of any such constant shape, resulting is a reduction of only 3% in sensitivity.

This residual offset needs to be "constant" only on the time scale of a day and one has to study the data to be sure this is indeed true.

- Bin in Sky Coordinates and Make a Map: one then pixelizes the sky- in the case of CAPMAP we have 20 pixels across the NCP and 72 around (corresponding to 20 minutes of observations each day). Then all the measurements at each pixel are co-added to make one grand map. Also for each pixel, we need to determine the error; this is taken from the scatter of the measurements on roughly 1 minute time scales. Since the ultimate error results from combining measurements taken over months, care must be taken to ensure that the data properly "integrates down."

- Run likelihood for Power in l-bands. The easiest way to think about fitting for the "fluctuation power" is the following: suppose you have a variable that is fluctuating with an unknown *signal* rms σ_S that you want to determine. Suppose further that the measurements also have a noise rms σ_N. With N samples of this variable, what is the best measure of σ_S?

  We can write the probability (likelihood) for σ_S given the N data measurements Q_i as follows:

$$L(\sigma_S) = \prod_{i=1}^{N} \frac{1}{\sqrt{\sigma_S^2 + \sigma_{N_i}^2}} e^{-\frac{Q_i^2}{2(\sigma_S^2+\sigma_{N_i}^2)}} \tag{45}$$

  In this simple example, all the measurements are independent. In practice, the measurements at the different pixels are coupled, i.e. the covariance matrix is non-diagonal and this has to be taken into account.

  The more general expression, where C is the covariance matrix (with a dimensionality of N×N) becomes:

$$L(\alpha) = [detC]^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\vec{\mathrm{x}}^{\mathrm{T}}\mathrm{C}^{-1}\vec{\mathrm{x}}\right] \tag{46}$$

  where $\vec{x}$ is the data vector, and α is a set of theory parameters for which one is fitting.

  Here the covariance matrix, C=S+N, is the sum of that from theory, $S(\alpha)$ and N, the experimental noise matrix. Often the latter can be taken as diagonal but the former encodes the expected pixel-pixel correlations, given by a theory power-spectrum.

- Perform "null" tests on the data. Finally it is important to perform various tests on the data to be sure that any signal is robust. For example, there should be no signal in a data set where the first half of the data is subtracted from the second. On shorter time-scales, it is useful to form difference data sets where one can similarly test for systematics. For example, no signal should be seen when the telescope is moving to the right when compared (subtracted) from when it is moving to the left. And data taken during daylight hours, when subtracted from that taken at night, should again similarly give a null result.

We now leave our study of experimental issues to discuss the most important questions for the next decade.

4 Lecture 3

In this final lecture, we consider the efforts for the coming decade. We will discuss the generation of gravity waves during the inflationary era and show how sensitive CMB experiments need to be to detect these waves. The polarization of the CMB must be detected and we describe how polarization is generated and the techniques for its detection. For these studies we will see that it will be important to understand the Gravitational Lensing of the CMB. After describing some detectors of the future, we close with some final remarks.

4.1 Primordial Gravity Waves

As was emphasized in the lecture of Michael Turner, tensor perturbations to the space-time metric would be generated during inflation, at least during slow-roll inflation. The mechanism is the same as that which produces density or scalar modes.

Just like the density perturbations, inflation cannot predict the level of the tensor modes. The parameter $r = T/S$ is the Tensor to Scalar ratio; it depends upon the energy scale of inflation V:

$$V^{0.25} = 0.003 M_{pl} r^{0.25} \tag{47}$$

A value of $r = 0.001$ corresponds to $E_{inflation} = 6.4 \times 10^{15}$GeV.

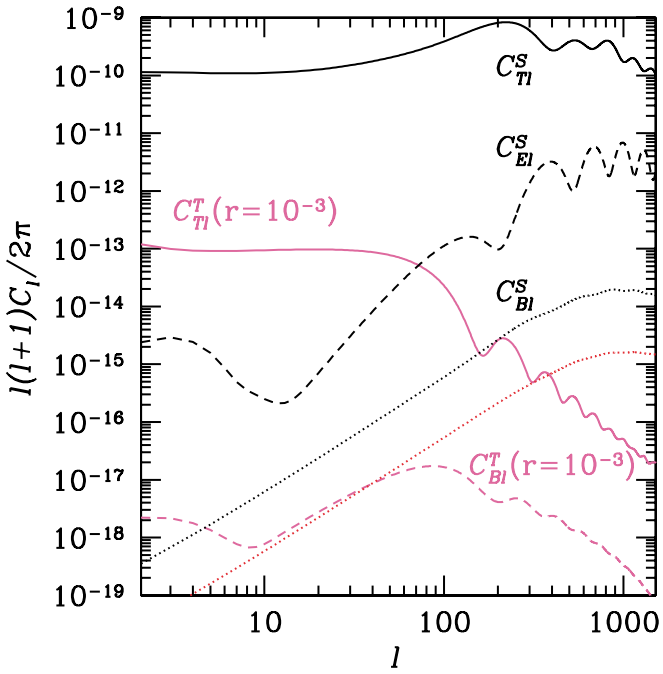

FIG. 1 (color online). Angular power spectra. The solid lines are for temperature anisotropies due to scalar perturbations, C_{Tl}^S and tensor perturbations C_{Tl}^T with $r = 10^{-3}$. The dashed lines are for the E modes from scalar perturbations C_{El}^S and the B modes from tensor perturbations C_{Bl}^T. The dotted lines are for the lensing-induced scalar B modes C_{Bl}^S before (above) and after (below) the cleaning that can be done by a perfect experiment. The feature at $\square < 10$ is due to reionization which we assume occurs at $z_{\rm ri} = 7$.

