RADIO SIGNATURES OF HI AT HIGH REDSHIFT: MAPPING THE END OF THE “DARK AGES”

PAOLO TOZZI,1,2 PIERO MADAU,1,3 AVERY MEIKSIN,4 AND MARTIN J. REES4

Received 1999 March 10; accepted 1999 August 23

ABSTRACT

The emission of 21 cm radiation from a neutral intergalactic medium (IGM) at high redshift is discussed in connection with the thermal and ionization history of the universe. The physical mechanisms that make such radiation detectable against the cosmic microwave background include Lyα coupling of the hydrogen spin temperature to the kinetic temperature of the gas and preheating of the IGM by the first generation of stars and quasars. Three different signatures are investigated in detail: (a) the fluctuations in the redshifted 21 cm emission induced by the gas density inhomogeneities that develop at early times in cold dark matter (CDM) dominated cosmologies, (b) the sharp absorption feature in the radio sky due to the rapid rise of the Lyα continuum background that marks the birth of the first UV sources in the universe, and (c) the 21 cm emission and absorption shells that are generated on several Mpc scales around the first bright quasars. Future radio observations with projected facilities like the Giant Meterwave Radio Telescope and the Square Kilometer Array may shed light on the power spectrum of density fluctuations at z > 5 and map the end of the “dark ages,” i.e., the transition from the postrecombination universe to one populated with radiation sources.

Subject headings: cosmology: theory — diffuse radiation — galaxies: evolution — quasars: general — radio lines: general

1. INTRODUCTION

The development of structure in the universe was well advanced at high redshifts. Quasars have been detected nearly to z = 5, and the most distant galaxies to even greater distances (Weymann et al. 1998). The Lyα forest in the spectra of high-redshift QSOs shows that the intergalactic medium (IGM) itself developed extensive nonlinear structures by early times. Still unknown, however, is the epoch during which the first generation of stars was formed and the timescale over which the transition from a neutral universe to one that is almost completely reionized took place. While numerical simulations have shown that the IGM is expected to fragment into structures at early times in cold dark matter (CDM) cosmogonies (Cen et al. 1994; Zhang, Anninos, & Norman 1995; Hernquist et al. 1996; Zhang et al. 1998), the same simulations are much less able to predict the efficiency with which the first gravitationally collapsed objects lit up the universe at the end of the “dark age” (Rees 1998). The nature of the first dominant ionizing sources is still unknown, and although IR observations will permit even higher redshift galaxies to be observed, detections become increasingly difficult because of the surface brightness dimming from cosmological expansion.

Madau, Meiksin, & Rees (1997, hereafter MMR) demonstrated that radiation sources prior to the epoch of reionization may be detected indirectly through their impact on the surrounding neutral IGM and the resulting emission or absorption against the cosmic microwave background (CMB) at the frequency corresponding to the redshifted 21 cm line. While its energy density is estimated to be about 2 orders of magnitude lower than the CMB, the 21 cm structural signature will display angular structure as well as structure in redshift space due to inhomogeneities in the gas density field (Hogan & Rees 1979; Scott & Rees 1990), hydrogen ionized fraction, and spin temperature (MMR). Because of the smoothness of the CMB sky, fluctuations in the 21 cm radiation will dominate the CMB fluctuations by about 2 orders of magnitude on several arcminute scales. Radio observations at meter wavelengths have, therefore, the potential to tell us how (as well as when) the primordial gas was reheated and reionized: whether the reheating was due to ultraluminous but sparsely distributed sources like QSOs or by fainter but more uniformly distributed objects like an abundant population of pregalactic stars.

In this paper we discuss possible ways in which the next generation of radio telescopes could provide the first observations of a neutral IGM at a level of a few tenths of μJy arcmin−2. The proposed Square Kilometer Array (SKA) would provide an enhancement of about 2 orders of magnitude in sensitivity relative to present-day radio telescopes (Taylor & Braun 1999). We discuss the technical requirements necessary to detect the 21 cm signature from the IGM and describe a possible design for a special-purpose telescope. We emphasize that within existing technological limits radio astronomy could effectively open up much of the universe to a direct study of the reheating and reionization epochs.

2. BASIC THEORY

The ground-state hyperfine levels of hydrogen will very quickly reach thermal equilibrium with the CMB. In order to detect the IGM in either emission or absorption, it is crucial that an effective mechanism be present to decouple the hydrogen spin states from the CMB. Following MMR, we briefly review below the mechanisms that determine the level populations of the hyperfine state of neutral hydrogen in the intergalactic medium (see also Meiksin 1999).

2.1. Spin Temperature

The emission or absorption of 21 cm radiation from a neutral IGM is governed by the hydrogen spin temperature
where \( n_0 \) and \( n_1 \) are the singlet and triplet \( n = 1 \) hyperfine levels, \( T_s \equiv h v_1 / k_B = 0.07 \) K, \( v_1 \) is the frequency of the 21 cm transition, \( h \) is Planck’s constant, and \( k_B \) is Boltzmann’s constant. In the presence of only the CMB radiation with \( T_{\text{CMB}} = 2.73(1 + z) \) K, the spin states will reach thermal equilibrium with the CMB on a timescale of \( T_s / T_{\text{CMB}} A_{10} \sim 3 \times 10^5 (1 + z)^{-1} \) yr \( (A_{10} = 2.9 \times 10^{-15} \) s\(^{-1} \) is the spontaneous decay rate of the hyperfine transition of atomic hydrogen), and intergalactic H I will produce neither an absorption nor an emission signature. A mechanism is required that decouples \( T_s \) and \( T_{\text{CMB}} \), e.g., coupling the spin temperature instead to the kinetic temperature \( T_K \) of the gas itself. Two mechanisms are available: collisions between hydrogen atoms (Purcell & Field 1956) and scattering by Lyγ photons (Wouthuysen 1952; Field 1958). The collision-induced coupling between the spin and kinetic temperatures is dominated by the spin-exchange process between the colliding hydrogen atoms. The rate, however, is too small for realistic IGM densities at the redshifts of interest, although collisions may be important in dense regions, \( \delta \rho / \rho \gtrsim 30 [(1 + z)/10]^{-2} \) (MMR).

