

REIONIZATION OF THE UNIVERSE AND THE EARLY PRODUCTION OF METALS

NICKOLAY Y. GNEDIN¹ AND JEREMIAH P. OSTRIKER²

Received 1996 December 13; accepted 1997 April 11

ABSTRACT

We simulate a plausible cosmological model in considerable physical and numerical detail through the successive phases of reheating (at $10 \lesssim z \lesssim 20$) and reionization at $z \approx 7$. We assume an efficiency of high-mass star formation appropriate to leave the universe, after it becomes transparent, with an ionizing background $J_{21} \approx 0.4$ (at $z = 4$), near (and perhaps slightly below) the observed value. Since the same stars produce the ionizing radiation and the first generation of heavy elements, a mean metallicity of $\langle Z/Z_{\odot} \rangle \sim 1/200$ is produced in this early phase, but there is a large variation about this mean, with the high density regions having $Z/Z_{\odot} \approx 1/30$ and the low density regions (or the Ly α forest with $N_{\text{HI}} \lesssim 10^{13.5} \text{ cm}^{-2}$) having essentially no metals.

When it occurs, reionization is very rapid (phase change-like), which will leave a signature that may be detectable by very large area meter-wavelength radio instruments. Also, the background UV radiation field will show a sharp drop of $\sim 10^{-3}$ from 1 to 4 ryd because of absorption edges.

The simulated volume is too small to form L_* galaxies, but the smaller objects that are found in the simulation obey the Faber-Jackson relation.

In order to explore theoretically this domain of “the end of the dark ages” quantitatively, numerical simulations must have a mass resolution of the order of $10^{4.5} M_{\odot}$ in baryons, have high spatial resolution ($\lesssim 1$ kpc) to resolve strong clumping, and allow for detailed and accurate treatment of both the radiation field and atomic/molecular physics.

Subject headings: cosmology: theory — early universe — galaxies: formation — galaxies: intergalactic medium — large-scale structure of universe

1. INTRODUCTION

In a previous paper (Ostriker & Gnedin 1996) we showed, for a specific cosmological model, how an early generation of stars would be expected to form from the first nonlinear self-gravitating clumps, which developed while the universe was still relatively cold and was primarily filled with neutral gas. This first generation, “Population III,” which would reside in widely and fairly smoothly distributed clumps of stars somewhat more massive than globular clusters, will reheat and reionize the universe, leading to an end of what Martin Rees has called the “dark ages,” during the interval $15 > z > 7$. Since the same stars produce both ionizing photons during their main sequence phase and metals when they end their lives as Type II supernovae, this early phase of the universe, the “clearing of the fog,” also and inevitably leaves an irregular contamination with metals at the low (average) level of $\langle Z \rangle / Z_{\odot} \sim 10^{-2.5}$.

In the current paper, we provide the physical basis for the conclusions reached in Ostriker & Gnedin (1996), and also address a number of related questions, the following being among them.

1. When, in terms of the Gunn-Peterson effect (Gunn & Peterson 1965), should the universe have become transparent? And how closely are the processes labeled “reheating” and “reionization” related?

2. How would one characterize the spatial variations in the metal abundance produced by the early generation of stars?

3. What was the degree of clumpiness in the gas during these early phases? (This is a question that was typically neglected in the previous treatments.)

4. Will specially designed large area telescopes looking at the (rest frame) 21 cm radio sky be able to detect this early development of structure?

5. How important were the various cooling processes (atomic, molecular, and dust), the various heating processes (gravitational collapse, UV photons, and supernovae), and radiative shielding of high density regions (by gas or dust)?

6. What were the properties of the stellar groups formed at high redshifts? Are they observable?

These and other questions are addressed in the context of a specific cosmological model. In a final section, we speculate on how robust our conclusions are to variations of the assumed model.

In discussing various physical effects that play a role at the end of the dark ages, we must note the many authors who have studied these processes before, from the pioneering work of Couchman & Rees (1986) and Shapiro & Giroux (1987) to the recent careful investigations by Shapiro, Giroux, & Babul (1994), Tegmark & Silk (1994, 1995), Kawasaki & Fukugita (1994), Shapiro (1995), Giroux & Shapiro (1996), Tegmark et al. (1997) to the one-dimensional simulations by Haiman, Thoul, & Loeb (1995) and Haiman & Loeb (1996). While varying amounts of physically detailed modeling were included in all of those papers, there was a critical factor that could not be followed in any of the semianalytic treatments: the clumping of the gas. However, *all* the relevant processes are dependent on the clumping factor $C_{bb} \equiv \langle \rho^2 \rangle / \langle \rho \rangle^2$. Included among these processes are recombination, cooling, gravitational collapse, molecular hydrogen formation, and numerous others. Whereas one-dimensional simulations are able to

¹ Astronomy Department, University of California, Berkeley, Berkeley, CA 94720; gnedin@astron.berkeley.edu.

² Princeton University Observatory, Peyton Hall, Princeton, NJ 08544; jpo@astro.princeton.edu.

TABLE 1
NUMERICAL PARAMETERS

Run	N	Box Size (Mpc)	Total Mass Resolution ($h^{-1} M_{\odot}$)	Spatial Resolution (h^{-1} kpc)	Dynamic Range
A.....	128^3	2	3.7×10^5	1.0	2000
B.....	64^3	2	2.9×10^6	3.0	640
C.....	64^3	1	3.7×10^5	1.5	640

account for the clumping around a single, spherically symmetric object, they must neglect the clumping within the flow, and therefore they cannot predict the average clumping of the gas in the universe and are not able to follow the evolution of the radiation field in any reasonable detail. Thus, even the qualitative conclusions of the previous papers may be in doubt until they are confirmed by a fully nonlinear, three-dimensional treatment.

The first attempts to investigate the formation of the first generation of collapsing cosmological objects in a realistic three-dimensional cosmological simulation were done by the NCSA/Urbana group (Anninos et al. 1996; Zhang et al. 1996; see also Abel 1995). Those simulations were designed to study the very onset of baryonic collapse of a single object in the early universe; they included neither star formation nor radiative transfer and therefore were not able to predict the evolution of average properties of the universe (such as the clumping of the gas, the ionization fraction, and the ionizing background) in a globally averaged fashion.

A legitimate question to ask at this point is, Does the present simulation reported on in this paper have sufficient spatial resolution to correctly compute the clumping? The answer is either yes or barely enough, depending on one's degree of optimism. On scales below the Jeans mass, clumping should be relatively unimportant as it cannot (by definition of the Jeans mass) be assisted by gravity. At a characteristic temperature of 10,000 K, the Jeans mass is around $10^{10} M_{\odot}$ at $z = 4$ for the average density of the universe, and is around $10^8 M_{\odot}$ for the overdensity of 10^4 (which is reliably resolved by our simulations), whereas our nominal total mass resolution is $10^{5.5} M_{\odot}$ (see Table 1). Numerical experiments that we have performed indicate that the resolution needed to permit the existence of a first (Population III) generation of stars is of the order of $10^{5.5} M_{\odot}$, which we just achieve in the computations reported on here. Work that is currently underway should allow us to push the limit to a comoving spatial scale of $\approx 0.5 h^{-1}$ kpc and a mass scale of $1.0 \times 10^5 M_{\odot}$, while simultaneously better allowing for the missing large-scale power.

2. METHOD

2.1. Physical Ingredients of a Numerical Simulation

In order to adequately simulate the evolution of the universe at high redshift to at least moderate accuracy, we must allow for all of the most important physical processes. Fortunately, by restricting ourselves to the early stages of the formation of structure, we can incorporate essentially all physics that affect formation of structure at moderate densities ($n \lesssim 1 \text{ cm}^{-3}$). This is in contrast to simulations of galaxy and large-scale structure formation at low redshifts, which are generically missing important physical ingredients. The main reason behind this difference between simulating a high- and a low-redshift universe is that most

of the poorly understood complicated physical processes (i.e., effects of heavy elements, dust, and magnetic fields) are thought to be less important at high redshift, where most of the intergalactic gas is of essentially primeval composition and magnetic fields have not yet had time to build up to a significant strength (see Kulsrud et al. 1996).

We have therefore undertaken to simulate numerically the evolution of the intergalactic medium at high redshift, including essentially all physical processes that we identified as important for the high-redshift evolution of the universe. We have used the SLH-P³M cosmological hydrodynamic code as described by Gnedin (1996) and Gnedin & Bertschinger (1996) to follow the evolution of the dark matter and the cosmological gas with high-Lagrangian resolution. We have also included detailed atomic and molecular physics of a gas of primeval composition, following in non-equilibrium, time-dependent fashion the ionization and recombination of hydrogen and helium as well as the formation and destruction of hydrogen molecules in the ambient radiation field. The radiation field, in turn, allows for sources of radiation (quasars and massive stars), sinks (due to continuum opacities), and cosmological effects. In particular, we have taken into account the fact that dense lumps will be shielded from the background radiation field. This reduces the heating rates for dense clumps and makes it nearly certain that once they have formed and started to collapse, the process will be irreversible. They will continue to collapse and fragment even though outer, lower density regions may reheat and be stabilized against further collapse. However, since the precise calculation of the shielding effect requires performing the whole radiative transfer in the highly inhomogeneous medium, which is well beyond the capabilities of existing computers, we adopted a simple approximation that we call the local optical depth approximation. In its essentials, this approximation treats absorption as localized but treats the sources of radiation as smoothly distributed. We derive radiative transfer equations in the expanding universe in Appendix A, and then introduce the local optical depth approximation in Appendix B. In Appendix C, we describe in detail our method of following formation and destruction of molecular hydrogen.

The ionizing radiation is emitted by stars and quasars. However, it is impossible to simulate star formation within the cosmological framework directly, and we therefore are forced to use a phenomenological approach to account for star formation and feedback processes. In regions that are cooling and collapsing, we have allowed the formation of pointlike "stellar" subunits according to the Cen & Ostriker (1992) algorithm as adapted to the SLH code by Gnedin (1996). However, we introduce the following change to the Cen & Ostriker algorithm. In their original paper, Cen & Ostriker introduced, as a numerical convenience, the density cutoff as a value below which no formation of stellar subunits was permitted. The value of the cutoff is arbitrary, and final results are somewhat dependent on its value (Gnedin 1996; Katz, Weinberg, & Hernquist 1996). In our improved version of the star formation algorithm, we abandon the density cutoff value in the belief that as long as the code has not reached its resolution limit, the evolution of the gas is followed accurately and there is no need to introduce this phenomenology altogether. Therefore, we allow for star formation only in the regions that have reached the resolution limit since, in those regions, the

further collapse of the gas cannot be followed accurately and stellar particles *must* be created in order to allow the gas to sink beyond the resolution limit of a simulation. We identify this limit as the limit when the Lagrangian behavior of the SLH code switches to the Eulerian one. Numerically, this corresponds to places where the maximum eigenvalue of the softening tensor σ_{ij} reaches 1/2 (see Gnedin & Bertschinger 1996 for the definition of the softening tensor σ_{ij}), and physically, it corresponds to the spatial resolution shown in Table 1 for the various runs we have performed.

As soon as a stellar particle forms, it releases radiation and (in proportion) metal-rich gas, which we have considered in the treatment of cooling. Since we cannot distinguish between stars and quasars in our treatment of star formation, we allow each stellar particle to emit a composite spectrum of radiation that consists of a mixture of both the quasar and young star spectra. The shape of the spectrum is shown in Figure 1 of Gnedin (1996). The amount of radiation emitted is, again, a free parameter of the phenomenological star formation algorithm, and we choose it so that we produce ionizing radiation at low redshift in the amount that is close to the observed value.³ We assume an efficiency $\varepsilon_{UV,S} = 6 \times 10^{-5}$ for the stellar component of the source spectrum and the equal value of $\varepsilon_{UV,Q} = 6 \times 10^{-5}$ for the QSO-like component of the source spectrum. These numbers are essentially our only free parameters. The sum of them is fixed by matching the background radiation field at $z = 4$ and the ratio determined by the softness of the ionizing background (see Fig. 11), which is also an observable.

We have also allowed for the injection of the thermal energy from supernova explosions into the intergalactic gas according to the methodology of Cen & Ostriker (1992). However, at the resolution we achieve in the simulations reported here, the gas density is so high that all this thermal energy is immediately radiated away, which produces a negligible net effect on the evolution of the universe. This result stems purely from the inadequate treatment of the supernova explosions, since gas inside the supernova bubbles is too hot to cool. Therefore, the appropriate treatment of the supernova explosions would require following hydrodynamics of the multiphase medium (Cowie, McKee, & Ostriker 1981) and modeling interactions between the phases, which is beyond the frame of this work. We therefore emphasize that while supernova explosions are formally incorporated in our simulations, they are not followed adequately and, therefore, their effects may be significantly underestimated in our treatment.

The only important piece of physics that we are totally missing in our simulations is the nonuniformity of the sources of radiation (as explained in Appendix B). One should therefore bear this in mind when interpreting the results of our simulations. But, since we compute all average quantities as appropriate, we expect that our simulations adequately reproduce the evolution of the universe as a whole, and, in particular, the evolution of average quantities.