Figure 25: Expected Power Spectra.[9]

The parameter r can be limited by studying ΔT but because of the shape of the temperature anisotropy from tensor modes, only a very weak (but still interesting) limit is possible: $E_{inflation} \leq 2 \times 10^{16}$GeV. More promising is the polarization field. Figure 25 shows the expected power spectra. C_{Tl}^S is the familiar temperature anisotropy from scalar modes; C_{Tl}^T is that from the tensor modes: it is easy to see that a small amount of tensor modes cannot be distinguished from the scalar modes. C_{El}^S and C_{Bl}^T are polarization spectra from scalar and tensor modes respectively. The latter holds the promise for a direct look at inflation.

4.2 Sensitivity

From the expected power spectrum in figure 25, we see that to detect the gravity waves, we need to concentrate on the range $50 \leq l \leq 120$. The upper range is, as we have seen, the horizon scale at de-coupling; the finer-scale modes are red-shifted away. The reason is simple (although, like most "simple" things, it took the author some time to understand): gravity waves cause a shear but not an over-density so the perturbations do not grow with time. The signal is, however, some 7 orders of magnitude smaller than that from the density fluctuations.

Let's calculate the sensitivity needed to make a $3-\sigma$ detection of this signal, with $r = 0.001$. We use equation 42 (which assumes an all-sky map) for the error on C_l given the experimental weight. From the figure we see that the peak of the power spectrum (at $l \approx 90$) is about 2×10^{-17}. Since what is plotted is $l(l+1)/2\pi)C_l T_{cmb}^2$, we have that at the peak $C_l = 0.12(nK)^2$. We can assume a $\Delta l = 70$: we are sampling the power in 70 separate and uncorrelated multipoles. Then the expression for C_N that would yield a $3-\sigma$ detection becomes[†]:

$$\sqrt{\frac{1}{90 \times 70}}\left[1 + \frac{C_N}{0.12}\right] = \frac{1}{3} \tag{48}$$

so that $C_N = 3(nK)^2$. Since the WMAP sensitivity for its first year is of order $21,000(nK)^2$, we see that the equivalent sensitivity of about 7000 "WMAP's" would be required for this measurement!

How much better could we do if we optimize the sky patch? For the calculation we just did, the error from detector noise is about 25 times greater than from cosmic variance. For the minimum error, it is easy to show that the two terms should contribute equally; hence, by reducing the coverage until the sample variance increases to the (decreasing) value of the detector noise (see equation 44), we will have an optimized experiment.

In this optimized case it is easy to show that for a $3-\sigma$ detection of a feature centered at multi-pole l_0, having a width $\Delta l = l_0$, we should have:

$$f_{sky}^{opt} = \left(\frac{6}{l_0}\right)^2, \text{and} \tag{49}$$

$$w_{opt}^{-1} = \frac{l_0}{6}\sqrt{\frac{C_l}{24\pi}} \tag{50}$$

[†]We are here assuming that the beam size is small enough to not compromise the sensitivity in this l-range.

where C_l is the power spectrum evaluated at l_0. For the case under consideration, we find $f_{sky}^{opt} = 4.4 \times 10^{-3}$ and $w_{opt}^{-1} = 2 \times 10^{-9}$K. This corresponds to about 400 WMAPs, or a "savings" of a factor of about 15 over the all-sky approach.

In general, the "savings" from reducing sky coverage to the optimum value turns out to be a factor of $\frac{l_0}{6}$, in running time or number of detectors.

We have been talking about the requirements for the detection of polarization but we have not yet mentioned how the polarization is generated. Hence the next section.

4.3 How Polarization Gets Generated

If there is a quadrupole anisotropy in the temperature field around a scattering center, even if the incident radiation is unpolarized, the scattered radiation will be as shown in figure 26: a linear polarization is generated. The quadrupole is scattered during decoupling as was shown is figure 8. Since the polarization arises from scattering but said scattering will dilute the quadupole, the polarization anisotropy is much weaker than that in the temperature field. The polarization also peaks at higher $l-$values: it is a small scale effect, the only scale that can support the quadrupole.

Two example polarization fields are depicted schematically in figure 27. The length of the lines is the strength of the polarization while their direction indicates how the electric field tends to line up. This (tensor) field can be decomposed into two components, in analogy with electrodynamics: "E" modes (with no handedness, shown on the left) and "B" modes (with a handedness, shown on the right). E-modes arise from the density perturbations that we have been treating while the B-modes come from the tensor distortions to the space-time metric induced by gravity waves (which do have a handedness).

The signature of inflation that we have been discussing could be detected if the B-modes can be effectively isolated, this in the presence of a larger E-mode field. It appears that this decomposition can be made at least at the level of a few percent and possibly smaller, depending upon the level of sensitivity and the geometry of the region observed.