In the low-density IGM, the dominant mechanism is the scattering of continuum UV photons redshifted by the Hubble expansion into local Lyγ photons. The many scatterings mix the hyperfine levels of neutral hydrogen in its ground state via intermediate transitions to the \( 2p \) state, the Wouthuysen-Field process. An atom initially in the \( n = 1 \) singlet state absorbs a Lyγ photon that puts it in an \( n = 2 \) state, allowing it to return to the triplet \( n = 1 \) state by a spontaneous decay. As the neutral IGM is highly opaque to resonant scattering, the shape of the continuum radiation spectrum near Lyγ will follow a Boltzmann distribution with a temperature given by the kinetic temperature of the IGM (Field 1959). In this case, the spin temperature of neutral hydrogen is a weighted mean between the matter and CMB temperatures:

\[
T_s = \frac{T_{\text{CMB}} + y_s T_K}{1 + y_s}, \tag{2}
\]

where \( y_s \equiv P_{10} T_s / A_{10} T_K \) (3)
is the Lyγ pumping efficiency. Here \( P_{10} \) is the indirect deexcitation rate of the triplet \( n = 1 \) state via absorption of a Lyγ photon to the \( n = 2 \) level, and \( T_s \gg T_K \) is assumed. If \( y_s \) is large, \( T_s \to T_K \), signifying equilibrium with the matter. A consideration of the net transition rates between the various hyperfine \( n = 1 \) and \( n = 2 \) levels above shows that the \( 1 \to 0 \) transition rate via Lyγ scattering is related to the total rate \( P_s \) by \( P_{10} = 4 P_s / 27 \) (Field 1958). This relation and equation (2) are derived in the Appendix. In the limit \( T_K \gg T_{\text{CMB}} \), the fractional deviation in a steady state of the spin temperature from the temperature of the CMB is

\[
\frac{T_s - T_{\text{CMB}}}{T_s} \approx \left( 1 + \frac{T_{\text{CMB}}}{y_s T_K} \right)^{-1}. \tag{4}
\]

There exists then a critical value of \( P_s \) which, if greatly exceeded, would drive \( T_s \to T_K \). This “thermalization” rate is (MMR)

\[
P_{th} \equiv \frac{27 A_{10} T_{\text{CMB}}}{4 T_s} \approx (7.6 \times 10^{-12} \) s\(^{-1}) \left( \frac{1 + z}{10} \right). \tag{5}
\]

### 2.2. Brightness Temperature

Consider a patch of neutral hydrogen with an angular diameter on the sky larger than the beamwidth of the telescope and a radial velocity width broader than the telescope. In an Einstein–de Sitter universe with \( H_0 = 100 \) km s\(^{-1} \) Mpc\(^{-1} \) and baryon density parameter \( \Omega_b \), the optical depth of the patch at 21\((1 + z)\) cm for a spin temperature \( T_s \neq T_{\text{CMB}} \) is

\[
\tau(z) = \frac{3}{32 \pi} \lambda_{10}^3 A_{10} \frac{T_s}{T_{\text{CMB}}} n_H(0) H_0 (1 + z)^{1.5} \approx 0.0033 \left( \frac{T_{\text{CMB}}}{T_s} \right) \left( \frac{\Omega_b h^2}{0.02} \right)^{1/2} \left( \frac{1 + z}{10} \right)^{1/2}. \tag{6}
\]

The optical depth will typically be much less than unity. Since the brightness temperature through the IGM is \( T_b = T_{\text{CMB}} e^{-\tau} + T_s (1 - e^{-\tau}) \), the differential antenna temperature observed at Earth between this region and the CMB will be

\[
\delta T_b = (1 + z)^{-1} (T_s - T_{\text{CMB}}) (1 - e^{-\tau}) \approx (9.0 \) mK) \left( \frac{\Omega_b h^2}{0.02} \right)^{1/2} \left( \frac{1 + z}{10} \right)^{1/2} \left( \frac{T_s - T_{\text{CMB}}}{T_s} \right). \tag{7}
\]

If the intergalactic gas has been significantly preheated, \( T_s \) will be much larger than \( T_{\text{CMB}} \), and the IGM will be observed in emission at a level that is independent of the exact value of \( T_s \). By contrast, when \( T_{\text{CMB}} \gg T_s \) (negligible preheating), the gas will appear in absorption, and \( | \delta T_b | \) will be a factor \( \sim T_{\text{CMB}} / T_s \) larger than in emission, so that it becomes relatively easier to detect intergalactic H I (Scott & Rees 1990).

The presence of a sufficient flux of Lyγ photons will thus render the neutral IGM “visible.” Without heating sources, the adiabatic expansion of the universe will lower the kinetic temperature of the IGM well below that of the CMB, and the IGM will be detectable through its absorption. If there are sources of radiation that preheat the IGM, it may instead be possible to detect the IGM in emission.

### 2.3. Preheating

The energetic demand for heating the IGM above the CMB temperature is meager, only \( \sim 0.004 \) eV per particle at \( z \sim 10 \). Consequently, even relatively inefficient heating mechanisms may be important warming sources well before the IGM has been reionized.\(^6\) Possible preheating sources are soft X-rays from an early generation of quasars or thermal bremsstrahlung emission produced by the ionized

\(^5\) In the presence of a radio source, the brightness temperature of the radio continuum emission should be added to \( T_{\text{CMB}} \) in eq. (2). The radio emission may make an important contribution in the vicinity of a radio loud quasar (Bahcall & Ekers 1969) and would itself permit the IGM to be detected in 21 cm radiation. This may be an important process if the emission of ionizing radiation is more tightly beamed than the radio continuum emission, allowing the gas to remain neutral near the quasar after it has switched on (MMR).