2.2. Cosmological Models

We have adopted a cold dark matter (CDM) + Λ cosmological model as a framework for our investigations.

³ However, since we cannot determine the final value for the ionizing intensity until after we have run the simulation, we can only approximately reproduce the observational value for the ionizing intensity.

However, we believe that qualitative results of our simulation are applicable to the whole family of the CDM-type models, and we emphasize those qualitative conclusions together with specific quantitative results for the model we use. We fix the cosmological parameters as

$$\Omega_0 = 0.35, \quad \Omega_L = 0.65, \quad h = 0.70, \quad \Omega_b = 0.03,$$

which are close to those used in the ‘‘concordance’’ model of Ostriker & Steinhardt (1995). We normalize the model to the COBE 10° measurement and allow for a small tilt with $n = 0.96$. This gives the value of $\sigma_8 = 0.67$ for the power spectrum normalization at $8 h^{-1}$ Mpc and $z = 0$. In order to compute the initial conditions accurately, we computed the linear transfer functions for the model using the linear Boltzmann code that is similar to (but different from) the COSMICS package (Bertschinger 1995). We show, in Figure 1, the transfer functions for both the dark matter (*solid line*) and the baryons (*dotted line*) in ratio to the BBKS transfer functions (Bardeen et al. 1986) at $z = 100$, where we start our simulations.

We have performed three runs with different box sizes and numerical resolutions to assess the importance of different scales and estimate the uncertainty introduced by the finite resolution of our simulations. The adopted parameters of those runs are compiled in Table 1. All runs were stopped at $z = 4$, since at that time the rms density fluctuation at $2 h^{-1}$ Mpc scales is 0.4 and the absence of nonlinear wavelengths larger than the box size renders the simulations senseless.

The largest of our simulations, run A, is our fiducial run against which to compare the results of the two smaller runs. However, it happened by accident that the particular realization of initial conditions we used for run A had substantially less power on large scales ($1\text{--}2 h^{-1}$ Mpc) than the true power spectrum, and we have to take this fact into account when interpreting our results. In particular, this affects the precise value for the redshift of reionization and the evolution at late times, but at earlier times ($z \gtrsim 10$),

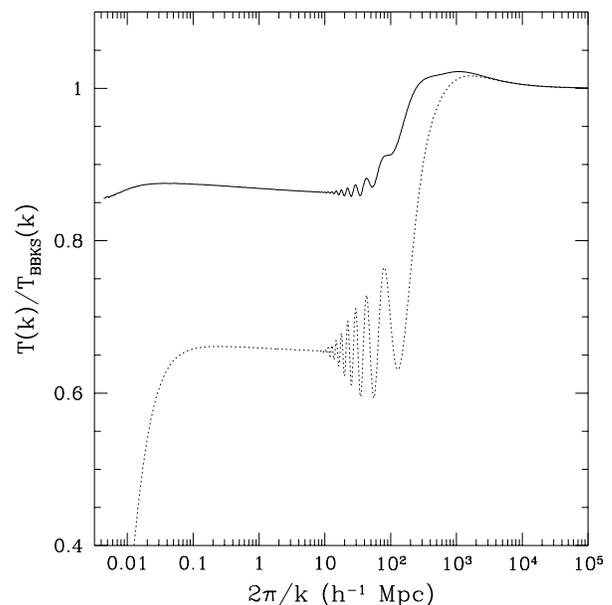


FIG. 1.—Transfer functions for the dark matter (*solid line*) and the baryons (*dotted line*) at $z = 100$ in ratio to the BBKS transfer function for the CDM + Λ cosmological model with $\Omega_0 = 0.35$, $h = 0.7$, and $\Omega_b = 0.03$.

when the density correlation length is much smaller than the box size, the absence of large waves has little effect.

3. RESULTS

3.1. Reheating and Reionization

It is often assumed in the current literature that reheating and reionization of the universe at high redshift are two sides of the same physical process, and the terms “reheating” and “reionization” are even sometimes used interchangeably. In this section we demonstrate that while the physical cause of both of those processes is the same—i.e., the high energy radiation field—they are nevertheless two different physical processes, separated in time as well as proceeding at different rates.

Since it is the stars that produce the ionizing radiation, we first show in Figure 2 the fraction f_* of the total baryonic mass that is locked in stars (*upper panel*) and the comoving star formation rate $d\rho_*/dt$ in units of $h^3 M_\odot \text{ yr}^{-1} \text{ Mpc}^3$, as a function of redshift (*middle panel*) from our run B. We also show in the middle panel (*dotted curve*) the comoving feedback rate, i.e., the rate of production of high-energy photons and metals (for convenience, we convert it to the same units). The feedback effects are delayed after the formation of a stellar particle according to the methodology of Cen & Ostriker (1992; see Gnedin 1996 for details). Note the existence of the first peak of Population III star formation at $z \approx 16$ in otherwise smooth star formation rate. The slow-down of the star formation at $z < 6$ is caused by the lack of large-scale waves in our simulation box.

Figure 3 shows the temporal evolution for the gas temperature (*lower panel*) and the neutral hydrogen fraction (*middle panel*). We also show the evolution of the ionizing radiation intensity as measured by the quantity J_{21} , which

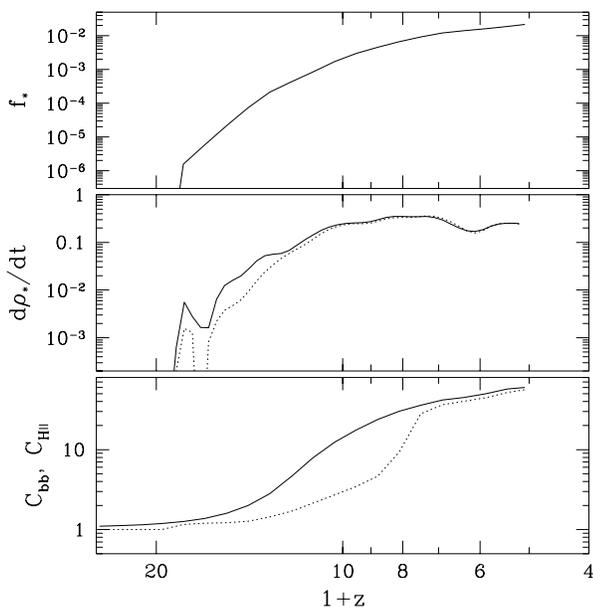


FIG. 2.—*Upper panel*: the fraction f_* of the total baryonic mass locked into stars as a function of redshift for run B. *Middle panel*: the comoving star formation rate in units of $M_\odot \text{ yr}^{-1} (h^{-1} \text{ Mpc}^{-3})$ (solid line) and the comoving stellar feedback rate, normalized to the same units (dotted line), as a function of redshift for the same run. *Lower panel*: the gas clumping $C_{bb} \equiv \langle \rho^2 \rangle / \langle \rho \rangle^2$ (solid line) and the ionized hydrogen clumping C_{HII} (discussed in § 3.1; dotted line) as a function for redshift for the same run.

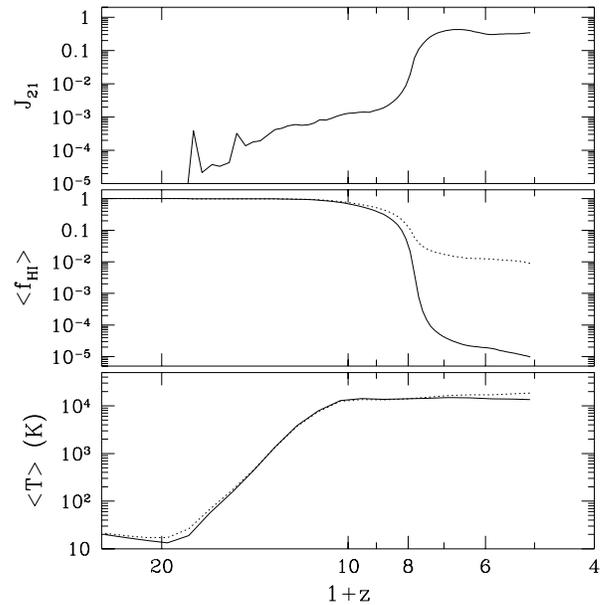


FIG. 3.—Evolution of the radiation intensity J_{21} (*upper panel*), average neutral hydrogen fraction f_{HI} (*middle panel*), and the average temperature (*lower panel*) as a function of redshift for run B. Solid lines show the volume average and dotted lines show the mass average for the temperature and neutral hydrogen fraction.

is called the ionizing intensity and defined as

$$J_{21} \equiv \frac{\int J_\nu \sigma_\nu d\nu/\nu}{\int \sigma_\nu d\nu/\nu} \times \frac{1}{10^{-21} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1}}, \quad (1)$$

where J_ν is the radiation intensity and σ_ν is the hydrogen photoionization cross section. The observed values for J_{21} lie between 0.3 and 1.5 (Savaglio et al. 1996; Cooke, Espey, & Carswell 1996) at $z \sim 3-4$.

We immediately note that there is a relatively long epoch from $z = 20$ to $z = 10$, encompassing many Hubble times at that redshift, when the volume and mass-weighted average temperature of the universe increases from the value $T = 11 \text{ K}$ at $z = 20$ (dictated by the pure adiabatic expansion of the universe) to the value $T = 14,000 \text{ K}$ (fixed by the atomic physics). This increase in temperature, or *reheating* of the universe, is accompanied by a slow increase in the value of ionizing intensity up to $J_{21} \sim 10^{-3}$, but the universe stays essentially neutral; the neutral hydrogen fraction decreases slowly until $z = 7$, when suddenly the neutral fraction drops by several orders of magnitude at essentially constant temperature and the ionizing intensity J_{21} increases to a value of around 0.5–1.0. This moment, which happens in a small fraction of a unit redshift (i.e., in a small fraction of the then Hubble time), is called *reionization* of the universe and occurs almost as a phase change.

We conclude that reheating and reionization are two different physical processes; reheating is a slow process happening at high redshift, $z > 10$, followed by a sudden reionization at a lower redshift. During reheating, the value of the ionizing intensity stays below $J_{21} < 10^{-3}$ because the universe remains mainly neutral at that time and the absorption is large. In order to understand why the absorption of ionizing photons is significant, let us estimate the recombination time, i.e., the time a hydrogen ion needs to recombine after being ionized by a high-energy photon. For the temperature $T = 10^4 \text{ K}$, the value of the recombination coefficient R is $4.2 \times 10^{-13} \text{ cm}^3 \text{ s}^{-1}$, assuming case A

recombination. The equation for the ionized hydrogen number density can be written as

$$\dot{n}_{\text{H II}} = -3Hn_{\text{H II}} + n_{\text{H I}}\Gamma - Rn_e n_{\text{H II}}, \quad (2)$$

where H is the Hubble constant, $\Gamma \equiv 4.3 \times 10^{-12} \text{ J}_{21} \text{ s}^{-1}$ is the photoionization rate, and n_e is the free electron number density. Equation (2) holds at every fluid element in the universe. In order to compute the average number density of the ionized hydrogen, we have to average equation (2) over the volume of the universe, obtaining

$$\dot{\bar{n}}_{\text{H II}} = -3H\bar{n}_{\text{H II}} + \langle n_{\text{H I}}\Gamma \rangle - \langle Rn_e n_{\text{H II}} \rangle, \quad (3)$$

where the overbar denotes the volume average. Assuming that R is constant in space and that helium is not ionized (i.e., $n_e = n_{\text{H II}}$), we can obtain the following expression for the recombination time:

$$t_R \equiv \frac{\bar{n}_{\text{H II}}}{R\langle n_{\text{H II}}^2 \rangle}; \quad (4)$$

or for the ratio of the recombination time to the Hubble time, $t_{\text{H}} = 1/H$ and

$$\frac{t_R}{t_{\text{H}}} = \frac{H}{\bar{n}_{\text{H II}} R C_{\text{H II}}} = 67 \frac{a^{3/2} h \sqrt{\Omega_0}}{x C_{\text{H II}}}, \quad (5)$$

where a is the scale factor, x is the volume average ionization fraction, $x \equiv \bar{n}_{\text{H II}}/\bar{n}_{\text{H}}$, and $C_{\text{H II}}$ is the ionized hydrogen clumping factor $C_{\text{H II}} \equiv \langle n_{\text{H II}}^2 \rangle / \langle n_{\text{H II}} \rangle^2$ (Fig. 2, *lower panel*, *dotted line*), which is somewhat lower than the total gas clumping factor $C_{\text{bb}} \equiv \langle \rho^2 \rangle / \langle \rho \rangle^2$ because higher density regions are typically less ionized. For the model under consideration, $h(\Omega_0)^{1/2} = 0.7(0.35)^{1/2} = 0.41$, and for the fully ionized gas, $x = 1$, with no clumping; the recombination time is equal to the Hubble time at the redshift $z_{\text{eq}} = 8$, which implies that if the universe stayed uniform, once it is ionized after $z = 8$ it will not be able to recombine ($t_R/t_{\text{H}} > 1$). This is a well-known result. However, since the universe is in fact strongly clumped, i.e., the clumping factor $C_{\text{H II}} \gg 1$, the absorption is much higher than that for the uniform universe. In particular, the clumping factor $C_{\text{H II}} = 12$ at $z = 8$ (and it is even larger in the real universe, since we underestimate clumping because of the finite resolution of our simulations) and the recombination time is equal to the Hubble time at $z = 8$ for the ionization fraction as low as $x = 1/C_{\text{H II}} = 0.08$. This means that it is sufficient to have only 8% of all hydrogen ionized in order to be able to absorb at least one ionizing photon per Hubble time per hydrogen atom (the corresponding number is 10% for the recombination case B).