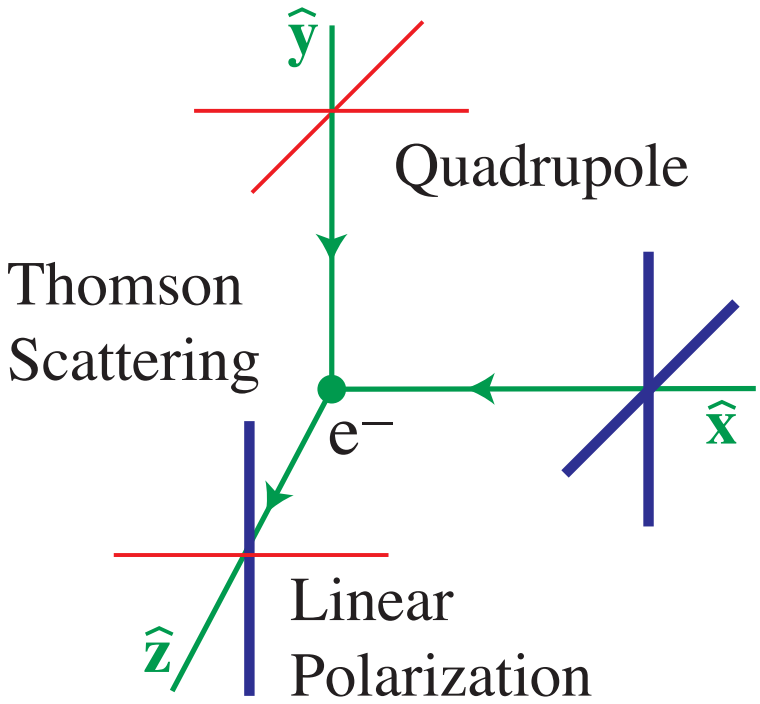

Figure 26: Generation of Polarization. Unpolarized but anisotropic radiation incident on a scattering center produces polarized radiation. Intensity is represented here by line thickness; the quadrupole pattern produces a linear polarization in the $\hat{z}$ direction. Figure from W. Hu.

4.4 The Measurement of Polarization

This section discusses one technique for the detection of polarization, the one used by the CAPMAP experiment. The idea is shown schematically in figure 28. The "brute-force" method would be the following: align a polarizing filter in the a direction and measure, after amplification the power. Do the same thing in the b direction where b is orthogonal to a. Then the difference is a measure of (one of the Stokes parameters of) the polarization.

The reason this is "brute-force" is because the result is dependent, among other things, on the stability of the amplifiers: any gain shift between the two measurements will generate a polarization where there is none to begin with. Many of the early polarization limits were obtained using just this technique.

CAPMAP and other experiments use what is called a "correlation polarimeter", which is in essence a sophisticated means of Dicke-switching. The radiation

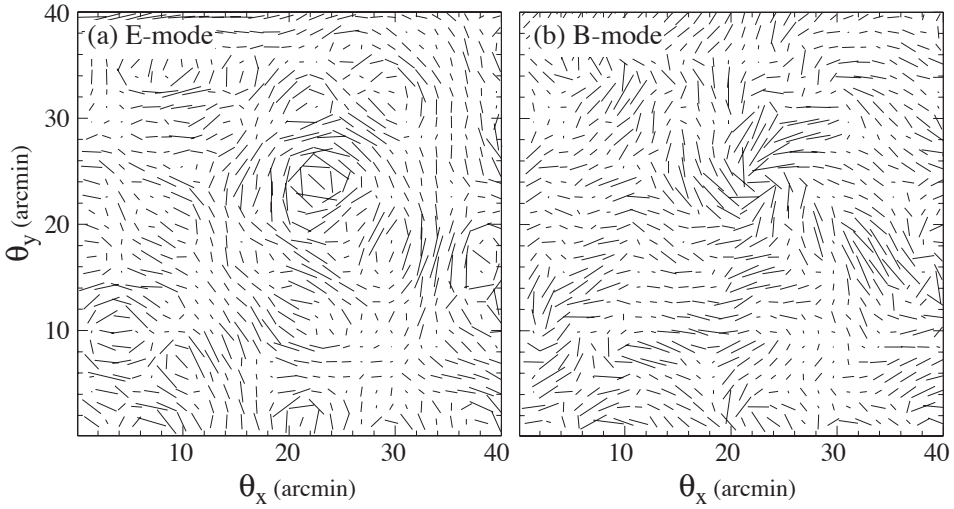

Figure 27: **Polarization Patterns.** "E" is shown on the left; "B" on the right. The pattern shown here (from W. Hu) is actually from a simulation of the lensing (described later in this lecture) of the CMB by the large-scale structure in the universe.

incident upon the horn passes through an "ortho mode transducer" (OMT) which magically sends polarization in the $x-$direction out one arm and that in the $y-$direction out the other. (See Figure 13.) These are separately amplified and then *multiplied* (in CAPMAP this is done at the IF stage). The result is:

$$G_x E_x G_y E_y \propto G_x G_y (E_b - E_a)(E_b + E_a) = G_x G_y (E_b^2 - E_a^2). \tag{51}$$

This is precisely what we were trying to measure with the "brute-force" experiment but this time we see that the result is proportional not to the difference but to the product of the amplifier gains. This product multiplies the small power difference on the sky so that if the gains drift, they no longer generate a fake signal.

4.5 Polarization Experiments

There are a large number of polarization experiments ongoing or in the planning stages. The current status is shown in figure 29, which shows the results from the DASI interferometric experiment together with those projected from capmap. Figure 7 showed the correlation between T and E as reported by WMAP. (WMAP has sensitivity to the EE power spectrum as well.)