\(^6\) In an inhomogeneous universe, reheating appears to be a slow process occurring at early epochs, followed by sudden reionization at later times (Gnedin & Ostriker 1997).
gas in the collapsed halos of young galaxies (MMR). While photons emitted by a quasar just shortward of the photoelectric edge are absorbed at the ionization front, photons of much shorter wavelength will be able to propagate into the neutral IGM much farther. Most of the photoelectric heating of the IGM by a QSO is accomplished by soft X-rays. The timescale required for the radiation near the light front to heat the intergalactic gas to a temperature above that of the CMB is typically 10% of the Hubble time. The H II region produced by a QSO will therefore be preceded by a warming front.

An additional heating source is the Lyα photon scattering itself (MMR). The heating rate by the recoil of scattered Lyα photons is

$$\dot{E}_h = \frac{h_p v_a}{m_p c^2} h_p v_a P_{th},$$

where $P_{th}$ is the Lyα scattering rate per hydrogen atom, and $m_p$ is the mass of a hydrogen atom. In the case of excitation at the thermalization rate $P_{th}$, equation (8) becomes

$$\dot{E}_{th} = \frac{27 h_0^2 c^2 A_{10}}{4 m_H c^2 T_{CMB}} \approx \left(315 \text{ K Gy}^{-1} \right) \left(\frac{1+z}{10}\right).$$

The characteristic timescale for heating the medium above the CMB temperature via Lyα resonant scattering at this rate is

$$\Delta t = \frac{2m_H c^2 v_1}{9 h_p v_a} A_{10}^{-1} \approx 10^8 \text{ yr},$$

about 20% of the Hubble time at $z = 8$. This result has an important consequence: if a substantial population of sources of UV radiation turn on quite rapidly at, say, $z \sim 10$, there exists a finite interval of time during which Lyα photons couple the spin temperature to the kinetic temperature of the IGM before heating the gas above the CMB (reionization occurs at even later times). A window is then created in redshift space where a large fraction of intergalactic gas may be observed at $\sim 160 \text{ MHz}$ in absorption against the CMB. We will show next how this transient phase may imprint a signature on the radio sky.

3. RADIO SIGNATURES

3.1. The “Cosmic Web” in Emission at 21 cm

Numerical N-body/hydrodynamics simulations of structure formation within the framework of CDM-dominated cosmologies (Cen et al. 1994; Zhang et al. 1995, 1998; Hernquist et al. 1996) have recently provided a definite picture for the topology of the IGM at $z \lesssim 5$, one of an interconnected network of sheets and filaments with viralized systems located at their points of intersection. In between the filaments are underdense regions. This “cosmic web” appears to be a generic feature of CDM (Pogosyan et al. 1998), a pattern imprinted in the matter fluctuations at decoupling and sharpened by gravity. In a UV photonizing background, modestly overdense filaments ($1 < \rho_{fil}/\rho_c < 5$) will give rise to the Lyα forest seen in the spectra of high-redshift quasars. As discussed below, radio observations at meter wavelengths on several arcminute scales may probe the cosmic web at times prior to the epoch of reionization.

Because of the clumpiness of the IGM, the 21 cm optical depth of equation (6) will vary for different lines of sight. To estimate the rms level of brightness temperature fluctuations at $21(1+z)$ cm, we shall consider three different cosmological models with parameters suggested by a variety of recent observations: an open model (ODM, $\Omega_M = 0.4$, $h = 0.65$, $n = 1$), a flat model (LDM, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$, $n = 1$), and an Einstein–de Sitter tilted model (tCDM, $\Omega_M = 1$, $\Omega_\Lambda = 0$, $h = 0.5$, $n = 0.8$). In all cases, the amplitude of the power spectrum or, equivalently, the value of the rms mass fluctuation in an $8 h^{-1}$ Mpc sphere, $\sigma_8$, has been fixed in order to reproduce the observed abundance of rich galaxy clusters in the local universe (e.g., Eke et al. 1998). The tilted model has been designed to also match COBE measurements of the CMB on large scales. Our choice of an open CDM cosmology (with a normalization which is 2 rms away from the value observed by COBE) is guided by the growing evidence in favor of a low value of $\Omega_M$ from cluster studies (e.g., Carlberg et al. 1997). The $\Lambda$-dominated model matches both the cluster constraint on small scales and the COBE constraint on large. It is also able to account for most known observational constraints, ranging from globular clusters’ ages to CMB anisotropies (Ostriker & Steinhardt 1995) to recent analyses of the Type Ia supernovae Hubble diagram (Perlmutter et al. 1998; Riess et al. 1998), and just fits within the constraints imposed by gravitational lensing (Kochanek 1996; Cooray, Quashnock, & Miller 1999). In all cosmologies, the baryon density is $\Omega_B h^2 = 0.02$ (Burles & Tytler 1998). The model parameters are summarized in Table 1.