It is also easy to understand why the reheating phase should precede the reionization phase. Let us consider the ratio of the photoheating time

$$t_{\text{heat}} \equiv \frac{3/2 k_{\text{B}} T}{E_J \Gamma}$$

to the photoionization time

$$t_{\text{ion}} \equiv \frac{1}{\Gamma},$$

where k_{B} is the Boltzmann constant and

$$E_J \equiv \frac{\int_{h\nu_0}^{\infty} (h\nu - h\nu_0) \sigma_{\nu} J_{\nu} d\nu / \nu}{\int_{h\nu_0}^{\infty} \sigma_{\nu} J_{\nu} d\nu / \nu}$$

is the average energy a hydrogen atom receives in act of ionization. For our case, $E_J = k_{\text{B}} 9.3 \times 10^4 \text{ K}$, and $T = 10^4 \text{ K}$ at $z = 10$ and

$$\frac{t_{\text{heat}}}{t_{\text{ion}}} = \frac{3}{2} \frac{k_{\text{B}} T}{E_J} = 0.16, \quad (6)$$

which implies that the radiation background is much more efficient (by a factor of 6) in heating the gas at high redshift than in ionizing it. The reason is simply that, unless the ionizing radiation field is very soft, excess heat is liberated for each act of photoionization.

Finally, we can understand why the reionization is so rapid. Since the recombination time is short compared to the Hubble time, there is an approximate balance between the emission of ionizing photons from the sources and the absorption (by ionizing a hydrogen atom which subsequently recombines). If the emission increases, the ionization fraction x increases, the recombination time decreases, and the absorption increases; inversely, the decrease in the emission leads to the subsequent decrease in the ionization fraction and an increase in the recombination time. Since the star formation rate is increasing with time (*middle panel*, Fig. 2), driven by the further collapse of hierarchical structure, the ionization fraction slowly increases in response to the increase in the emission of ionizing photons until it reaches values close to unity. At this point, a further increase in the recombination rate cannot be achieved, emission and absorption get out of balance, and the universe ionizes at the timescale of the emission, which at the point of loss of balance is approximately equal to the recombination time. We therefore conclude that the sharp decrease in the neutral hydrogen fraction from the values of ~ 0.01 at the beginning of (complete) reionization to final values of the order of 10^{-5} is achieved on a timescale of

$$t_{\text{rei}} \sim t_R = \frac{t_{\text{H}}}{C_{\text{H II}}} \left(\frac{1 + z_{\text{rei}}}{9} \right)^{-3/2},$$

which for our case of $z_{\text{rei}} = 7$ has a value of

$$t_{\text{rei}} \sim \frac{t_{\text{H}}}{10}.$$

This value is in excellent agreement with the actual reionization time measured from Figure 3. This effect of fast reionization is entirely caused by the high value of the clumping factor. Were the clumping ignored in our consideration, we would conclude that the reionization time is more than 10 times larger than the actual value is and that reionization is a slow process continuing for more than one Hubble time (as long as the source evolution is also slow as in Fig. 2). This is precisely the conclusion achieved in some of the previous work on reionization where the clumping was ignored (see, for example, Haiman & Loeb 1996).

We point out here that the specifics of the cosmological model do not enter this conclusion and the prediction that slow reheating happens first, followed later by sudden reionization, is a generic prediction for any cosmological model (with UV photons providing both heat and ionization), not necessarily a CDM-type one.

3.2. The Gunn-Peterson Effect in the Intergalactic Medium

The complete reionization of the universe manifests itself in the absence of the Gunn-Peterson absorption trough

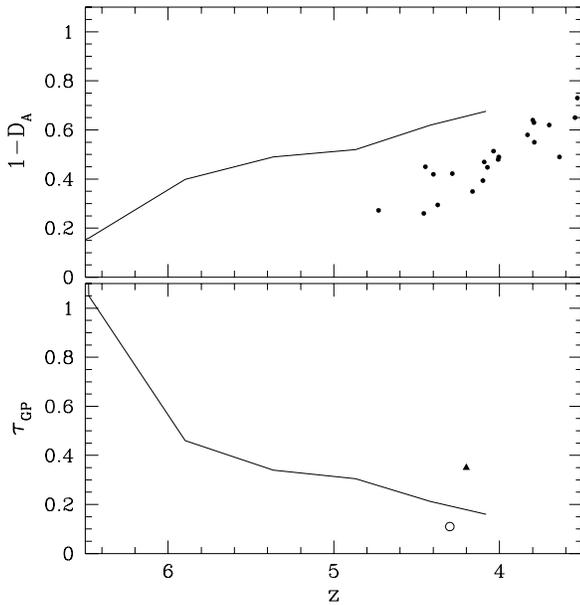


FIG. 4.—*Upper panel*: continuum suppression factors $1 - D_A$ as a function of redshift for run B (*solid line*) together with the data from Jenkins & Ostriker (1991) (*filled circles*). *Lower panel*: the Gunn-Peterson optical depth as a function of redshift for run B (*solid line*). Shown as a solid triangle is the observational upper limit from Jenkins & Ostriker (1991) and as an open circle is the upper limit from Giallongo et al. (1994).

(Gunn & Peterson 1965) in the quasar spectra.⁴ We can use the observational limits on the Gunn-Peterson optical depth as a test for cosmological models. In order to compute the Gunn-Peterson optical depth, we generate random lines of sight through our simulation box⁵ and compute the Ly α absorption $\tau(\lambda)$ along each line of sight. We then compute the average opacity of the intergalactic medium,

$$\exp(-\bar{\tau}) \equiv \frac{1}{N_{\text{ray}}} \sum_{j=1}^{N_{\text{ray}}} \frac{1}{\lambda_2^{(j)} - \lambda_1^{(j)}} \int_{\lambda_1^{(j)}}^{\lambda_2^{(j)}} \exp[-\tau(\lambda)] d\lambda, \quad (7)$$

where $N_{\text{ray}} \sim 1000$ is the number of lines of sight, and $\lambda_1^{(j)}$ and $\lambda_2^{(j)}$ are wavelength limits for the j line of sight. This computation can be compared with observations by Jenkins & Ostriker (1991) when expressed as $1 - D_A \equiv \exp(-\bar{\tau})$.

We show in Figure 4 (*upper panel*) the data from Jenkins & Ostriker (1991) as filled circles together with our computation (*solid line*). We note that our simulation predicts somewhat too little absorption compared to the Jenkins & Ostriker data, but the general behavior is quite similar to the observations. We reiterate here that we do not expect to get an exact match with all existing observational data on high-redshift gas contents of the universe, since we simulate only one cosmological model that may not be correct and our simulations are subject to uncertainties owing to the phenomenological description of star formation.

We can also express the average absorption along the line of sight as the average Gunn-Peterson optical depth.

⁴ Residual clumps (filaments) of neutral gas remain and will produce the Ly α forest.

⁵ Those lines of sight are taken to be random with respect to the box orientation since at lower redshifts, when the nonlinear scale approaches the box size, there appear purely artificial correlations between the density distribution and the orientation of the computational box.

However, we must exclude the Ly α forest from the total absorption since it is not counted toward the total Gunn-Peterson optical depth (Giallongo et al. 1994). We then compute the average Gunn-Peterson absorption τ_{GP} as the amount of absorption in the parts of the spectrum that lie above the average opacity,

$$\exp(-\tau_{\text{GP}}) \equiv \frac{1}{N_{\text{ray}}} \sum_{j=1}^{N_{\text{ray}}} \frac{\int_{\lambda_1^{(j)}}^{\lambda_2^{(j)}} \exp[-\tau(\lambda)] \theta[\bar{\tau} - \tau(\lambda)] d\lambda}{\int_{\lambda_1^{(j)}}^{\lambda_2^{(j)}} \theta[\bar{\tau} - \tau(\lambda)] d\lambda}, \quad (8)$$

where $\theta(x)$ is a step function.⁶

We show in Figure 4 (*lower panel*) the Gunn-Peterson optical depth as a function of redshift from our run B. We also show observational upper limits from Jenkins & Ostriker (1991) (*triangle*) and from Giallongo et al. (1994) (*circle*). We again note that we predict too little absorption compared to the Jenkins & Ostriker data but too much absorption compared to the Giallongo et al. data. Since the Giallongo et al. (1994) limit is the most stringent one, we adopt it for our analysis. We note that the result of our simulations, while broadly consistent with observations, goes above the observational upper limit. This implies that the real universe is more ionized than predicted by our simulations. The final value for J_{21} is 0.34 at $z = 4$, which is slightly below the currently favored observational value, as we have noted above. The more realistic value of ~ 0.7 – 1.0 would decrease the neutral hydrogen fraction and give a lower Gunn-Peterson optical depth. However, we cannot rescale the ionizing intensity a posteriori and preserve the accuracy of our simulations since the radiation field is tied up in the complex nonlinear evolution together with the intergalactic medium⁷, and we would have to rerun the whole simulation with new values for the radiative efficiencies for our stellar particles, as explained in § 2.1.

Given both the observational and theoretical uncertainties, the agreement between the predicted and observed opacity of the intergalactic medium seems quite adequate. It will be interesting to see if this rough agreement survives when other cosmological models are considered.

3.3. Metal Enrichment of the Intergalactic Medium

Since, in our simulations, we follow the star formation with the phenomenological approach, we also account for the dynamics of gas that was processed in stars and is therefore enriched with heavy elements (*metals*). However, since the initial mass function (IMF) at high redshift is not known and may deviate significantly from the standard IMF, we have no way of assigning the value for the metallicity in the processed gas reliably, but we still can study the distribution of this metal-enriched gas as compared to the distribution of the primeval gas.

In Figure 5 we plot the fraction f_Z of the metal-enriched gas in the total intergalactic gas as a function of redshift for our fiducial run A. The three different curves correspond to

⁶ This method of computing the Gunn-Peterson optical depth is similar to the Miralda-Escude et al. (1996) method for identifying Ly α lines.

⁷ In particular, even if most of the low-density intergalactic medium is in ionization equilibrium with the radiation field, the value of the temperature at every fluid element depends on the previous evolution of the density and the radiation field in this fluid element (Miralda-Escude & Rees 1994); since we explicitly allow for the nonuniform radiation background, this dependence cannot be parametrized in a simple form.

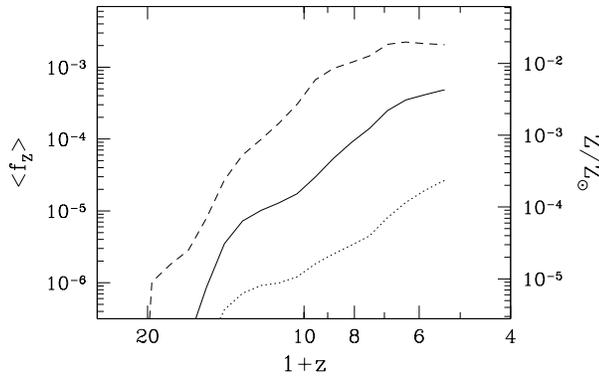


FIG. 5.—The mass weighted (*solid line*) and the volume weighted (*dotted line*) fraction of the intergalactic gas that was processed in stars as a function of redshift. The *dashed line* shows $\rho^{5/3}$ weighting, which roughly corresponds to the weighting by the column density of a Ly α absorber. The right-hand scale corresponds to the normalization with the adopted value for the yield $y = 0.020$.

different weights: the mass-weighted value is shown with the solid line, the volume-weighted value is shown with the dotted line, and the dashed line corresponds to the weighting with the $\rho^{5/3}$ weight, which roughly reflects the weighting by the column density of a Ly α absorber. Here f_z is the ratio of the mass of the gas that was processed in stars⁸ to the total mass of the intergalactic gas. If the “yield” for metal production with standard definition is y , then the “metallicity,” or a fraction of mass in elements heavier than helium, is $9yf_z$, given our definitions. Thus, for $y = 0.020$, which is characteristic for protoellipticals as measured by the metallicity of the intracluster gas (see Silk 1996), the final value of $\langle f_z \rangle \approx 5 \times 10^{-4}$ corresponds to $\bar{Z} \approx 9 \times 10^{-5}$ or $\bar{Z}/Z_\odot \approx 4.3 \times 10^{-3}$. For a more standard value for the yield ($y = 0.005$), we would obtain the final metallicity of the order of $\bar{Z}/Z_\odot \approx 1.0 \times 10^{-3}$. We also note that by $z = 4$, the volume average fraction of the metal-enriched gas may reach a value of about 0.05%.