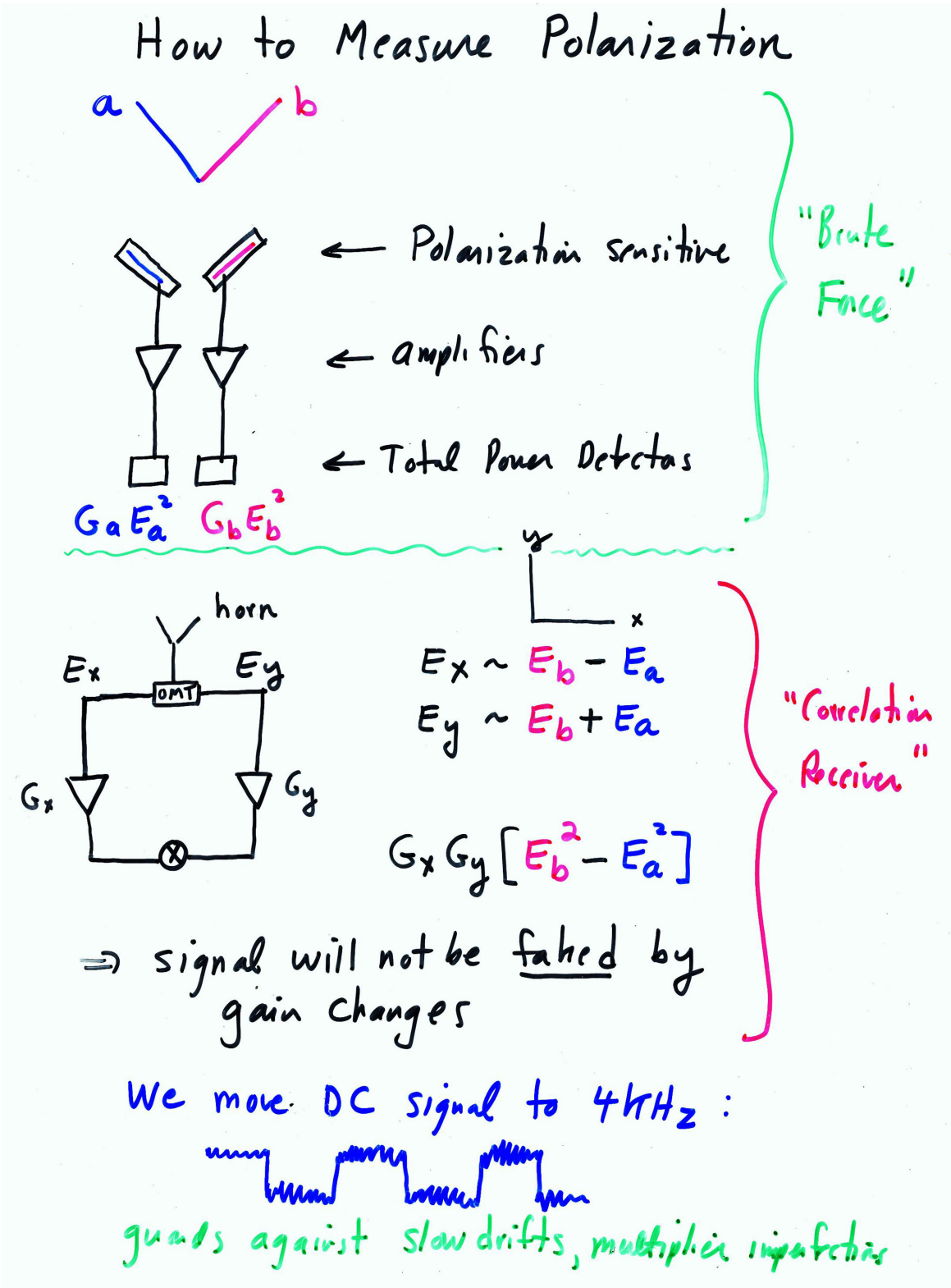

Figure 28: **The "Brute Force" and Systematically Freer Approach to Measuring Polarization.**