As an illustrative example, consider a scenario in which sources of Lyα photons are in sufficient abundance throughout the universe to preheat the IGM to a temperature well above that of the CMB and to couple the spin temperature to the kinetic temperature of the intergalactic gas everywhere. While the average 21 cm signal is 2 orders of magnitude lower than that of the CMB (and will be swamped by the much stronger nonthermal backgrounds that dominate the radio sky at meter wavelengths and that must be removed), its fluctuations relative to the mean of the surveyed area—induced in this example by density inhomogeneities only—will greatly exceed those of the CMB. Brightness temperature fluctuations will be present both in frequency and in angle across the sky. At high redshifts, and with a spatial resolution of $\sim 1'$, a patchwork of radio emission will be produced by regions with overdensities still in the linear regime and by voids. From equations (6) and (7), the rms temperature fluctuation relative to the mean, neglecting angular correlations in the projected IGM density, is

$$\langle \delta T_{\delta}^2 \rangle^{1/2} \approx (2.9 \text{ mK}) \ h^{-1} \ x \ \left(\frac{\Omega_B h^2}{0.02}\right) \ \left[\Omega_M (1+z)^3 + \Omega_\Lambda (1+z)^3 + \Omega_\Lambda \right]^{1/2},$$

where $\Omega_k = 1 - \Omega_M - \Omega_\Lambda$ is the curvature contribution to the present density parameter. This corresponds to an rms

| Model    | $\Omega_M$ | $\Omega_\Lambda$ | $h$  | $n$  | $\sigma_8$ |
|----------|------------|-------------------|------|------|------------|
| tCDM     | 1.0        | 0.0               | 0.5  | 0.8  | 0.55       |
| ODM      | 0.4        | 0.0               | 0.65 | 1.0  | 0.92       |
| LDM      | 0.3        | 0.7               | 0.7  | 1.0  | 1.1        |

TABLE 1

Cosmological Parameters
fluctuation in the observed flux of
\[ \langle dl^2 \rangle^{1/2} = (15 \, \mu Jy \, \text{arcmin}^{-2}) \, h^{-1} \]
\[ \times \left( \frac{\Omega_b h^2}{0.02} \frac{\sigma_o}{[\Omega_M(1 + z)^3 + \Omega_K(1 + z)^2 + \Omega_{\Lambda}]^{1/2}} \right). \]  

Here \( \sigma_o \) is the variance of the density field in a volume corresponding to a given bandwidth \( [\Delta \nu / \nu = \Delta z/(1 + z)] \) and angular size \( \Delta \theta \). The former corresponds to a comoving length \( L \approx (1 + z)h(1 / \Delta \nu) \, \text{Mpc} \), where \( h(z) \) is the Hubble constant and \( \nu = 1.4/(1 + z) \, \text{GHz} \) is the observation frequency, and the latter corresponds to a comoving radius \( R = 0.5 \Delta \theta (1 + z) / d_A(z) \), where \( d_A \) is the angular diameter distance.

On the scales of interest here, one can use the linearly evolved power spectrum \( P(k) \) (Bardeen et al. 1986) and a filter function that is a cylinder of radius \( R \) and angular size \( \theta \) to compute the density variance:

\[ \sigma_o^2 = \frac{8D^2(z)}{\pi^2 R^2 L^2} \int_0^\infty dk \int_0^1 dx \frac{\sin^2(k L x / 2 J_1(k R / \sqrt{1 - x^2}))}{x^2(1 - x^2)} \frac{P(k)}{k^2} \times (1 + f x^2)^2, \]

where \( J_1 \) is the Bessel function of order 1, \( D(z) \) is the linear growth factor (e.g., Peebles 1993), and \( f \approx \Omega_M \delta \) takes into account the effect of the peculiar velocities in compressing the same emitting volume into a narrower bandwidth (as the Hubble flow is reduced in overdense regions; see Kaiser 1987). In this regime, it is also legitimate to assume that the baryons follow the dark matter, since segregation effects are important only on smaller scales. We have verified, using N-body simulations, the validity of the linear approximation and the neglect of angular correlations on scales of \( \sim 1' \).

Figure 1 shows the fluctuation levels at 150 MHz \( (z = 8.5) \) as a function of beam size for the models listed in Table 1. At a fixed bandwidth of 1 MHz, \( \langle \delta T^2 \rangle^{1/2} \) increases with decreasing angular scale from about 1 to 10 mK as \( \sigma_o \) increases with decreasing linear scale. Since the growth of density fluctuations ceases early on in an open universe (and the power spectrum is normalized to the abundance of clusters today), the signal at a given \( \Delta \theta \) is much larger in OCDM than in tCDM at high redshift. In a \( \Lambda \)CDM universe, the level of fluctuations is comparable with OCDM. The range of density fluctuations that would be detectable at the 5 \( \sigma \) level by a projected facility like the SKA is shown in Figure 2 as a function of beam size and frequency bandwidth for several assumed integration times. The detection thresholds are scaled according to a theoretical rms continuum noise in an 80 MHz bandwidth at 150 MHz of 64 nJy over an 8 hr integration, allowing for only 21 of the 30 dishes in a compact array configuration (Taylor & Braun 1999),

\[ \text{rms} = (0.23 \, \mu Jy)(150 \, \text{MHz})^{1/2} \left( \frac{100 \, \text{hr}}{\tau} \right)^{1/2}. \]

The central compact array in the “straw man” SKA design proposal has a diameter of \( \sim 50 \, \text{km} \) in order to achieve 1”–10” resolution. A more appropriate resolution for measuring fluctuations in the IGM intensity is \( \sim 1' \). The

\[ \text{See also http://www.nfna.nl/skai/science/}. \]
compact array would then have to be scaled down to \( D_c \approx 7 \text{ km (} \Delta \theta/\text{arcmin})^{-1} \) to be optimized for such an experiment. The dashed lines show curves of constant rms fluctuation of the signal in the corresponding beam. Each angular resolution observation corresponds to a different compact array diameter \( D_c \) (in other words, the array is being optimally resized to match its resolution to the desired angular scale). Because the received flux is proportional to the solid angle of the beam, the signal decreases with decreasing angular scale until it falls below the detection threshold, indicated by the solid lines. With 100 hr of observation in a tCDM model, 2 \( \mu Jy \) fluctuations will be detectable only on scales exceeding 3'. In OCDM, and similarly in LCDM, a signal of \( \sim 2 \mu Jy \) may be detected at significantly higher resolution with the same integration time.