While Figure 5 shows a rough consistency with observations, it does not show us where this metal-enriched gas is actually located, and it understates the variability in heavy element abundances at a given epoch. In Figure 6 we show the probability distribution to find a fluid element with the particular values of the fraction of the metal-enriched gas and the total density of the intergalactic gas, as measured by the overdensity $\delta \equiv \rho/\bar{\rho} - 1$. The upper panel shows the mass-weighted distributions, and the lower panel shows the volume-weighted distributions. We note that most of the volume has f_z well below 0.1%, whereas a considerable fraction of mass has f_z of the order of 0.1–1.0%. We also note that at low overdensities, which are characteristic of the low column density Ly α forest (see Miralda-Escude et al. 1996; Hui, Gnedin, & Zhang 1996), the range of existing metallicities is extremely broad, while the average value is quite low. This is in a qualitative agreement with the conclusion by Rauch, Haehnelt, & Steinmetz (1996) who found that observed scatter in the abundance ratios of the QSO metal absorption systems exceeds that predicted by simulations with the uniform metallicity.

⁸ We assume here that average stars lose 10% of their total mass into the intergalactic medium; this amount does not include the mass loss that is confined in the interstellar medium. If this amount is different from 10%, the plotted f_z should be adjusted simply in proportion to the total fraction of stellar mass lost to the intergalactic medium.

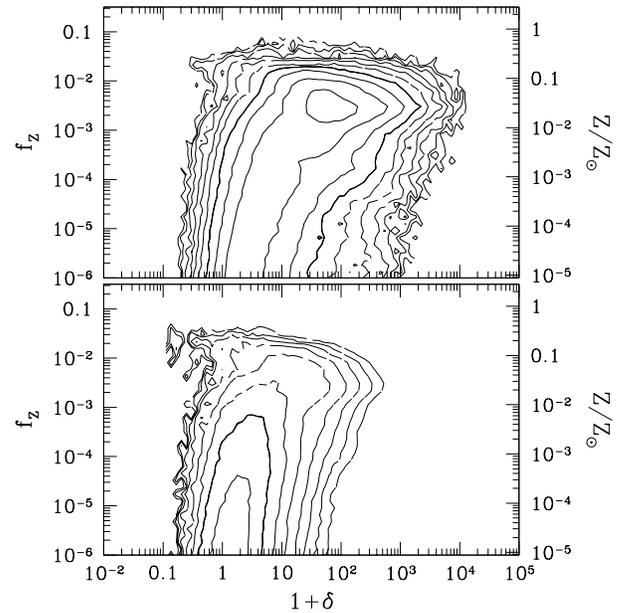


FIG. 6.—The mass-weighted (*upper panel*) and volume-weighted (*lower panel*) probability distribution to find a fluid element with given values of the fraction of metal-enriched gas and the total gas density at $z = 4$. The right-hand scale corresponds to the normalization with the adopted value for the yield $y = 0.020$. The bold line shows the average probability, solid contours mark the probability above average, and the dashed contours show the probability below average; contour spacing is logarithmic with the increment of one-third.

To quantify this effect further, we plot in Figure 7 the average fraction of the metal-enriched gas as a function of the total gas density at four different redshifts. In other words, Figure 7 is Figure 6 collapsed along the y -direction. It is interesting to note that the final values of the metallicity in the high-density regions from which galaxies will be made approach $Z/Z_\odot \approx 1/30$, which is close to the minimum metallicity found in the oldest and most metal-poor Population II component of our Galaxy. More than that, this value does not change significantly with time after $z \sim 8$. We also point out that there seems to be steady diffusion of heavy elements into low-density regions.

It is remarkable that the average metallicity is only a weak function for wide range of overdensities above

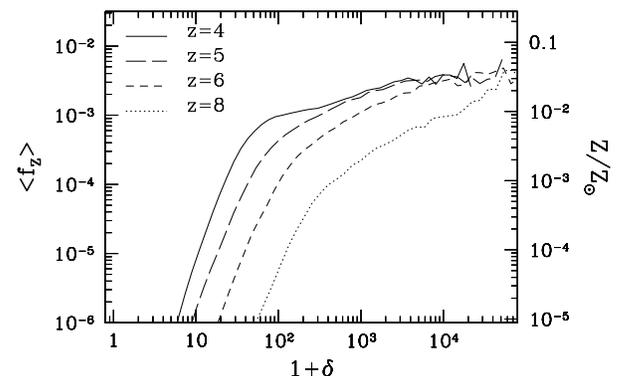


FIG. 7.—The average fraction of the metal-enriched gas as a function of total gas density at four different redshifts $z = 4, 5, 6,$ and 8 . The right-hand scale corresponds to the normalization with the adopted value for the yield $y = 0.020$.

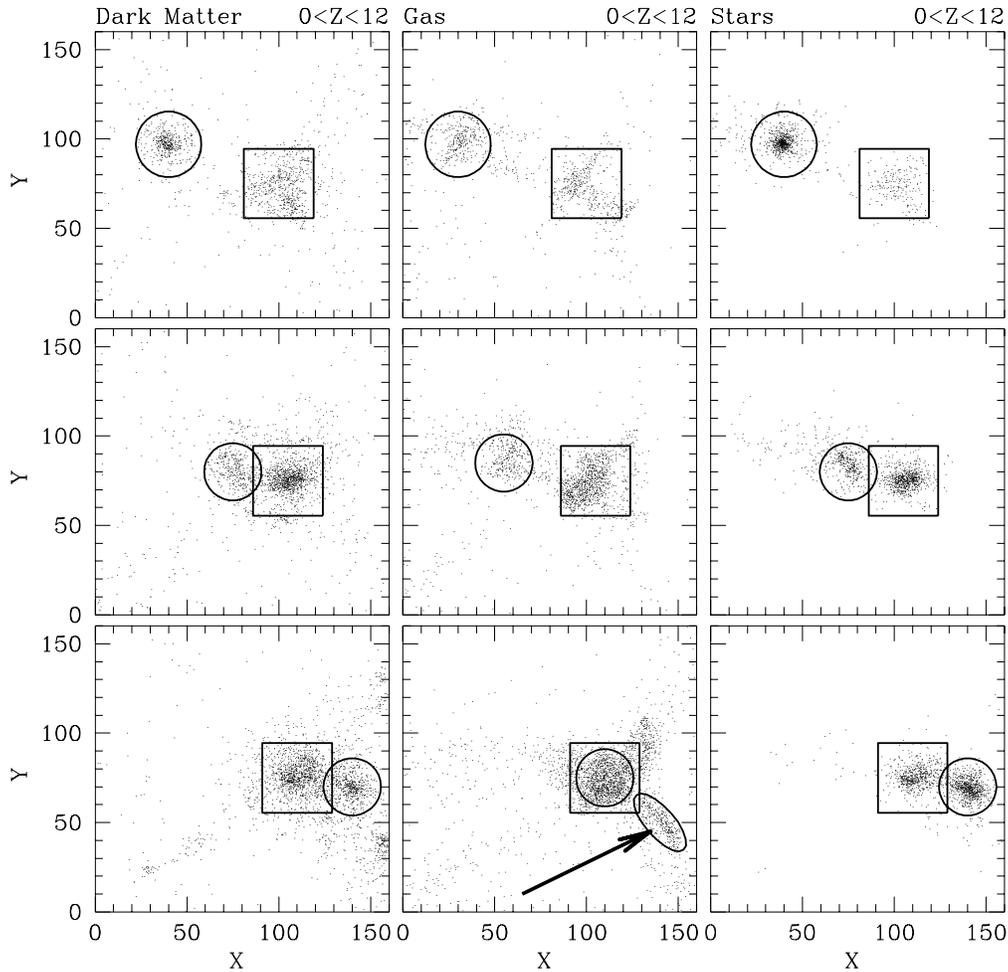


FIG. 8.—Distributions of the dark matter (*left column*), the cosmic gas (*middle column*), and the stars (*right column*) at three different redshifts: $z = 5.4$ (*upper row*), $z = 4.9$ (*middle row*), and $z = 4.3$ (*lower row*) represented by particles in a thin slice with the width of $12 h^{-1}$ comoving kpc. The square and the circle mark two merging objects, and the oval pointed to by an arrow marks ejection of the metal-enriched gas from the merger product.

$\delta \sim 10^2$, and it decreases very fast as the overdensities become smaller. This effect allows us to make a specific prediction that the metallicity in the Ly α forest should decrease sharply for the column densities below about $10^{13.5} - 10^{14.5} \text{ cm}^{-2}$ as measured at $z = 3$ (which roughly corresponds to $\delta \sim 30$; see Hui et al. 1996). Here we allow ourselves a broad error bar in the quoted value of the characteristic column density to account for the possible range of cosmological models and uncertainties in the simulations, both numerical uncertainties (such as the limitations of the phenomenological star formation algorithm) and uncertainties in the input physics (such as our lack of knowledge of the IMF at high redshift or the precise value of the cosmic baryon density). Recent measurements of metal abundances in the Ly α forest (Songaila & Cowie 1996) found heavy elements in a large fraction ($\geq 75\%$ of all Ly α systems with column densities in excess of $10^{14.5} \text{ cm}^{-2}$). Our prediction then implies that when those studies are extended to lower column densities, the sharp decline in the fraction of Ly α systems with measurable heavy element abundances will be found at column densities of about $10^{13.5} - 10^{14.0} \text{ cm}^{-2}$. The specific value of the column density at which the metal abundance of the Ly α forest declines can serve as a new test for cosmological models.

The heavy elements are produced in the regions of star formation, where the gas density is high and the potential

well is deep, and, therefore, one can expect that they are effectively trapped inside the (proto)galaxies, contrary to our finding that they are widely distributed even in the low-density regions. We must therefore address the question of how heavy elements eventually get into the low-density regions. Since our simulations do not include any special diffusion processes that may be responsible for distributing the metal-enriched gas over a large range of densities, we conclude that there must exist a mechanism that works entirely within the framework of hierarchical clustering. Figure 8 demonstrates how this may have happened. We show the distribution of the dark matter (*left column*), the cosmic gas (*middle column*), and the stars (*right column*) at three different redshifts ($z = 5.4$ [*upper row*], $z = 4.9$ [*middle row*], and $z = 4.3$ [*lower row*]) represented by particles⁹ in a thin slice with the width of $12 h^{-1}$ comoving kpc.¹⁰ All distances in the slice are in h^{-1} comoving kpc. There are two separate objects at $z = 5.4$ that are marked

⁹ While dark matter and stars are numerically represented as particles in the SLH approach, the gas component is numerically represented by a distorted quasi-Lagrangian mesh; we convert (for display only) cells of this mesh into gas particles for the sake of simplicity and uniform presentation.

¹⁰ While dark matter particles all have the same mass, the gas particles and the stellar particles have different masses, and the visual appearance of the particle density does not necessarily correspond to the actual gas or stellar density in the slice.

with the square and the circle that are undergoing merger at lower redshift.¹¹ In the lower row ($z = 4.3$), the dark matter and the star components of the smaller object (*circle*) passed close to the center of the larger object, while the gas components of both objects collided in a gigantic shock and merged at the first impact; as a result, the stream of high-density gas was ejected from the merged object (*lower middle panel, oval and arrow*). The ejected gas came mostly from the larger object and was heavily enriched by heavy elements. In the next Hubble time or so, it will be dispersed and mixed with the surrounding primeval intergalactic gas, producing high metal abundance even in low-density regions lying close to the merged object. But the low-density gas far from any merger would still be uncontaminated. Thus, the metal enrichment of the intergalactic gas will be highly inhomogeneous, as shown in Figure 5.

We also note that the fact that we demonstrated a working mechanism for the heavy element enrichment does not, of course, exclude the possibility that there exist other mechanisms that distribute metals in the low density intergalactic gas. In particular, supernova explosions, which are included in the work of Cen & Ostriker (1992), are effective at dispersing metal-enriched gas.

3.4. The 21 cm Emission from High Redshift

The neutral gas at high redshift is surrounded by the thermal bath of cosmic microwave background (CMB) photons. Depending on the spin temperature of neutral hydrogen atoms, the 21 cm line radiation can be seen either in emission or in absorption. At a sufficiently late epoch ($z = 15$ in our simulation), the kinetic temperature of the matter rises above ~ 150 K, and after that moment the 21 cm line radiation becomes visible in emission (Scott & Rees 1990). This radiation, which is being constantly redshifted during the evolution of the universe, produces a continuum signal at the telescope. As the universe evolves and structure develops, velocity focusing in the converging flows will produce enhanced emission at lower redshifts until reionization, when the neutral hydrogen fraction drops by several orders of magnitude and 21 cm emission declines sharply at the frequency corresponding to the epoch of reionization.