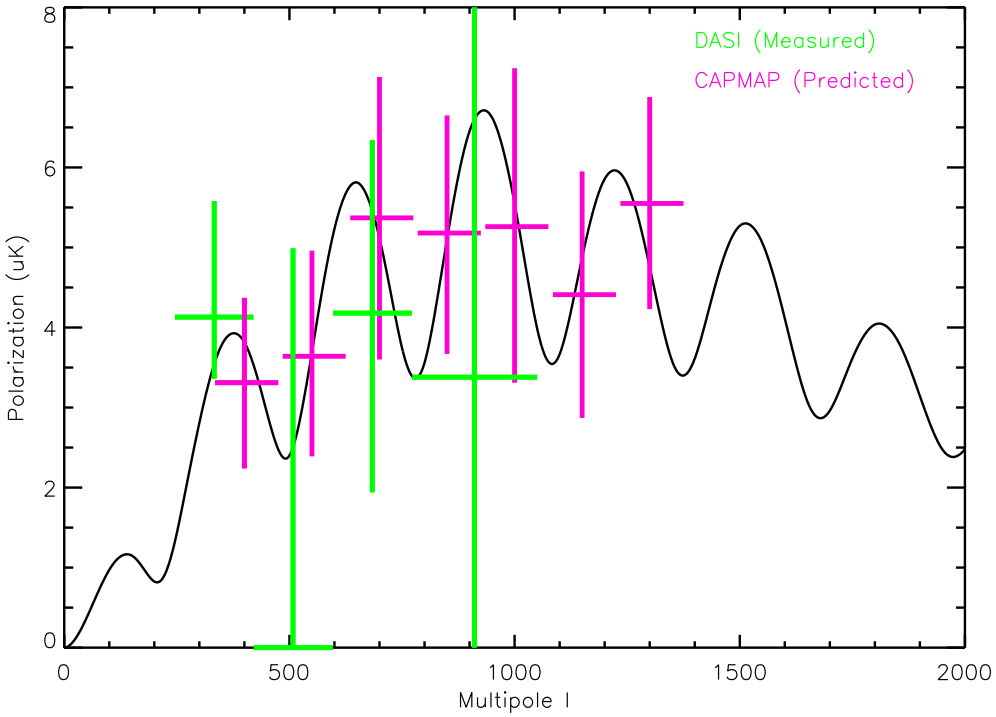

Figure 29: **DASI Results and CAPMAP Projections on the Polarization Power Spectrum.** The solid line is close to the expected spectrum.

These experiments, and all others that have reported significant limits, use coherent detectors. Many experiments plan on using the inherently more sensitive bolometric detectors. These are very stable and have been used successfully for T-anisotropy experiments such as Boomerang and Maxipol. These are incoherent so in essence they use what we have called the "brute-force" technique although very clever means have been devised to both make them polarization sensitive and able to modulate the polarization sufficiently rapidly.

Figure 30 shows for many of the experiments the frequency coverage, the size of the beam, the site, and the technique used. It is good for determining the best techniques that these experiments span so many parameters differently.

4.6 Contaminants to Gravity-Wave B-Mode Polarization

As we have shown, the B-mode signal is likely to be very small so that one has to worry about foregrounds. The galaxy may very well dominate at these signal

	Frequencies (#)	*Beam*	*Site*	*Technique*
POLAR	30 (1)	7°	WI	Correl. Rad., axial spin
COMPASS	30 (1), 90 (1)	20', 7'	WI	Correl. Rad., NCP scan
PIQUE	40 (1), 90 (1)	30', 15'	NJ	Correl. Rad., NCP chop
CAPMAP	40 , 90	13', 6'	NJ?	Correl. Rad. Array
DASI	30 (13)	20', 7'	S. Pole	Interferometer
CBI	30 (13), 90 (13)?	3'	Atacama	Interferometer
VLA	8.4	6''	Socorro	Interferometer
Polatron	90 (1)	2'	OVRO	Bolo,1/2 λ plate
QUEST	150 , 225 (~30)	4', 3'	Chile?	Bolo Array, 1/2 λ plate
POLARBEAR	150 … (3000 dt'rs)	10'	S. Pole or M. Kea	Bolo Array
BOOM2K	150 (4), 240 (4), 340 (4)	10'	Antarctic LDB	Bolo Array
MAXIPOL	150 (12), 420 (4)	10'	US-Balloon	Bolo Array,cold 1/2 λ plate
BaR-SPOrt	32, 90	30', 12'	Antarctic LDB	Correl. Rad. Array
MAP	22, 30, 40(2), 60(2), 90(4)	13'	L2, full-sky	Correl. Rad. Array*
SPOrt	22, 32, 60, 90	7°	ISS, full-sky	Correl. Rad. Array
PLANCK-LFI	30(4), 44(6), 70(12),100(34)	33',23',13', 10'	L2, full-sky	Correl. Rad. Array
PLANCK-HFI	100(4), 143(12), 217(12), 353(6), 545(8), 857(6)	11', 8', 6', 5', 5', 5'		Bolo Array

Figure 30: **A Compilation (by Peter Timbie) of Most Current Polarization Experiments.**

levels: dedicated studies are needed but it is likely that we won't know the level of the galactic foregrounds until sensitive experiments are underway. WMAP will be able to help guide the way with its next release of its polarization data, but not at the sensitivities we ultimately will need.

One of the foregrounds, the one that has been most studied, is also of cosmological significance. It is discussed in the following section.

4.6.1 Gravitational Lensing of the CMB

The CMB, on its way from the surface of last scattering, does not quite travel unimpeded to today's detectors. When it passes huge concentrations of matter, it gets "lensed" by the gravitational potential. This lensing effect causes deflections of a few minutes of arc, deflections that are coherent over degree scales-the (projected) scale of the large-scale structure in the universe. Figure 31 shows a simulation of these effects: the lensing potential distorts both the temperature field and the E-polarization field but has its largest effect in generating a