Figure 3 shows radio maps at \( z = 8.5 \) for tCDM and OCDM, assuming a beamwidth of 2' and a frequency window of 1 MHz around 150 MHz. Here a collisionless \( N \)-body simulation with 64\(^3 \) particles has been performed with Hydra (Couchman, Thomas, & Pearce 1995). The simulation box size is 20 \( h^{-1} \) comoving Mpc, corresponding to 17' (11') in tCDM (OCDM). The baryons are assumed to trace the dark matter distribution without any bias. From a visual inspection of the two simulated images, as well as from Figure 2, it seems possible that observations with a modified (much less spread out) SKA design may be used to reconstruct the matter density field at redshifts between the epoch probed by galaxy surveys and recombination, on scales as small as 0.5–2 \( h^{-1} \) comoving Mpc, i.e., masses in the range between \( 10^{12} \) and \( 10^{13} \) \( h^{-1} M_\odot \).

3.2. Observing the Epoch of the First Stars

In the real universe, the observed radio pattern on the sky depends on a combination of the underlying Gaussian density field and the distribution and luminosities of the sources that cause the preheating and reionization of the IGM. In this and the following sections we shall discuss the detectability of the first radiation sources in the universe through their impact on the surrounding neutral IGM.

In hierarchical clustering theories for the origin of cosmic structure, massive objects grow “bottom-up”—low-mass perturbations collapse early, then merge into progressively larger systems. In such a scenario, one naturally expects the high redshift universe to contain small-scale substructures, systems with a virial temperature well below 10\(^8 \) K. The resultant virialized systems will remain as neutral gas clouds unless they can cool because of molecular hydrogen, in which case they may form Population III stars. These stars will produce a background of UV continuum photons near the Ly\( \alpha \) frequency which will immediately escape into the IGM while at the same time ionize the gas in their vicinity. Ly\( \alpha \) photons will propagate into uncollapsed, largely neutral regions of the IGM where the kinetic temperature in the absence of preheating will be

\[
T_k(z) \approx 26 \text{ mK}(1+z)^2
\]

(Couchman 1985), well below \( T_{\text{CMB}} \) because of adiabatic cooling during cosmic expansion. If Population III sources are in sufficient abundance throughout the universe, the Ly\( \alpha \) flux will couple the spin temperature to the kinetic temperature, and \( T_k \) will be pulled below \( T_{\text{CMB}} \) everywhere. After about 10\(^8 \) yr (see eq. [10]), the same Ly\( \alpha \) photons responsible for the coupling will warm the IGM to a temperature well above that of the CMB. Eventually, a sufficient number of photons above 1 ryd may be able to penetrate all of the surrounding neutral gas in the cloud, reach the external intergalactic gas, and create expanding intergalactic H\( \Pi \) regions.

The crucial assumption we make here is that the Ly\( \alpha \) radiation field switches on at a redshift \( z > z_{\text{th}} \), following the formation of the first stars and reaches the thermalization value \( P_{\text{th}} \) at \( z_{\text{th}} \) on a timescale much shorter than the Hubble time at that epoch. While in hierarchical models of structure formation the spatial number density of halos of a given mass grows initially exponentially (Press & Schechter 1974), the rise time of the UV radiation background will be modulated by gasdynamical processes. In the model of Gnedin & Ostriker (1997), for instance, a Ly\( \alpha \) metagalactic flux is rapidly produced in the range \( 7 < z < 10 \). In this case, coupling between and is almost instantaneous at \( z = z_{\text{th}} \), and the IGM is suddenly detectable in absorption. For \( z < z_{\text{th}} \), Ly\( \alpha \) photons begin to heat the IGM, and the kinetic temperature increases at the rate

\[
\frac{dT_k}{dz} = \frac{2\mu}{3} \frac{\dot{E}_{\text{th}}}{k_B} \frac{dt}{dz} + 2 \frac{T_k}{(1+z)},
\]

until \( T_k > T_{\text{CMB}} \) and the IGM is detectable in emission. Here \( \mu = 16/13 \) is the mean molecular weight for a neutral gas with a fractional abundance by a mass of hydrogen equal to 0.75. Such a chain of events will leave a strong imprint on the CMB and mark the epoch of reheating by the first generation of stars. At 150 MHz \( (z_{\text{th}} = 8.5) \), the brightness temperature (shown in Fig. 4) shows an absorption feature of 40 mK, with a width of about 10 MHz. Such a depression in the CMB is much easier to observe compared with brightness fluctuations, as the isotropic signal may be recorded with a larger beam area. Assuming \( \Delta T \approx 40 \) mK as a fiducial estimate, the signal-to-noise ratio for the SKA is

\[
S/N \approx 9 \left( \frac{\Delta v}{1 \text{ MHz}} \right)^{1/2} \left( \frac{\tau}{100 \text{ hr}} \right)^{1/2} A_b,
\]

where \( \Delta v \) is the frequency resolution, \( A_b \) is the beam area in arcmin\(^2 \), and \( \tau \) is the observation time. While the strength of this feature depends mainly on the timescale \( t_{\text{th}} \) over which the continuum Ly\( \alpha \) background field reaches the thermalization value, we have checked that the absorption dip is well above the detection limit for all timescales \( t_{\text{th}} \leq 0.1 t_{\text{H}} \), where \( t_{\text{H}} \) is the then Hubble time. Note that similar sharp signatures in the radio background at meter wavelengths could also be produced at the reionization epoch, as recently discussed by Gnedin & Ostriker (1997) and Shaver et al. (1999).