We show in Figure 9 the spectrum of the redshifted 21 cm line as would be observed on Earth (*solid line*) at frequencies where this radiation could be seen in emission, together with the contribution from the cosmic microwave background, reduced by a factor of 100 to fit on the same scale (*dotted line*). The very different spectral signature of the 21 cm radiation makes it possible, at least in principle, to distinguish it from the CMB. However, the amplitude of the signal is still substantially lower than the sensitivity of all existing or proposed radio telescopes (including the Square Kilometer Array). Nevertheless, when (and if) this measurement is performed, the reward will be quite substantial: since reionization is so sudden, and, therefore, the break in the spectrum at the frequency $\nu_b = \nu_0/(1+z_{\text{reion}})$ is very sharp, the measurement of the break frequency will give a very accurate value for the redshift of reionization, which in

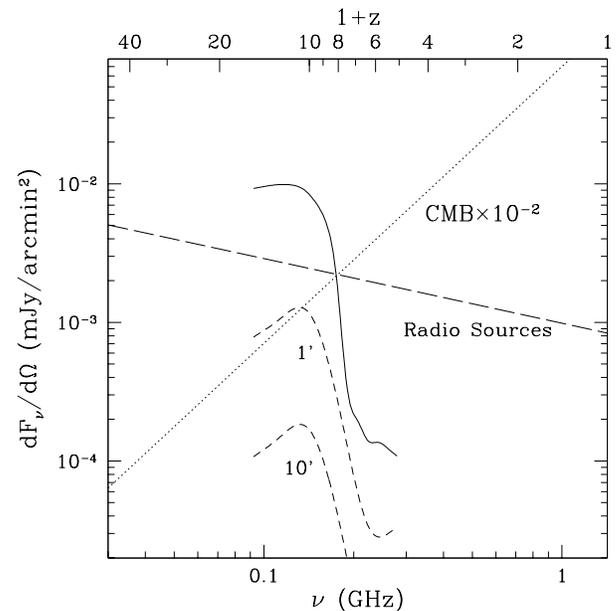


FIG. 9.—The spectrum of the 21 cm emission from the intergalactic gas at high redshift (*solid line*). The sharp break in the spectrum corresponds to the redshift of reionization. The two dashed lines show the expected rms fluctuations in this radiation on the $1'$ (*upper curve*) and $10'$ (*lower curve*) angular scale. Only the spectrum at frequencies above 92 MHz ($z < 15$) is shown where the 21 cm line is seen in emission. Shown with the dotted line is the cosmic microwave background radiation intensity reduced by a factor of 100 to fit onto the same scale.

turn can place a strong constraint on both cosmological models and physical parameters at high redshift.

We also show (*short-dashed line*) the rms fluctuations in the 21 cm radiation on $1'$ and $10'$ angular scales. Since $1'$ corresponds to the comoving scale of around $1.7 h^{-1}$ Mpc at $z = 7$ in the adopted cosmological model, the density fluctuations on this scale are on the order of 20%; in the sky, those fluctuations will result in the fluctuations in the redshifted 21 cm radiation. Even if the isotropic component of the 21 cm radiation is only 1% of the CMB fluctuations, because of the incredible smoothness of the CMB sky, $\Delta T/T \sim 10^{-5}$, fluctuations in the 21 cm will dominate the CMB fluctuations by a factor of $10^5 \times 10^{-2} \times 0.2 = 200!$ Since the density fluctuations are smaller at lower frequencies (earlier times), there is a peak in the spectrum of rms fluctuations in the 21 cm radiation close to the frequency corresponding to the 21 cm frequency redshifted to z_{reion} . This may provide an alternative (and, likely, much more efficient) way to measure the redshift of reionization from the 21 cm from high-redshift gas. However, the possibility of measuring these fluctuations will crucially depend on the ability to remove foregrounds, the dominant one of which is confusion noise from discrete radio sources. The long-dashed line shows a level of fluctuations on the $1'$ scale expected from extragalactic radio sources (as extrapolated from Franceschini et al. 1989). Unless this contamination is removed (by using differences in spectral signatures), genuine fluctuations in the 21 cm radiation will be unobservable.

The spectrum in Figure 9 shows the 21 cm emission as measured by a broad filter with 200 km s^{-1} width. However, since after reionization most of the neutral hydrogen is in optically thick lumps (protogalaxies), the final structure of the 21 cm emission at lower redshift would consist of a

¹¹ The objects may have different appearance at different redshifts owing to the thinness of the slice: particles composing the objects may enter and leave the slice at different redshifts. We show only a subset of all particles in each object for the sake of clarity.

combination of several lines rather than of a continuum emission. Figure 10 serves to illustrate this conclusion. We show the 21 cm emission as a function of frequency measured in velocity units at three different redshifts, each centered at the frequency corresponding to the redshifted 21 cm line. The emission terminates at velocities larger than 200 km s⁻¹ because of the final size of the simulation box. We note that while at high redshift, emission is almost uniform in frequency, at lower redshift, after reionization, it consists of several overlapping lines.

3.5. Evolution of the Background Radiation

The radiation emitted by starlike and quasarlike sources participates in a complex interaction with the highly inhomogeneous intergalactic medium. The interplay between ionization and recombination processes in the intergalactic gas coupled with star formation and development of cosmic structure determine the shape and the amplitude of the background radiation, which are not guaranteed to be monotonic in frequency or time. As an illustration, we show in Figure 11 the spectra of the background radiation at three different redshifts. This is qualitatively very similar to Figure 8 of Ostriker & Cen (1996) despite the considerably more detailed treatment of radiative processes in this paper. We note that in the 13.6 eV < hν ≤ 1 keV range, the background radiation spectrum has a complex shape with pronounced ionization edges at 13.6, 24.6, and 54.4 eV. The neutral helium ionization edge at 24.6 eV becomes weak after helium becomes at least singly ionized somewhat prior to hydrogen reionization at z = 7, but the hydrogen ionization edge at 13.6 eV is still easily observed even at z = 4, well after reionization of the universe. Examination of Figure 11 shows that the assumption of a power law (or other simple form) for the background radiation field, which is an approximation adopted by several authors, would have been extremely inaccurate owing to the strong effects of atomic ionization edges.

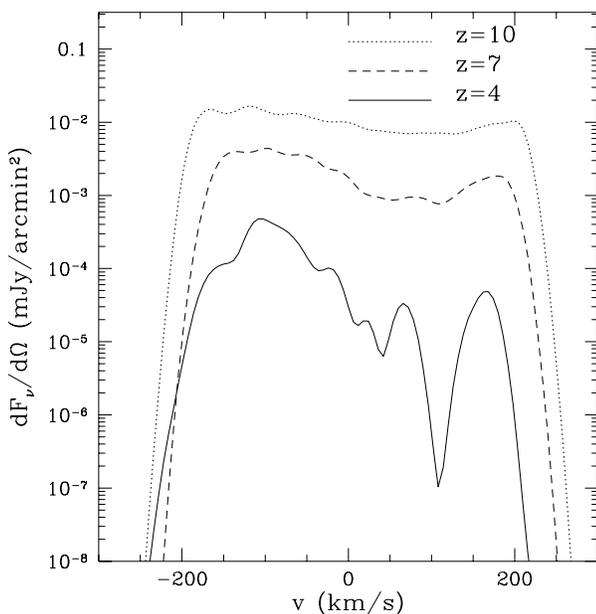


FIG. 10.—The fine structure of the 21 cm emission at three different redshifts. Each spectrum is centered at the redshifted frequency of the 21 cm. The emission terminates at high velocity values due to the final size of the simulation box.

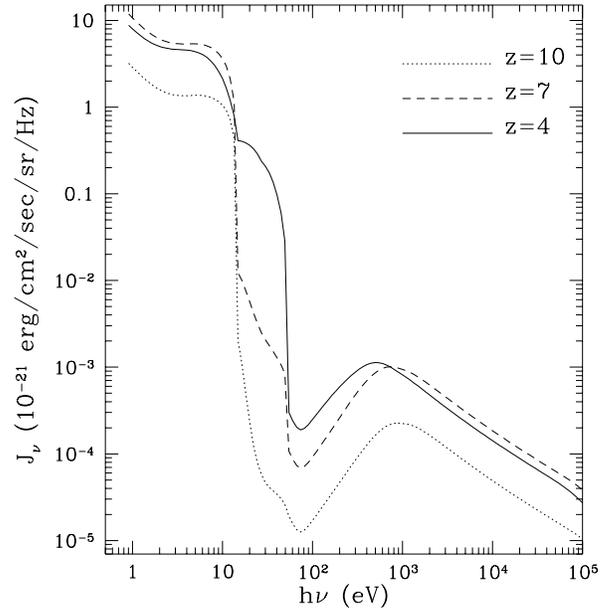


FIG. 11.—The spectra of the background radiation at three different redshifts. Ionization edges corresponding to the hydrogen and helium ionization are well observed at all times.

3.6. Properties of Protogalaxies

We now concentrate on general properties of objects formed in our simulations, which we identify with small protogalaxies. Ostriker & Gnedin (1996) addressed in some detail the question of the various stellar Populations and star formation rates at different epochs. Here we will discuss some of the internal properties of those objects.

Figure 12 shows the fraction of metal-enriched gas as a function of total mass for all objects with the total mass M_{tot} > 10⁸ h⁻¹ M_⊙ and stellar mass M_{*} > 10⁵ h⁻¹ M_⊙ at three different redshifts. We note that there exists a large fraction of objects with as much as 10% of metal-enriched

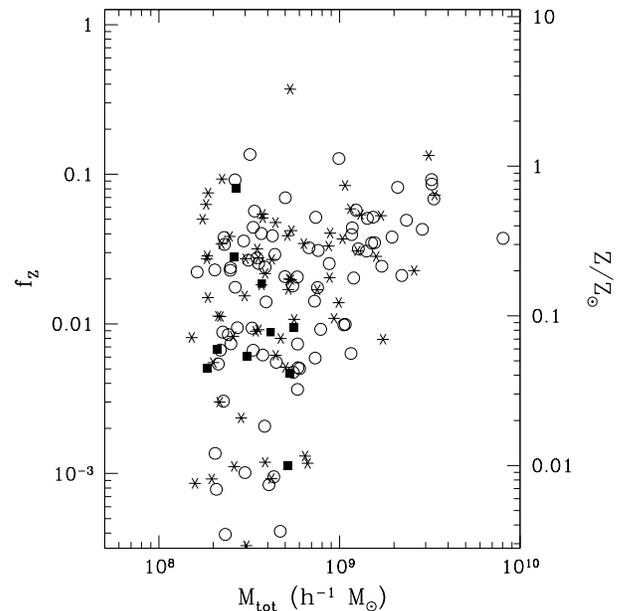


FIG. 12.—The fraction of metal-enriched gas as a function of the total mass for all bound objects with M_{tot} > 10⁸ h⁻¹ M_⊙ and M_{*} > 10⁵ h⁻¹ M_⊙ at z = 9.3 (solid squares), z = 5.9 (stars), and z = 4.3 (open circles). The right-hand scale corresponds to the normalization with the adopted value for the yield y = 0.020.

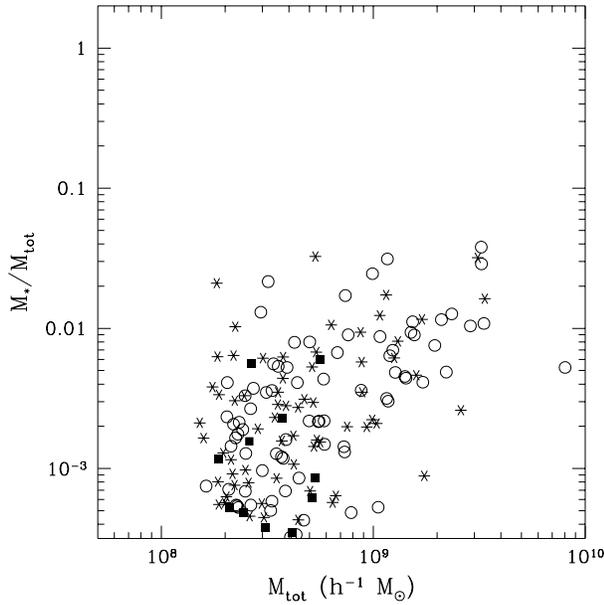


FIG. 13a

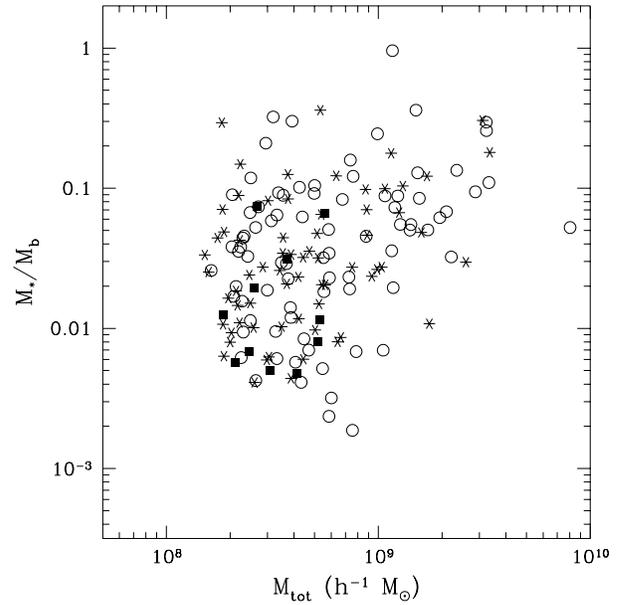


FIG. 13b

FIG. 13.—The fraction of stars in the total mass (a) and in the total baryonic mass (b) as a function of the total mass of an object for all massive objects at three different redshifts as explained before.

gas, which corresponds to solar metallicity if the high value of the yield $y = 0.020$ is adopted as discussed above. We also point out the existence of an object at $z \approx 6$ with at least the solar metallicity (if the standard value for the yield $y = 0.005$ is adopted) or several times the solar metallicity if the higher value of yield is accepted. We therefore conclude that hierarchical clustering models have no difficulty in explaining the high metallicity observed in quasars at high redshifts. We emphasize here that, because of the small size of our computational box, our simulations are strongly biased *against* finding very bright massive quasars, as those are expected to be rare events in the Gaussian initial conditions.