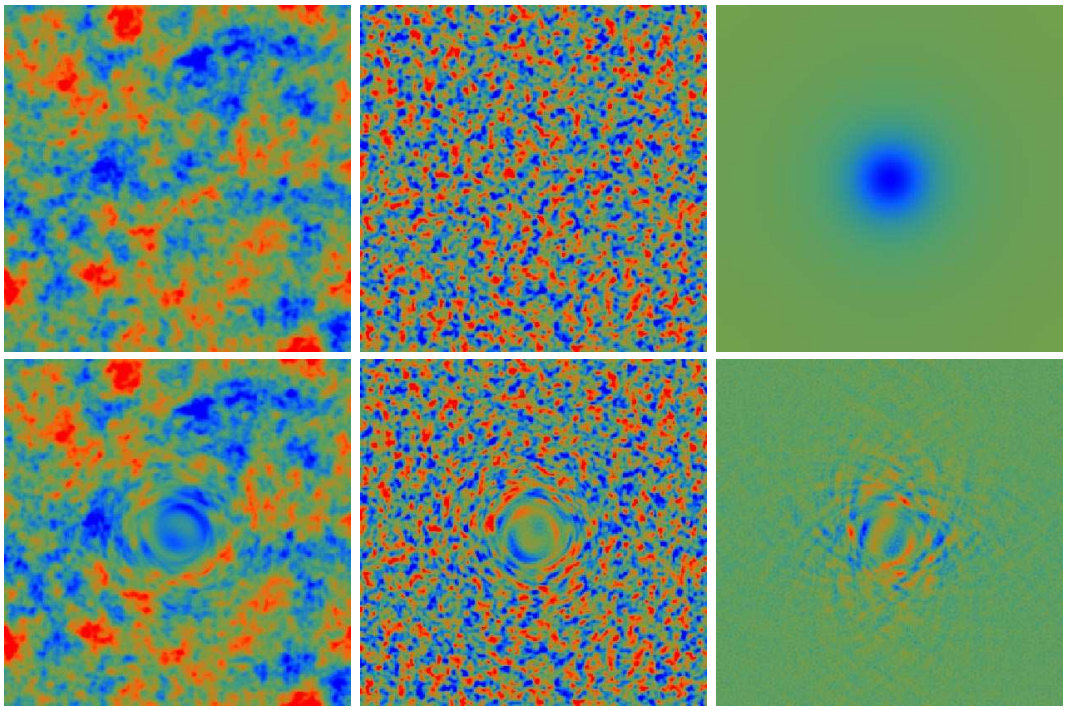

Figure 31: **Gravitational Lensing of the CMB.** This graphic (taken from[10]) gives insight into the lensing of the CMB. All panels are for a 10^0 by 10^0 field. The upper-left panel shows a simulated temperature field; upper-middle the associated E-polarization field. The lower-left panel shows the lensed temperature field; lower-middle the lensed E-polarization field; and lower-right the lensed B-polarization field. The deflection field for this simulation, an appropriate convolution of the matter distribution between the surface of last scattering and today, is shown in the upper-right panel.

B-polarization field. This B-field can be "subtracted" to some extent (see figure 25) by cleverly using the information in very deep surveys with excellent angular resolution but this may represent an ultimate limit for the detection of the primordial B-modes.

Figure 25 shows the power spectrum of these B-modes and the level to which they can be "cleaned" using one technique. For these modes, because of the coherence of the lensing potential, there is more information than just the power spectrum and work is ongoing to characterize the expected cross-correlation between different multi-pole bands.

4.7 Sensitivities Needed

We can now compile the sensitivities needed to be able to detect the primordial B-modes at various levels. We showed this calculation in section 4.2 and we extend it here, including the lensing contaminant. We see that, to reach the lowest levels that our theorists are challenging us to reach, we will need great advances in overall sensitivity.

Table 3: Sensitivities to Detect Polarization Signals. N is the equivalent number of "WMAP's" needed for a 3 standard deviation effect, given for both the full sky and optimized scenarios. The W values in this table are the total required sensitivities of the detectors.

Signal	W (full)	N_{wmap}^{full}	W (opt)	f_{sky}^{opt}	N_{wmap}^{opt}
"E", $l = 1000$			$300nK$	1.3×10^{-3}	.02
lensing, $l = 1000$			15 nK	1.3×10^{-3}	8
"B", $r = 10^{-3}$	500 pK	6,400	2 nK	4.4×10^{-3}	400
"B" , $r = 10^{-4}$	170 pK	64,000	630 pK	4.4×10^{-3}	4000
"B", cleaned $r = 10^{-4}$	100 pK	150,000	370 pK	4.4×10^{-3}	9500

It is sobering to realize that to see inflation at a level even as large as 6.4×10^{15} GeV, corresponding to T/S=0.0001, one needs, in an optimized experiment, the equivalent sensitivity of 10,000 WMAPs. Since WMAP is at the level of 100 times COBE, it is likely to take some time before such a sensitivity is reached.

4.8 The Opportunity of Reionization

If the WMAP indication of the optical depth to reionization is correct, it means there is a significant additional scattering surface for the generation of polarization. The gravity wave signature will show up at much lower $l-$values, corresponding to the horizon scale at reionization: instead of peaking near $l = 100$, the best place to see the B-modes from reionization is actually at $l = 3$! The only means of detecting such a signal is from space; and even there it will be very hard. Nevertheless, if the gravity waves are truly to be believed, it is wonderful to think of them showing up with the right relative strengths at these very different scales.

4.9 Future Detectors

It is generally thought that to make the greatest advance in sensitivity, large bolometric arrays will be needed. Figure 32 shows a picture of one such detector. The bolometer is a plastic with a gold coating coupled to a thermistor. The time constant is of the order of milliseconds and this can influence the rates with which one can scan. The sensitivity can be dominated by the photon noise itself; bolometers are comparable to HEMT detectors at 100 GHz but are significantly better at higher frequencies. They are also very stable. By using a wire grid for the absorber, one minimizes the effects of cosmic rays; and by appropriate metalizations, the detectors can be made polarization sensitive.