3.3. Observing the Epoch of the First Quasars

Very little is actually known observationally about the nature of the first bound objects and the thermal state of the universe at early epochs. If ionizing sources are uniformly distributed, like an abundant population of pregalactic stars, the ionization and thermal state of the IGM will be the same everywhere at any given epoch, with the neutral fraction decreasing rapidly with cosmic time. As discussed above, this may be the case in CDM-dominated cosmologies, since bound objects are sufficiently massive (\( \sim 10^6 M_\odot \)) to make stars form at high redshift. On the other hand, reionization may also occur in a highly inhomogeneous
Fig. 3a

Fig. 3a—(a) Radio map of redshifted 21 cm emission against the CMB in a tCDM cosmology at $z = 8.5$. The linear size of the box is $20 \, h^{-1}$ (comoving) Mpc. The point-spread function of the synthesized beam is assumed to be a spherical top hat with a width of 2'. The frequency window is 1 MHz around a central frequency of 150 MHz. The color intensity goes from 1 to 6 kJy per beam. For clarity, the contour levels outline regions with signal greater than 4 $\mu$Jy per beam. (b) Same for OCDM.
fashion, as widely separated but very luminous sources of photoionizing radiation such as QSOs, present at the time the IGM is largely neutral, generate expanding HI regions on Mpc scales: the universe will be divided into an ionized II regions that are shrinking neutral phase. If the ionizing sources are random-phase, whose filling factor increases with time, and an ever-increasing HI bubble, there is an inner thin shell of neutral gas where the IGM is heated to $T_{\text{HI}} \approx 10^4$ K and Kelvin (see text), coupling with $T_k$ with $T_g$. Bottom: Corresponding evolution of the kinetic (dashed line), spin (solid line), and CMB (dotted line) temperatures.

![Image of a graph showing temperature profiles](image-url)

**Fig. 4.—** Top: Mean $\Delta T = T_k - T_{\text{CMB}}$ vs. frequency for an angular resolution of 1° and frequency resolution of 1 MHz (solid line), together with the rms fluctuation value (dashed line). The IGM is reheated at $z_{\text{eq}} = 9$. The strong absorption feature in the mean antenna temperature is associated with the fast rise of a Lyα continuum background on a timescale of $\sim 10^6$ yr (see text), coupling with $T_k$ with $T_g$. Bottom: Corresponding evolution of the kinetic (dashed line), spin (solid line), and CMB (dotted line) temperatures.

The IGM is unveiled both in emission and in absorption. The size and intensity of the detectable 21 cm region depend on the QSO luminosity and age. In Figure 5 the signal ranges from $-3$ to $3 \mu$Jy per beam.

Note that, although the quasar was placed in a high-density region in the corner of the simulation volume, the figure can equally be viewed as the emission due to heating by a beam of soft X-rays from the quasar with an opening angle of 90°. Thus, imaging the gas surrounding a quasar in 21 cm emission could provide a direct means of measuring the opening angle of quasar emission.

4. A COMPACT, SPECIAL-PURPOSE, LOW-FREQUENCY ARRAY

The requirements for observing the epoch of IGM reheating and reionization are a bandwidth of $\Delta \nu \approx 1$ MHz, to achieve adequate resolution in velocity space at the relevant redshift, and an angular resolution of a few arcminutes, the typical scale of the expected density fluctuations. The straw man proposal for the SKA suggests a large diameter for its central compact array in order to achieve a resolution of $\sim 1°$—10°. While gas density fluctuations would persist to smaller scales in the IGM, Figure 2 shows that the associated signal would drop well below the threshold for a practical detection. For a successful experiment, it is necessary to significantly decrease the diameter of the compact array.

The current design for the SKA (and the Giant Meterwave Radio Telescope [GMRT]) has too widely distributed “subapertures” for a program oriented toward detecting arcminute-scale structures.

An alternative possibility is to build a compact, special-purpose radio telescope dedicated to searching for the expected reheating and reionization signals at low frequencies. A single dish with a diameter of 200 m (smaller than Arecibo) would resolve scales of 30°, corresponding to $35-50 h^{-1}$ comoving Mpc. This is comparable to the largest scale expected for coherent primordial density fluctuations (Pogosyan et al. 1998). Figure 1 shows that the expected brightness temperature fluctuations on this scale will be in excess of 1 mK. The rms brightness temperature is given by the radiometer equation as

$$\Delta T_b = (0.33 \, \text{mK}) \left( \frac{T_{\text{sys}}}{200 \, \text{K}} \right) \left( \frac{1 \, \text{MHz}}{\Delta \nu} \right)^{1/2} \left( \frac{100 \, \text{hr}}{\tau} \right)^{1/2} \quad (18)$$

(Burke & Graham-Smith 1997) for a bandwidth $\Delta \nu$ and integration time $\tau$. The system temperature $T_{\text{sys}}$ is sky-limited to 200 K in the coldest directions through the Galaxy (Burke & Graham-Smith). If a density fluctuation filled the beam, it would be detectable at the $5 \sigma$ level in a 300 hr integration.

Since such structures tend to be filamentary, a more typical covering factor may be $\sim 0.3$, increasing the required integration time by a factor of 10. This may be offset by increasing the bandwidth. The velocity width of a $50 h^{-1}$ comoving Mpc filament at $z = 8.5$ corresponds to a frequency width at 150 MHz of about 8 MHz. A bandwidth of 5 MHz would then require an integration time of 500 hr. New spectrometers will be able to map the same area of sky in many bands at once, allowing a wide range of redshifts to be probed simultaneously.