In Figure 13 we show the mass in stars M_* as a fraction of the total mass M_{tot} (a) and the baryonic mass M_b (b) of a given object as a function of the total mass of the object at three different redshifts, as described above. A considerable fraction of all objects have already turned almost half of their gas into stars by $z \approx 4$, but even at the centers, those objects are still dominated by the dark matter; stars constitute on average one-third of the total mass density at the center by $z \approx 4$ for objects with the total mass above $10^9 h^{-1} M_\odot$. This is mostly because of our finite resolution ($1 h^{-1}$ comoving kpc) (which corresponds to the physical scale of around 300 pc at $z \approx 4$) and partly because of the fact that objects in Figure 13 are still efficiently forming stars and

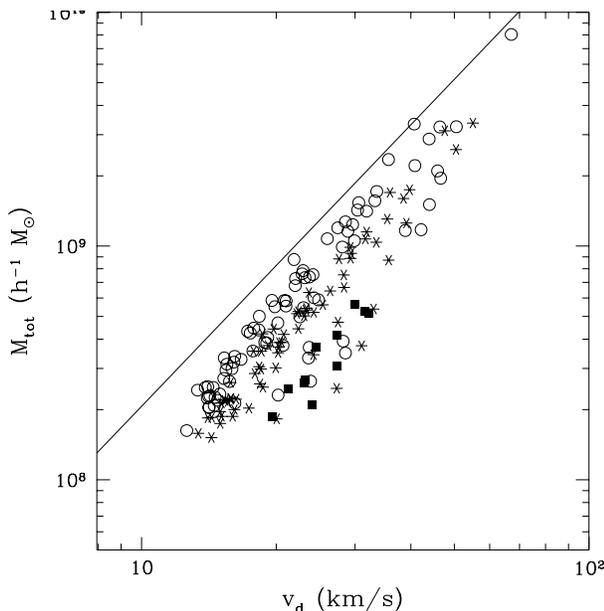


FIG. 14a

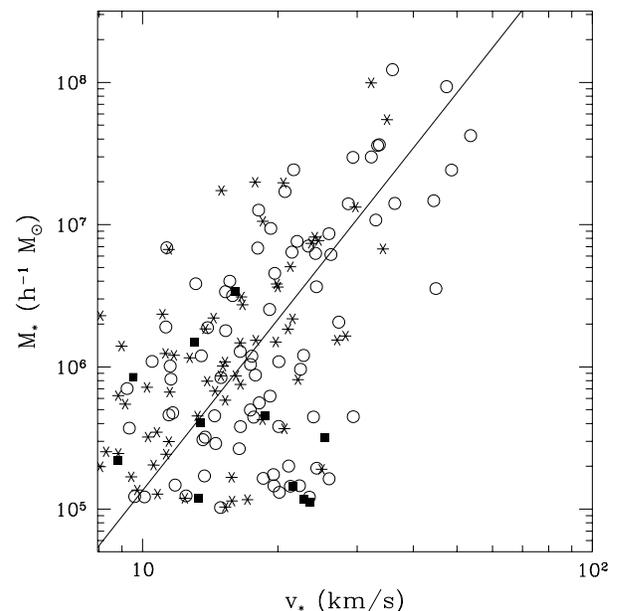


FIG. 14b

FIG. 14.—The total mass as a function of the dark matter velocity dispersion (a) and the stellar mass as a function of stellar velocity dispersion (b) for all massive objects at three different redshifts as explained before. The solid lines have a slope of 2 in (a) and 4 in (b).

actively accreting intergalactic gas that is cooling and concentrating at the center of objects, thus increasing the role of baryons compared to the dark matter with time. Therefore, if we were able to continue our simulations further in time with the appropriate allowance for the missing large-scale waves, we would observe the increasing fraction of baryons at the centers of the protogalaxies.

Let us now discuss the gravitational structure of protogalaxies. We show in Figure 14a the total mass for all objects as a function of dark matter velocity dispersion at three different redshifts, as explained above. The solid line shows the power-law dependence with the power-law index of 2. The fact that our objects follow this power-law shape indicates that they are close to the virial equilibrium. In Figure 14b we plot the total stellar mass as the function of the stellar velocity dispersion, and the solid line is the following law:

$$M_{\text{tot}} = 10^{10.5} h^{-1} M_{\odot} \left(\frac{v_*}{220 \text{ km s}^{-1}} \right)^4. \quad (9)$$

The power-law index is now 4 instead of 2. If we compare this dependence with the Tully-Fisher or Faber-Jackson relation, we will get good agreement provided that the average stellar mass-to-light ratio is $Y = 10^{0.5} h = 2.2$, which is reasonably consistent with the solar neighborhood. We also point out that the normalization of our “Faber-Jackson” relation is roughly independent of redshift. The scatter is apparently substantially larger than the observed scatter in large galaxies (~ 1.5 – 1.7 magnitudes), but it mainly arises from the finite mass resolution of our simulations: if the most massive objects in our simulations contain many thousands of dark matter particles and an even larger number of gas particles, they contain a much smaller number of stellar particles.¹² We also emphasize that active merging is occurring around all massive objects in our simulation, which tends to increase the scatter.

4. CONCLUSIONS

We demonstrate by means of a high-resolution, numerical hydrodynamic simulation how reheating and reionization of the universe occurs in the CDM + Λ cosmological model. We incorporate in our treatment essentially all relevant physical processes that are important for assessing the evolution of average properties of the universe, with one free parameter—the efficiency of high mass star formation—determined by matching the computed and observed background at redshift $z = 4$.

¹² The largest object in the simulation has 21,000 dark matter particles, 23,000 gas particles and 2200 stellar particles, whereas the objects at the bottom of Fig. 14b have only 30–40 stellar particles.

We conclude that a generic prediction of models where the universe is reheated and reionized by high-energy radiation from starlike and quasarlike sources is that slow reheating happens first, followed by sudden reionization. We also demonstrate that hierarchical clustering models are broadly consistent (with the appropriate choice of parameters) with the measured Gunn-Peterson opacity of the intergalactic medium.

Clumping of the gas (with $\langle \rho^2 \rangle / \langle \rho \rangle^2 > 10$) must be allowed for in any computation to realistically estimate the epoch and nature of the reionization process.

The average heavy element abundance in the intergalactic medium can reach values of 1/200 of solar by $z = 4$, with the large dispersion in the low density regions ($\delta \lesssim 30$). The metal abundances in the high-density regions are more uniform and can reach 1/30 on average, but there exist individual bound objects (protogalaxies) that can have solar or higher abundances of heavy elements at $z \sim 6$. This implies that hierarchical clustering models such as the one studied in this paper can easily explain observed high metallicity of quasars at $z = 5$. We also propose a working mechanism based on mergers for transporting the heavy elements, initially locked in the high density regions, into the lower density intergalactic medium, where they are observed in (sufficiently high column density) Ly α absorbers.

Reionization, which suddenly reduces the neutral hydrogen abundance by several orders of magnitude, leaves a distinct signature in the redshifted 21 cm emission from the intergalactic hydrogen. While this emission cannot be measured by existing or proposed meter-wavelength radio telescopes, which lack sensitivity by a factor of several hundred, future ultrahigh-sensitivity experiments may detect this emission, giving us a very strong (because of the suddenness of reionization) constraint on the evolution of the universe at high redshift.

The Faber-Jackson relation, i.e., the correlation between the luminosity and the velocity dispersion of a galaxy, is reproduced by our simulations (assuming constant mass-to-light ratio) with correct amplitude and slope, which indicates that this phenomenon can be explained entirely based on hierarchical clustering model and simple cooling arguments. Both the slope and the amplitude of our Faber-Jackson relation are independent of redshift within numerical uncertainties.

The authors would like to express their gratitude to Prof. Martin Rees for numerous fruitful discussions and valuable comments. We are also thankful to the anonymous referee for valuable comments that substantially improved the manuscript. This work was supported by NSF grant AST 93-18185 awarded to the Grand Challenge Cosmology Consortium.

APPENDIX A

RADIATIVE TRANSFER EQUATIONS IN THE EXPANDING UNIVERSE

Let us start by deriving the full radiative transfer equation in comoving coordinates. Let $F(t, x^i, p^k)$ be the distribution function for photons in comoving coordinates x^i and comoving momentum

$$p^k = a \frac{h\nu}{c} n^k,$$

where $a \equiv a(t)$ is the scale factor, h is the Planck constant, ν is the photon frequency, and n^k is the unit vector in the direction of photon propagation. By definition (Peebles 1980),

$$N_\gamma = \int F(t, x^i, p^k) d^3x d^3p$$

is the number of photons in the universe; N_γ changes only in response to photon absorption or emission in physical processes (ionization, recombination, cooling, etc.). We can, therefore, write down the continuity equation for the function F in the phase space (x^i, p^k) :

$$\frac{\partial F}{\partial t} + \frac{\partial}{\partial x^i} (\dot{x}^i F) + \frac{\partial}{\partial p^k} (\dot{p}^k F) = \text{RHS}(F), \tag{A1}$$

where the term $\text{RHS}(F)$ denotes photon absorption and emission in physical processes.

It is customary, however, to use specific intensity $J_\nu(t, x^i, n^k)$ of radiation instead of $F(t, x^i, p^k)$, where, by definition, $J_\nu d\nu d\Omega dA dt$ is the energy of photons with frequencies from ν to $\nu + d\nu$ passing through the area dA in the time interval dt in the solid angle $d\Omega$ around n^k . It is easy to convert from F to J_ν if one recalls that $a^3 d^3x = c dt dA$ and $d^3p = p^2 dp d\Omega$:

$$J_\nu = h\nu F \frac{d^3x d^3p}{d\nu d\Omega dA dt} = \frac{h^4 \nu^3}{c^2} F. \tag{A2}$$

Substituting F from equation (A2) into equation (A1), we easily obtain the following most general radiative transfer equation:

$$\frac{\partial J_\nu}{\partial t} + \frac{\partial}{\partial x^i} (\dot{x}^i J_\nu) - H \left(\nu \frac{\partial J_\nu}{\partial \nu} - 3 J_\nu \right) = -k_\nu J_\nu + S_\nu, \tag{A3}$$

where H is the Hubble constant, $H \equiv \dot{a}/a$. We specified physical processes including absorption with the coefficient k_ν , and emission (sources) S_ν .

We now consider the mean specific intensity in the universe, averaged over all space and all directions,

$$\bar{J}_\nu(t) \equiv \langle J_\nu(t, x^i, n^k) \rangle, \tag{A4}$$

where the averaging operator $\langle \rangle$ acting on a function $f(x^i, n^k)$ of position and direction is defined as

$$\langle f(x^i, n^k) \rangle = \text{a number} = \lim_{V \rightarrow \infty} \frac{1}{4\pi V} \int_V d^3x \int d\Omega f(x^i, n^k). \tag{A5}$$

Applying the averaging operator to the equation (A3), we obtain the following equation for the mean specific intensity:

$$\frac{\partial \bar{J}_\nu}{\partial t} - H \left(\nu \frac{\partial \bar{J}_\nu}{\partial \nu} - 3 \bar{J}_\nu \right) = -\bar{k}_\nu \bar{J}_\nu + \bar{S}_\nu, \tag{A6}$$

where, by definition,

$$\bar{S}_\nu \equiv \langle S_\nu \rangle,$$

and

$$\bar{k}_\nu \equiv \frac{\langle k_\nu J_\nu \rangle}{\bar{J}_\nu}.$$

We would like to stress here that we use the overbar merely as a part of notation and not to denote the space average; in particular, \bar{k}_ν is not a space average of k_ν , since it is weighted by the local value of specific intensity J_ν .