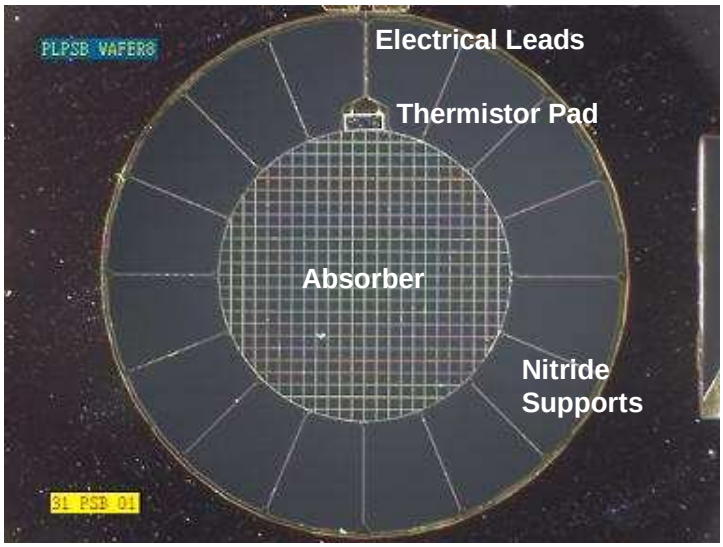

Figure 32: A Bolometric Detector for Boomerang

Figure 33 shows how two polarization sensitive bolometers are illuminated using the optics designed for the Planck mission.

Once one is at the sensitivity limit of the photon noise itself, the only way to increase sensitivity is to add more detectors. Figure 34 shows a bolometric array (for infrared radiation), the Sharc-II array. Experiments of the future are planning focal planes consisting in a few hundred to a few thousand detectors.

From space, these detectors will likely be the technology of choice. But many studies on the ground before the next satellite experiment will be able to advance the field; and from the ground, bolometric detectors may not be so superior.

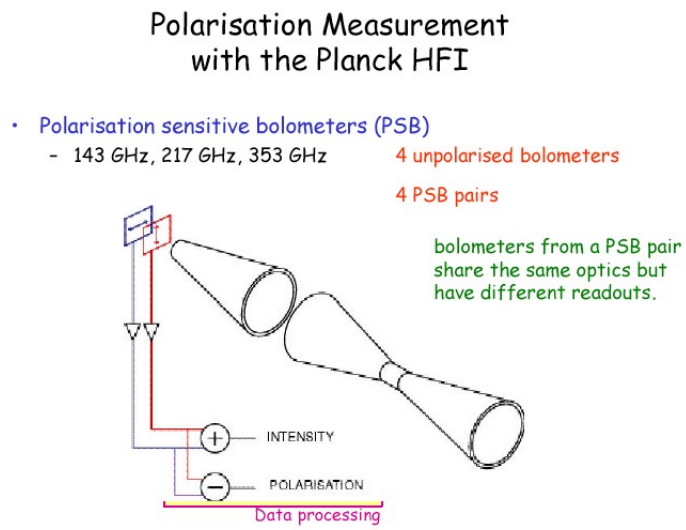

Figure 33: **Polarization Optics for the HFI Instrument on Planck.** Orthogonally oriented Bolometers detect the two powers through identical optics.

Recently there has been an advance in the mass production of coherent radiometer elements which are in fact full correlation receivers. The leader in this development is JPL. The QUIET collaboration [‡] has recently formed and is proposing to fully exploit this very attractive possibility. The QUIET group intends to move the Crawford Hill telescope, Figure 16, to the CBI station in the Atacama desert in Chile. Figure 35 shows how a focal plane of nearly 100 correlation polarimeters might be assembled. The collaboration would ultimately like to field 1000 or more element arrays.

4.10 The CMB and Particle Physics

In this section I want to briefly mention three other "applications" that the CMB has to particle physics.

The CMB is sensitive to the contribution to the matter density from neutrinos and hence to their masses. Non-zero masses change the expansion history and hence the time (or z) of decoupling and the subsequent formation of the large scale structure. The relevant scale is the temperature at decoupling: $T_{dec} \approx 0.30$ ev. Currently there is a limit from a joint analysis of the CMB and large scale

[‡]Berkeley, CalTech, Chicago, Columbia, Goddard, JPL, Miami, Princeton

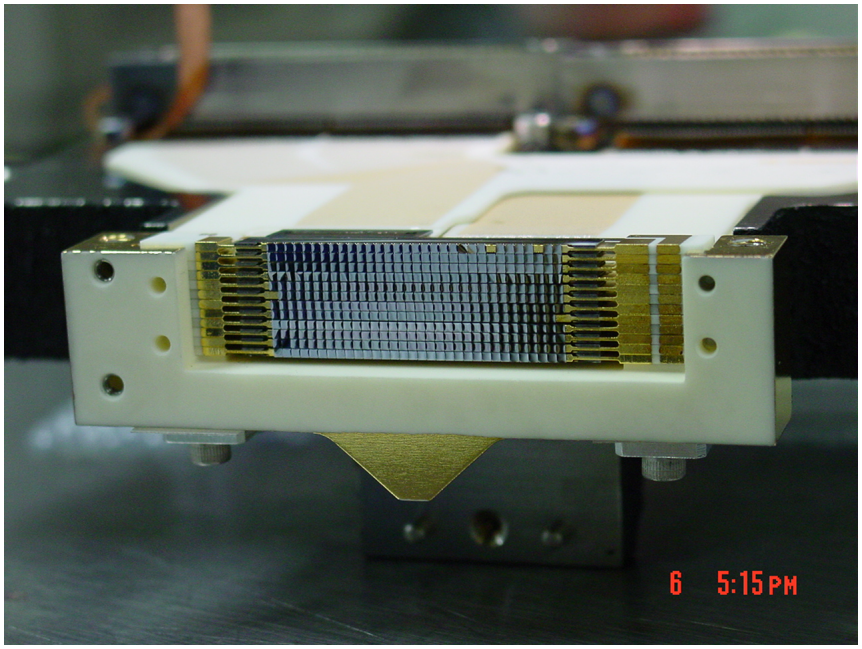

Figure 34: **The SHARC-II Bolometer Array.** This array, optimized for radiation around 400μm, efficiently fills the focal plane with 384 1mm by 1mm detectors.

structure of about 0.26 ev.