5. CONCLUSIONS

The history of the universe between the epoch of recombination and the birth of bright sources of radiation by
Fig. 5a—21 cm emission and absorption against the CMB from the region surrounding a quasar source (lower right corner, in the center of an H II zone) revealed once the IGM is heated above the CMB by soft X-rays from the quasar. The linear size of the box is of $20 \, h^{-1}$ comoving Mpc (tCDM), the angular resolution is $2\,\text{arcmin}$, and the bandwidth is 1 MHz. The color levels range from $-3$ to $3$ $\mu Jy$ per beam. The dark contour demarcates the absorption ring. The quasar turns on at $z = 8.5$ with an ionizing photon luminosity of $10^{57}$ photons $s^{-1}$ and is observed after (a) 7 and (b) 10 Myr. The temperature of the IGM beyond the light radius is assumed to be $T_k = 26 \, \text{mK} \approx 2.3 \, \text{K} < T_{\text{CMB}}$. 

Fig. 5b
Radio Signatures of H I at High Redshift 605

$z \lesssim 5$ is currently shrouded in mystery: the absence of radiation sources leaves this period a “dark age.” Yet spectra of QSOs show not only metal absorption features, the signature of ongoing and past star formation, but evidence for nonlinear density fluctuations on large scales as well, a “cosmic web” in the form of the Lyman forest.

We have shown that it may be possible to detect both the development of the cosmic web and the epoch of the first radiation sources, whether stars or QSOs, using 21 cm measurements from intergalactic gas. Prior to reionization, the radio signal will display emission and/or absorption structures both in angle across the sky and in redshift. Fluctuations in 21 cm emission will result from nonuniformities in the distribution of Lyman radiation sources required to decouple the spin temperature from the CMB temperature. With an instrument like the Square Kilometer Array, such fluctuations would be detectable on arcminute scales, and so reveal whether the reheating is due to galaxies or QSOs.

Absorption measurements may permit the epoch of the first stellar sources to be determined. The continuum radiation from an early generation of stars, redshifted to Ly$\alpha$, will couple the spin temperature to the low kinetic temperature of the IGM for a transitory interval before heating the gas above the CMB temperature. The result is a detectable absorption signature against the CMB that would flag the first epoch of massive star formation in the universe. If reheating were due to ultraluminous but sparsely distributed sources (e.g., the first bright QSOs), the resultant patches would have a larger scale than any gravitationally induced inhomogeneities. Because of the very large scales involved, there is no difficulty in principle in achieving a sufficiently narrow bandwidth and an adequate angular resolution with the SKA or GMRT; the main limitations are the low flux levels and the radio background subtraction. Alternatively, a special-purpose telescope dedicated to scanning the sky in frequency along cold lines of sight through the Galaxy may be used to detect the epochs of reheating. Subsequent high-resolution measurements with the SKA would provide a means of identifying the nature of the first radiation sources and so shed new light on the thermal and reionization history of the universe.

Support for this work was provided by NASA through ATP grant NAG5-4236 (P. M., A. M., and P. T.) and by the Royal Society (M. J. R.). We thank the anonymous referee for clarifying the technical requirements needed for this project. The results in this paper made use of the Hydra N-body code (Couchman et al. 1995).

Appendix A

The Wouthuysen-Field Mechanism

Electric dipole radiation (Ly$\alpha$ photons) can induce a spin transition through spin-orbit coupling. The relevant quantum number is the total angular momentum $F = I + J$, where $I$ is the proton spin and $J$ is the total angular momentum of the electron, $J = S + L$. There are six hyperfine states in total that contribute to the Ly$\alpha$ transition. Only four of these, the $I = 1$ singlet $S_{1/2}$ and triplet $S_{1/2}$ states (the notation is $P_L$) and the two triplet $I = 2$ states, $P_{1/2}$ and $P_{3/2}$, contribute to the excitation of the 21 cm line by Ly$\alpha$ scattering. The selection rule $\Delta F = 0, 1$ permits the transitions $S_{1/2} \rightarrow P_{1/2}, P_{3/2}$ and $P_{1/2} \rightarrow S_{1/2}$, and so effectively $S_{1/2} \rightarrow S_{1/2}$ occurs via one of the $n = 2$ states. Denoting the occupation number of $S_{1/2}$ by $n_0$ and that of $S_{1/2}$ by $n_1$, the rate equation for $n_0$ is

$$\frac{dn_0}{dt} = A_{10} \left(1 + \frac{T_R}{T_0}\right)n_1 - 3A_{10} \frac{T_R}{T_0} n_0 + P_{10}^s n_1 - P_{01}^s n_0,$$

where the radiation intensity at the 21 cm frequency has been expressed in terms of the antenna temperature $T_R$, $I_{10} = 2k_B T_R/\lambda_{21}^2$. The ratio $T_R/T_0$ is the number of 21 cm photons per mode. Here $P_{01}^s$ and $P_{10}^s$ are the effective excitation and de-excitation rates due to Ly$\alpha$ scattering. These may be related to the total scattering rate of Ly$\alpha$ photons, $P_a$, as shown below.