The purpose of this section is, however, not to derive equation (A6), but rather an equation for fluctuations in the specific intensity around the mean. It is convenient, therefore, to introduce a new quantity called relative (to the mean) specific intensity, $f_\nu(t, x^i, n^k)$, as a measure of deviation of the specific intensity in a particular position in space x^i and in a particular direction n^k :

$$J_\nu \equiv f_\nu \bar{J}_\nu, \tag{A7}$$

so that $\langle f_\nu \rangle = 1$. It is straightforward to derive the following equation for f_ν :

$$\frac{\partial f_\nu}{\partial t} + \frac{\partial}{\partial x^i} (\dot{x}^i f_\nu) = H\nu \frac{\partial f_\nu}{\partial \nu} - (k_\nu - \bar{k}_\nu) f_\nu + \psi_\nu, \tag{A8}$$

where

$$\psi_\nu \equiv \frac{S_\nu - \bar{S}_\nu}{\bar{J}_\nu}. \tag{A9}$$

Equation (A8) is tedious and by no means simpler than the original equation (A3). However, if we restrict ourselves to scales significantly smaller than the horizon size and matter velocities much smaller than the speed of light (Newtonian limit), then

we can simplify equation (A8) substantially by noting that the relative specific intensity does not change substantially and the universe does not expand substantially over the period of time a photon needs to travel over the scale under consideration. Therefore, we can ignore the time derivative and the derivative with respect to the frequency. Under these assumptions, equation (A8) reduces to the following simple equation:

$$\frac{c}{a} n^i \frac{\partial f_\nu}{\partial x^i} = -(k_\nu - \bar{k}_\nu) f_\nu + \psi_\nu. \quad (\text{A10})$$

Equations (A6) and (A10) represent the radiative transfer equations in the expanding universe in the Newtonian limit.

APPENDIX B

LOCAL OPTICAL DEPTH APPROXIMATION TO THE RADIATIVE TRANSFER EQUATIONS

Our goal in this appendix is to describe numerical techniques for solving equations (A6) and (A10) within the framework of simulating large-scale structure and galaxy formation in the universe. The former of those two equations can be reduced to an ordinary differential equation that is solved without difficulty (see Gnedin 1996). The latter, however, is beyond practical implementation for the current computer capabilities and resolution requirements for galaxy formation simulations.

The main difficulty comes from the fact that f_ν is a function of six variables: three spatial coordinates and three coordinates in the momentum space. In addition, unless J_ν by itself is needed in a simulation, only direction-averaged quantities are required to calculate photoionization and photoheating rates that are needed to accurately compute gas temperature, ionization fractions, cooling rates, etc. But even if we approximate the directional dependence of f_ν , it still will be of little use since f_ν will remain a function of four arguments; storing such a function with sufficient number of resolution elements in each of four directions is a challenging problem for modern galaxy formation simulations.

The second practical complication of implementing the radiative transfer equation (A10) is that the ionization state of cosmic gas (and, therefore, the absorption coefficient k_ν) can change on a very short timescale, one substantially shorter than the hydrodynamic Courant timescale. Equation (A10) must therefore be solved millions of times for a typical large cosmological simulation, which is obviously well beyond existing computer power.

Fortunately, there exist two physical reasons why a reasonable approximation to equation (A10) can be found that will employ only functions of three arguments.

There are two main physical effects which produce spatially variable specific intensity: the optical depth of the gas, which can shield high density regions from external radiations (k_ν is a function of position), and nonuniformly distributed sources of radiation (ψ_ν is a function of position). The optical depth is indeed a function of seven arguments: the frequency ν and two spatial positions (optical depth from one point to another); however, its dependence on frequency is uniquely determined since, for known column densities of neutral hydrogen $N_{\text{H I}}$, neutral helium $N_{\text{He I}}$, and singly ionized helium $N_{\text{He II}}$ (we consider only hydrogen and helium contributions to the optical depth since heavy element abundances are small), the optical depth is

$$\tau_\nu \equiv N_{\text{H I}} \sigma_\nu^{\text{H I}} + N_{\text{He I}} \sigma_\nu^{\text{He I}} + N_{\text{He II}} \sigma_\nu^{\text{He II}}, \quad (\text{B1})$$

where σ_ν^i is the photoionization cross section for a species $i = \text{H I}, \text{He I}, \text{He II}$ and is a known function of frequency. In addition, there is no need to know τ_ν with high accuracy, since if the optical depth is small, it does not matter if it is 10^{-3} or 10^{-5} , and if it is large, it again does not matter if it is 10^3 or 10^5 . It is only $\tau_\nu \sim 1$ that should be computed accurately. However, since $\tau_\nu \sim 1$ primarily happens at the boundary of an optically thick region, the total volume occupied by regions with $\tau_\nu \sim 1$ is not large and little error in the physical state of the gas is introduced from a large error in the optical depth calculation.¹³

The second important physical reason that equation (A10) can be reasonably approximated with functions of positions only is that source function ψ_ν dependence on frequency, if known, does not change with time since photons do not have enough time to redshift significantly while traveling a distance short compared to the horizon scale. Since ψ_ν includes only fluctuations in the source function, it is mainly contributed to by a few nearby sources, while the redshift of radiation emitted by distant (and, therefore, uniform) sources is already taken into account by a derivative with respect to ν in equation (A6). Thus, the frequency dependence of relative specific intensity f_ν at every particular point and at every moment in time can be easily calculated from scratch and needs not be saved as an additional dimension for f_ν (in particular, it is the same for all positions).

Let us now proceed further to deriving the simplest possible approximation that has any sense, which we call the local optical depth approximation. First, let us consider the homogeneous equation (A10) when no spatial dependence of sources is taken into account. Let $f_\nu^{(0)}$ be its solution. It is convenient and instructive to introduce optical depth τ_ν , as

$$f_\nu^{(0)}(t, x^i, n^k) \equiv \exp \{ -[\tau_\nu(t, x^i, n^k) - \bar{\tau}_\nu(t)] \}, \quad (\text{B2})$$

where $\bar{\tau}_\nu(t)$ is defined as

$$e^{-\bar{\tau}_\nu} \equiv \langle e^{-\tau_\nu} \rangle$$

¹³ Those handwaving arguments cannot serve as a proof to the approximation described below, but are merely justification for searching for reasonable approximations.

because $\langle f_v \rangle = 1$. Substituting (B2) into (A10) with $\psi_v = 0$, we obtain the following equation:

$$n^i \frac{\partial \tau_v}{\partial x^i} = \frac{a}{c} (k_v - \bar{k}_v). \quad (\text{B3})$$

This equation has the formal solution (omitting the time dependency for clarity) of

$$\tau_v(x^i, n^k) = \tau_v^{(0)}[x^i - n^i(x^k n^k), n^k] + \frac{a}{c} \int_{-(x^k n^k)}^0 [k_v(x^i + n^i l) - \bar{k}_v] dl, \quad (\text{B4})$$

where $\tau_v^{(0)}$ is an arbitrary function of two-dimensional vectors that is specified from boundary conditions. One can see that τ_v is indeed an optical depth in the sense that it is an integral of an absorption coefficient over a distance.

However, as we discussed before, there is no practical sense to implement this equation in a real code, or even a direction-averaged version of the exact solution (in fact, it is not even clear how to calculate the direction-averaged f_v in a closed form from eq. [B4]). More than that, we would eventually like to get the completely local dependence of f_v on properties of the gas as an approximation to f_v ; this would allow us to consider cooling evolution within a hydrodynamic time step of each resolution element independently and to use an ordinary differential equation solver to resolve smallest time-scales in the problem (see Gnedin 1996). We therefore introduce an approximation to f_v by defining τ_v in equation (B2) as a function of position and frequency only by the following *Ansatz*:

$$\tau_v(x^i) \equiv [n_{\text{H I}} \sigma_v^{\text{H I}} + n_{\text{He I}} \sigma_v^{\text{He I}} + n_{\text{He II}} \sigma_v^{\text{He II}}] L, \quad (\text{B5})$$

where $n_{\text{H I}}$, $n_{\text{He I}}$, and $n_{\text{He II}}$ are local number densities of neutral hydrogen, neutral helium, and singly ionized helium, and L is a characteristic distance (which is a function of position, but not direction; the resulting $f_v^{(0)}$ is therefore a function of position and frequency, but not direction). The justification to this *Ansatz* is that it is likely that ionization fractions are similar in the neighboring points, and if we choose the characteristic distance L appropriately, we will obtain a successful approximation to f_v . Since L is obviously related to the characteristic scale over which the density changes, we choose the following form for L :

$$L = \frac{\rho}{\sqrt{\alpha |\nabla \rho|^2 + \beta \rho |\Delta \rho|}}, \quad (\text{B6})$$

where the parameters $\alpha = 2.11$ and $\beta = 0.27$ are fitted to produce the correct value for the column density for a spherical density distribution with inverse square law profile in two limits, $r \rightarrow 0$ and $r \rightarrow \infty$ (two constraints produce two parameter values); these values for α and β also give a very close match to the Gaussian density profile and for r^{-4} (photoionized inverse square law) profiles.

The local optical depth approximation is an uncontrolled approximation in the sense that it seems impossible to define its range of applicability; we can guarantee that it is not zero since it gives correct results for a spherical density distribution with the inverse square law profile and quite accurate results for other density profiles. In addition, we have tested the approximation by solving equation (A10) numerically for density distributions extracted at a few fixed epochs from simulations discussed in this paper; the correspondence between the exact result and the approximate f_v is remarkably good. The epochs, however, have been chosen after the universe reionizes, so f_v deviates from unity appreciably only within isolated dense optically thick clouds, and the local optical depth approximation seems to give reasonable results in this limit. When the most of the universe is “optically thick” (i.e., when $\bar{\tau}_v > 1$), the approximation is less accurate, but since $\tau_v > 1$ in most of the volume of the universe in that case, big errors in τ_v do not lead to substantial errors in ionization abundances, as has been argued above.

Now, equipped with the approximation to the solution $f^{(0)}$ for the homogeneous radiative transfer equation, we can obtain solution for the whole equation (A10). Since the direction n^k enters equation (A10) only as a parameter, we can introduce a distance l along the direction n^k and equation (A10) becomes an ordinary differential equation that can be easily solved (we assume further that ψ_v does not depend on direction, i.e., that radiation sources are isotropic; in the opposite case dependence on direction can not be eliminated and the described approximation becomes invalid):

$$f_v(x^i, n^k) = e^{-[\tau_v(x^i) - \bar{\tau}_v]} + \frac{a}{c} \int_0^\infty dl \psi_v(x^i + n^i l) e^{-\tau_v(x^i, x^i + n^i l)}, \quad (\text{B7})$$

where we specify $\tau_v(x^i, x^i_1)$ below. Since no directional dependence is required in order to calculate effects of radiation on the gas, we average equation (B7) over all directions; in order to keep notation compact, we retain the same symbol $f_v(x^i)$ for the direction averaged $f_v(x^i, n^k)$, i.e.,

$$f_v(x^i) \equiv \frac{1}{4\pi} \int f_v(x^i, n^k) d\Omega.$$

Then equation (B7) becomes

$$f_v(x^i) = e^{-[\tau_v(x^i) - \bar{\tau}_v]} + \frac{a}{4\pi c} \int d\Omega \int_0^\infty dl \psi_v(x^i + n^i l) e^{-\tau_v(x^i, x^i + n^i l)}. \quad (\text{B8})$$

If we introduce a new vector $x_1^i \equiv x^i + n^k l$ and notice that $l = |x^i - x_1^i|$ and $d^3 x_1 = l^2 dl d\Omega$, we obtain the following equation for $f_v(x^i)$:

$$f_v(x^i) = e^{-[\tau_v(x^i) - \bar{\tau}_v]} + \frac{a}{4\pi c} \int d^3 x_1 \frac{\psi_v(x_1^i)}{(x^i - x_1^i)^2} e^{-\tau_v(x^i, x_1^i)}, \quad (\text{B9})$$

where $\tau_v(x^i, x_1^i)$ is the true optical depth from point x^i to point x_1^i .

The integral of equation (B9), except for the exponential factor, is a potential for the force with $1/r^3$ force law, which can be solved for by known numerical techniques (P³M, for example). The exponential factor, however, offers a considerable difficulty in solving integral (B9) faster than in $O(N^2)$ operations. We therefore leave this form of the solution to the radiative transfer equation (A10) until further work.

APPENDIX C

MOLECULAR HYDROGEN

In this appendix, we describe our method for solving for nonequilibrium time-dependent molecular hydrogen abundance. The method we adopt is exact but assumes that the molecular hydrogen abundance is small compared to the neutral hydrogen abundance; thus, we can follow the ionization balance without taking molecular hydrogen formation into account.

This approach significantly reduces the computational expense of adding three more equations in computing ionization balance and chemical evolution. Without including the molecular hydrogen and H_2^+ and H^- ions, there are five evolution equations for five ionization states of hydrogen and helium. The SLH code uses a Rosenbrock implicit method to solve those equations, which, in general, requires an arbitrary, not necessarily symmetric, five-dimensional matrix at each ionization time step to be inverted. However, since the total number of hydrogen and helium species does not change, it is possible to rewrite the Rosenbrock method so that it is necessary to invert only a three-dimensional matrix at each time step, while still storing all five ionization species separately to avoid numerical loss of accuracy. This simplification is not possible when molecular hydrogen chemistry is included, and using the Rosenbrock method for all eight evolution equations directly would require inverting an eight-dimensional matrix at each ionization time step—unaffordable given the current level of computer resource availability. With the approach described below, we can solve all eight evolution equations and still invert only a three-dimensional matrix at each ionization time step.

An alternative approach to including molecular hydrogen chemistry is presented in Abel et al. (1996) and Anninos et al. (1996).