The affects of a finite mass on structure will then show up in the lensing of the CMB. It has been claimed that it is possible to detect masses in the range of 0.03 ev, the range suggested by atmosphere neutrino results.

The CMB also has implications for SUSY particles. In particular, the very precise value for Ω_{cdm} from WMAP puts significantly tighter limits on the masses of the lightest super-symmetric particle.

Finally we mention that the CMB may have sensitivity to what is called "Trans-Planckian Physics." Indeed, the modes we detect today started with wavelengths smaller than the Planck length itself! Models of such physics can be limited by precise studies of the CMB.

4.11 Final Remarks

It should be clear that the cosmic microwave background has had a seminal impact on our views of cosmology. There is the possibility that in the future we will be able to see signals from the Universe when it went through its inflationary era

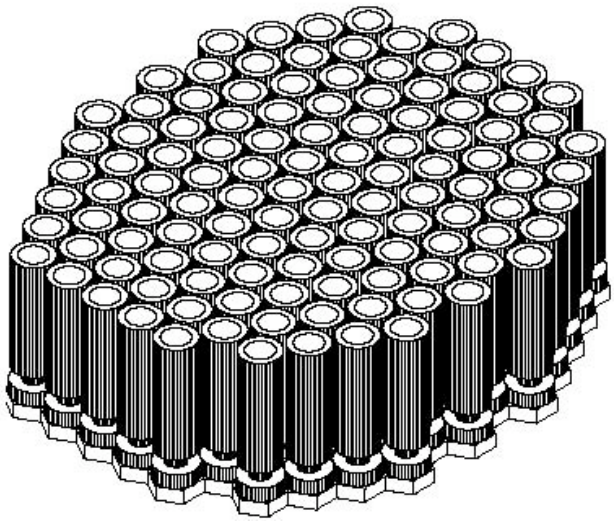

Figure 35: **The QUIET 91 Element Array.** Shown schematically are the feed horns spaced about 1.5 in apart.

(if indeed it had one). This depends critically on what the scale of inflation is and whether it has anything to do with the grand unification scale. Also, is slow-roll inflation the "true" mechanism? It seems that the search for these gravity waves have an analogy with the search for the decay of the proton. The latter, which arguably is even more significant science than a signature of inflation, has no particularly natural scale. An order of magnitude improvement might cost $500M and would be hard to justify on its own, given the uncertain payoff. That important neutrino physics can be done with the same detector will be critical in the initiation of such an effort. A satellite to look for the B-modes might cost about the same; such a mission will also need its "neutrino physics", in this case studies of the lensing of the CMB.

Currently three NASA Inflation Probe studies are underway to map out the issues which need addressing for such a mission. Many other ground based efforts are in progress so the climate is as exciting as ever.

4.12 Acknowledgements

I am indebted to several sources in the preparation of these lectures.

First off, I want to mention the very best book on the subject for beginners. It is by Bruce Partridge[1] who provides an historical perspective together with a

scholarly approach, including critical analyses of many of the early experiments in the field. In spite of its being out of date on both measurements and theory, I strongly recommend the book.

Several scientists were very giving of their time in the weeks before the lectures. These include Steve Meyer, Wayne Hu, and my colleagues in CAPMAP.

Tom Crawford deserves special mention: his knowledge, experience and willingness to be reading and discussing the same papers with which I was struggling was most helpful and much appreciated.

I want to thank the organizers of this SLAC summer school for the opportunity to try to summarize and expand my own knowledge on the subject: as they say, you never fully understand anything (or know how little you really do understand anything) until you try to teach it.

Finally I want to acknowledge Suzanne Staggs, Lyman Page, and Dave Wilkinson for welcoming me into a most stimulating environment at Princeton during 1999-2000, resulting in a career change for the author.

References

[1] Bruce Partridge, 3K: The Cosmic Microwave Background Radiation, Cambridge University Press, 1995.

[2] E. L. Wright et al., 1994, ApJ 420:450:456

[3] http://background.uchicago.edu/ whu/intermediate/intermediate.html

[4] L. Page et al., ApJ Supplement Series, 148:233-241, 2003 September

[5] Chicago, JPL, Miami, Princeton, http://cfcp.uchicago.edu/capmap/

[6] Carlton Caves, Phys. Rev. D26, 1817 (1982).

[7] Tools of Radio Astronomy, K. Rohlfs and T.L. Wilson, 3rd Edition, Springer 1999

[8] SLAC E-158, M. Woods et al.

[9] Lloyd Knox and Yong-Seon Song, PRL 89, July 2002

[10] W. Hu and T. Okamoto, Astrophys. J. 574 566 (2002)