It is convenient to label the $n = 2$ levels as states $2-5$, from lowest energy to highest. Then the external radiation intensity at the frequency $v_{ij}$ corresponding to the $i \rightarrow j$ transition may be expressed as the antenna temperature $T_R^i$, as above. The temperature corresponding to the energy difference itself, $h \nu_{ij}$, is expressed as $T_{ij} = h \nu_{ij}/k_B$. If $A_{ij}$ denotes the spontaneous decay rate for the transition $j \rightarrow i$, then $P_{01}^s$ and $P_{10}^s$ are given by

$$P_{01}^s = 3 \frac{T_R^3}{T_0^3} A_{30}A_{31} + 3 \frac{T_R^4}{T_0^4} A_{40}A_{41} + A_{10}A_{11},$$

$$P_{10}^s = \frac{T_R^2}{T_1^2} A_{20}A_{21} + \frac{T_R^3}{T_1^3} A_{30}A_{31} + \frac{T_R^4}{T_1^4} A_{40}A_{41}.$$

The ratios $A_{ij}/A_s$, where $A_s$ is the total Ly$\alpha$ spontaneous decay rate, may be solved using a sum rule for the transitions. This states that the sum of all transitions of given $nFJ$ to all the $n'J'$ levels (summed over $F$ and $m_F$) for a given $n'J'$ is proportional to $2F + 1$ (e.g., Bethe & Salpeter 1957). There are four sets of downward transitions to $n' = 1, J' = 1/2$, corresponding to the decay intensities $I_{51}, I_{50}(=0), I_{41}, I_{40}, I_{31}, I_{30}, I_{21},$ and $I_{20}(=0)$. These give

$$I_{51} \over I_{40} + I_{41} = 3, \quad I_{40} + I_{41} \over I_{30} + I_{31} = 1, \quad I_{30} + I_{31} \over I_{21} = 3.$$
Similarly, there are four sets of upward transitions to \( n' = 2, J' = 1/2 \) or \( 3/2 \), giving the intensities \( I_{30}, I_{20}(=0), I_{31}, I_{21}, I_{50}(=0), I_{40}, I_{51}, \) and \( I_{41} \). These ratios are

\[ \frac{I_{40}}{I_{41} + I_{51}} = \frac{1}{3}, \quad \frac{I_{30}}{I_{21} + I_{31}} = \frac{1}{3}. \]  

(A5)

Using \( I_{kj} / I_a = (g_k / g_a)(A_{kj} / A_a) \), where \( I_a \) is the total Ly\( \alpha \) decay intensity summed over all the hyperfine transitions, and \( g_k = 2J_k + 1 \) is the statistical weight of level \( k \) \((g_{tot} \text{ is the sum of the weights of the upper levels})\), gives

\[ A_{25}/A_a = A_{50}/A_a = 0, \quad A_{21}/A_a = A_{51}/A_a = 1, \quad A_{40}/A_a = A_{41}/A_a = 1/3, \quad \text{and} \quad A_{31}/A_a = A_{40}/A_a = 2/3. \]

We then obtain

\[ P_{10} / P_{01} = n_1 / n_0 = 3 \exp (-T_\alpha / T_K). \]

Solving the rate equation (A1) for \( T_R \) \( > 0 \) gives equation (2), where \( T_R = T_{\text{CMB}} \), and where \( T_5 \gg T_\alpha, T_K \gg T_\alpha \), and \( T_R \gg T_\alpha \) have been assumed.

REFERENCES

Bahcall, J. N., & Ekers, R. D. 1969, ApJ, 157, 1055
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Bethe, H. A., & Salpeter, E. E. 1957, Quantum Mechanics of One- and Two-Electron Atoms (New York: Academic)
Burke, B. F., & Graham-Smith, F. 1997, An Introduction to Radio Astronomy (Cambridge: Cambridge Univ. Press)
Burles, S., & Tytler, D. 1998, ApJ, 507, 732
Carlberg, R. G., Morris, S. L., Yee, H. K. C., & Ellingson, E. 1997, ApJ, 479, 119
Cen, R., Miralda-Escude, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9
Cooray, A. R., Quashnock, J. M., & Miller, M. C. 1999, ApJ, 511, 562
Couchman, H. M. P. 1985, MNRAS, 214, 137
Couchman, H. M. P., Thomas, P. A., & Pearce, F. R. 1995, ApJ, 452, 797
Eke, V. R., Cole, S., Frenk, C. S., & Henry, J. P. 1998, MNRAS, 298, 1145
Field, G. B. 1958, Proc. IRE, 46, 240
———. 1959, ApJ, 129, 51
Gnedin, N. Y., & Ostriker, J. P. 1997, ApJ, 486, 581
Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escude, J. 1996, ApJ, 457, L51
Hogan, C. J., & Rees, M. J. 1979, MNRAS, 188, 791
Kaiser, N. 1987, MNRAS, 227, 1
Kochanek, C. S. 1996, ApJ, 466, 638
Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429 (MMR)
Meiksin, A. 1999, in Perspectives in Radio Astronomy: Scientific Imperatives at cm and m Wavelengths, ed. M. P. van Haarlem & J. M. van der Hulst (Dwingeloo: NFRA), in press
Ostriker, J. P., & Steinhardt, P. J. 1995, Nature, 377, 600
Peebles, P. J. E. 1993, Principles of Physical Cosmology (Princeton: Princeton Univ. Press)
Perlmutter, S., et al. 1998, Nature, 391, 51
Pogosyan, D., Bond, J. R., Kolman, I., & Wadsley, J. 1998, in Wide Field Surveys in Cosmology (Paris: Editions Frontieres), 61
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Purcell, E. M., & Field, G. B. 1956, ApJ, 124, 542
Rees, M. J. 1998, in The Next Generation Space Telescope: Science Drivers and Technological Challenges (Noordwijk: ESA), 5
Riess, A., et al. 1998, AJ, 116, 1009
Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510
Shaver, P. A., Windhorst, R. A., Madau, P., & de Bruyn, A. G. 1999, A&A, 345, 380
Taylor, A. R., & Braun, R. 1999, Science with the Square Kilometer Array
Weymann, R. J., Stern, D., Bunker, A., Spinrad, H., Chaffee, F. H., Thompson, R. I., & Storrie-Lombardi, L. J. 1998, ApJ, 505, L95
Wouthuysen, S. A. 1952, AJ, 57, 31
Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57
Zhang, Y., Meiksin, A., Anninos, P., & Norman, M. L. 1998, ApJ, 495, 63