We use Shapiro & Kang (1987) as our source of reaction rates and Lepp & Shull (1983) as the source for cooling rates. Table 2 contains all relevant reactions together with the number from Shapiro & Kang (1987); we use this number as an identification for a reaction.

Assuming that we know the time dependence of hydrogen and helium ion abundances as a function of time, we can write down the following three equations for three chemical species under consideration: H_2 , H_2^+ , and H^- .

$$\frac{1}{n_b} \frac{dX_{\text{H}^-}}{dt} = -X_{\text{H}^-}(k_8 X_{\text{H}^+} + k_{16} X_e + k_{17} X_{\text{H}^+} + k_{18} X_{\text{H}^+} + k_{19} X_{\text{H}^+} + k_{21} X_{\text{H}_2^+} + p_{27}) + (k_7 X_{\text{H}^+} X_e + k_{13} X_{\text{H}_2} X_e), \quad (\text{C1a})$$

$$\frac{1}{n_b} \frac{dX_{\text{H}_2^+}}{dt} = -X_{\text{H}_2^+}(k_{10} X_{\text{H}^+} + k_{20} X_e + k_{21} X_{\text{H}^-} + p_{28} + p_{30}) + (k_9 X_{\text{H}^+} X_{\text{H}^+} + k_{12} X_{\text{H}_2} X_{\text{H}^+} + k_{19} X_{\text{H}^+} X_{\text{H}^-} + p_{29} X_{\text{H}_2}), \quad (\text{C1b})$$

and

$$\begin{aligned} \frac{1}{n_b} \frac{dX_{\text{H}_2}}{dt} = & -X_{\text{H}_2}(k_{11} X_{\text{H}^+} + k_{12} X_{\text{H}^+} + k_{13} X_e + k_{14} X_e + k_{15} X_{\text{H}_2} + p_{29} + p_{31}) \\ & + (k_8 X_{\text{H}^+} X_{\text{H}^-} + k_{10} X_{\text{H}^+} X_{\text{H}_2^+} + k_{21} X_{\text{H}_2^+} X_{\text{H}^-}), \end{aligned} \quad (\text{C1c})$$

where the fractional abundance X_j for a species j is defined as

$$n_j \equiv X_j \frac{\rho}{m_{\text{H}}}, \quad (\text{C2})$$

where m_{H} is the mass of a hydrogen atom and ρ is the gas mass density. The total baryon number density n_b is

$$n_b \equiv \frac{\rho}{m_{\text{H}}}, \quad (\text{C3})$$

TABLE 2
MOLECULAR HYDROGEN CHEMISTRY RATES

Number ^a	Reaction	Rate Notation
7	$\text{H I} + e \rightarrow \text{H}^-$	k_7
8	$\text{H I} + \text{H}^- \rightarrow \text{H}_2 + e$	k_8
9	$\text{H I} + \text{H II} \rightarrow \text{H}_2^+$	k_9
10	$\text{H I} + \text{H}_2^+ \rightarrow \text{H}_2 + \text{H II}$	k_{10}
11	$\text{H I} + \text{H}_2 \rightarrow 3\text{H I}$	k_{11}
12	$\text{H}_2 + \text{H II} \rightarrow \text{H I} + \text{H}_2^+$	k_{12}
13	$\text{H}_2 + e \rightarrow \text{H I} + \text{H}^-$	k_{13}
14	$\text{H}_2 + e \rightarrow 2\text{H I} + e$	k_{14}
15	$\text{H}_2 + \text{H}_2 \rightarrow \text{H}_2 + 2\text{H I}$	k_{15}
16	$\text{H}^- + e \rightarrow \text{H I} + 2e$	k_{16}
17	$\text{H}^- + \text{H I} \rightarrow 2\text{H I} + e$	k_{17}
18	$\text{H}^- + \text{H II} \rightarrow 2\text{H I}$	k_{18}
19	$\text{H}^- + \text{H II} \rightarrow \text{H}_2^+ + e$	k_{19}
20	$\text{H}_2^+ + e \rightarrow 2\text{H I}$	k_{20}
21	$\text{H}_2^+ + \text{H}^- \rightarrow \text{H}_2 + \text{H I}$	k_{21}
27	$\text{H}^- + \gamma \rightarrow \text{H I} + e$	k_{27}
28	$\text{H}_2^+ + \gamma \rightarrow \text{H I} + \text{H II}$	k_{28}
29	$\text{H}_2 + \gamma \rightarrow \text{H}_2^+ + e$	k_{29}
30	$\text{H}_2^+ + \gamma \rightarrow 2\text{H II} + e$	k_{30}
31	$\text{H}_2 + \gamma \rightarrow 2\text{H I}$	k_{31}

^a As in Shapiro & Kang 1987.

and temperature and X_j for $j = \text{H I}, \text{H II}, \text{He I}, \text{He II}, \text{He III}, e$ are known functions of time.

Equations (C1a)–(C1c) are at a disadvantage in that they are nonlinear and, therefore, do not offer a simple solution. The nonlinearity comes in with two terms: $k_{15} X_{\text{H}_2}^2$ and $k_{21} X_{\text{H}_2^+} X_{\text{H}^-}$. However, there is a lucky coincidence that each of those terms is superseded by another term in each of equations (C1a)–(C1c):

1. $k_{21} X_{\text{H}_2^+} X_{\text{H}^-} \ll k_{18} X_{\text{H II}} X_{\text{H}^-}$ in equation (C1a) for H^- since $k_{21} \approx k_{18}$;
2. $k_{21} X_{\text{H}_2^+} X_{\text{H}^-} \ll k_{20} X_{\text{H}_2^+} X_e$ in equation (C1b) for H_2^+ since $k_{21} \sim k_{20}$;
3. $k_{21} X_{\text{H}_2^+} X_{\text{H}^-} \ll k_8 X_{\text{H I}} X_{\text{H}^-}$ and $k_{15} X_{\text{H}_2}^2 \ll k_{11} X_{\text{H}_2} X_{\text{H I}}$ in equation (C1c) for H_2 since $k_{21} \lesssim 10^3 k_8$ and $k_{15} \sim k_{11}$, provided that $X_{\text{H}_2^+} \ll 10^{-3} X_{\text{H I}}$ (this condition is always satisfied in simulations described in this paper).

Neglecting two nonlinear terms, we get a system of three *linear* ordinary differential equations which offers an *analytical* solution.

Let us introduce a three-dimensional vector, Y :

$$Y = \begin{pmatrix} X_{\text{H}^-} \\ X_{\text{H}_2^+} \\ X_{\text{H}_2} \end{pmatrix}, \quad (\text{C4})$$

Then equations (29)–(31) can be written in a matrix form as

$$\dot{Y} = -AY + S, \quad (\text{C5})$$

where A is a matrix,

$$A = n_b \begin{pmatrix} D_1 + k_8 X_{\text{H I}} + k_{19} X_{\text{H II}} & 0 & -k_{13} X_e \\ -k_{19} X_{\text{H II}} & D_2 + k_{10} X_{\text{H I}} & -k_{12} X_{\text{H II}} - p_{29} \\ -k_8 X_{\text{H I}} & -k_{10} X_{\text{H I}} & D_3 + k_{12} X_{\text{H II}} + k_{13} X_e + p_{29} \end{pmatrix}, \quad (\text{C6})$$

D is a vector,

$$D = \begin{pmatrix} k_{17} X_{\text{H I}} + k_{16} X_e + k_{18} X_{\text{H II}} + p_{27} \\ k_{20} X_e + p_{28} + p_{30} \\ k_{11} X_{\text{H I}} + k_{14} X_e + p_{31} \end{pmatrix}, \quad (\text{C7})$$

and S is a vector,

$$S = n_b \begin{pmatrix} k_7 X_{\text{H I}} X_e \\ k_9 X_{\text{H I}} X_{\text{H II}} \\ 0 \end{pmatrix}. \quad (\text{C8})$$

Equation (C5) can be solved *analytically* in an ionization time step Δt from t_0 to $t_1 = t_0 + \Delta t$, in which both A and S do not change significantly, with the $O(\Delta t^2)$ accuracy, to obtain:

$$Y(t_1) = e^{-\bar{A}\Delta t} Y(t_0) + (1 - e^{-\bar{A}\Delta t}) \bar{A}^{-1} \bar{S}, \quad (\text{C9})$$

where

$$\bar{A} \equiv \frac{1}{\Delta t} \int_{t_0}^{t_1} A(t) dt \quad (\text{C10})$$

and

$$\bar{S} \equiv \frac{1}{\Delta t} \int_{t_0}^{t_1} S(t) dt . \quad (\text{C11})$$

Equation (C9) then gives the molecular hydrogen and H_2^+ and H^- fractions at the time moment t_1 .

REFERENCES

- Abel, T. 1995, PhD thesis, Univ. of Regensburg
 Abel, T., Anninos, P., Zhang, Y., & Norman, M. L. 1996, *New Astronomy*, in press (astro-ph 9608040)
 Anninos, P., Zhang, Y., Abel, T., & Norman, M. L. 1996, *New Astronomy*, in press (astro-ph 9608041)
 Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, *ApJ*, 304, 15
 Bertschinger, E. 1995, astro-ph 9506070
 Cen, R., & Ostriker, J. P. 1992, *ApJ*, 399, L113
 Cooke, A. J., Espey, B., & Carswell, B. 1996, *MNRAS*, 269, 552
 Couchman, H. M. P., & Rees, M. J. 1986, *MNRAS*, 221, 53
 Cowie, L. L., McKee, C. F., & Ostriker, J. P. 1981, *ApJ*, 247, 908
 Franceschini, A., Toffolatti, L., Danese, L., & De Zotti, G. 1989, *ApJ*, 344, 35
 Giallongo, E., D'Odorico, S., Fontana, A., McMahon, H. G., Savaglio, S., Cristiani, S., Molaro, P., & Treverse, D. 1994, *ApJ*, 425, L1
 Giroux, M. L., & Shapiro, P. R. 1996, *ApJS*, 102, 191
 Gnedin, N. Y. 1996, *ApJ*, 456, 1
 Gnedin, N. Y., & Bertschinger, E. 1996, *ApJ*, 470, 115
 Gunn, J. E., & Peterson, B. A. 1965, *ApJ*, 142, 1633
 Haiman, Z., & Loeb, A. 1996, *ApJ*, submitted (astro-ph 9611028)
 Haiman, Z., Thoul, A. A., & Loeb, A. 1995, *ApJ*, 464, 523
 Hui, L., Gnedin, N. Y., & Zhang, Y. 1996, *ApJ*, submitted (astro-ph 9608157)
 Jenkins, E. B., & Ostriker, J. P. 1991, *ApJ*, 376, 33
 Katz, N., Weinberg, D. H., & Hernquist, L. 1996, *ApJS*, 105, 1
 Kawasaki, M., & Fugita, M. 1994, *MNRAS*, 269, 563
 Kulsrud, R. M., Cen, R., Ostriker, J. P., & Ryu, D. 1996, *ApJ*, submitted (astro-ph 9607141)
 Lepp, S., & Shull, J. M. 1983, *ApJ*, 270, 578
 Miralda-Escude, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, *ApJ*, 471, 582
 Miralda-Escude, J., & Rees, M. J. 1994, *MNRAS*, 266, 343
 Ostriker, J. P., & Cen, R. 1996, *ApJ*, 464, 27
 Ostriker, J. P., & Gnedin, N. Y. 1996, *ApJ*, 472, L63
 Ostriker, J. P., & Steinhardt, P. J. 1995, *Nature*, 377, 600
 Peebles, P. J. E. 1980, *The Large Scale Structure of the Universe* (Princeton: Princeton Univ. Press)
 Rauch, M., Haehnelt, M. G., & Steinmetz, M. 1996, *ApJ*, submitted (astro-ph 9609083)
 Savaglio, S., Cristiani, S., D'Odorico, S., Fontana, A., Giallongo, E., Molaro, P. 1996, in print (astro-ph 9606063)
 Scott, D., & Rees, M. 1990, *MNRAS*, 247, 510
 Shapiro, P. R. 1995, in *ASP Conf. Proc. 80, The Physics of the Interstellar Medium and the Intergalactic Medium*, ed. A. Ferrara, C. F. McKee, C. Heiles, & P. R. Shapiro (San Francisco: ASP), 53
 Shapiro, P. R., & Giroux, M. L. 1987, *ApJ*, 321, L107
 Shapiro, P. R., Giroux, M. L., & Babul, A. 1994, *ApJ*, 427, 25
 Shapiro, P. R., & Kang, H. 1987, *ApJ*, 318, 32
 Silk, J. 1996, *ApJ*, 481, 703
 Songaila, A., & Cowie, L. L. 1996, *AJ*, 112, 335
 Tegmark, M., & Silk, J. 1994, *ApJ*, 423, 529
 ———. 1995, *ApJ*, 441, 458
 Tegmark, M., Silk, J., Rees, M., Blanchard, A., Abel, T., & Palla, F. 1997, *ApJ*, 474, 1
 Zhang, Y., Norman, M. L., Anninos, P., & Abel, T. 1996, in *Star Formation, Near and Far*, ed. S. S. Holt & L. G. Mundy (Woodbury: AIP), in press