Reionization of the Inhomogeneous Universe

Jordi Miralda-Escudé[1,4], Martin Haehnelt[2,3], Martin J. Rees[2]
[1] Univ. of Pennsylvania, Dept. of Physics and Astronomy
[2] Institute of Astronomy, University of Cambridge
[3] Max-Planck-Institut für Astrophysik, Garching
[4] Alfred P. Sloan Fellow

ABSTRACT

A model of the density distribution in the intergalactic medium, motivated by that found in numerical simulations, is used to demonstrate the effect of a clumpy IGM and discrete sources on the reionization of the universe. In an inhomogeneous universe reionization occurs outside-in, starting in voids and gradually penetrating into overdense regions. Reionization should not be sudden but gradual, with a continuous rise of the photon mean free path over a fair fraction of the Hubble time as the emissivity increases. We show that a hydrogen Gunn-Peterson trough should be present at $z \simeq 6$ unless the emissivity increases with redshift at $z > 4$. However, the epoch of overlap of cosmological H II regions could have occurred at a higher redshift if sources of low luminosity reionized the IGM; the Gunn-Peterson trough at $z \sim 6$ would then appear because even the most underdense voids have a large enough neutral fraction in ionization equilibrium to be optically thick to Lyα photons. Cosmological H II regions near the epoch of overlap can produce gaps of transmitted flux only if luminous quasars contributed to the reionization. Despite the clumpiness of the matter distribution, recombinations are not very important during the reionization of hydrogen because the high density gas is not ionized until a late time. We show that the He II reionization was most likely delayed relative to the hydrogen reionization, but should be completed by $z \sim 3$, the redshift where observations are available. The reported large optical depth fluctuations of He II are probably not due to an incomplete He II reionization, but arise from a combination of density fluctuations and the variations in the intensity of the ionizing background due to luminous QSO's.

Subject headings: cosmology: theory — intergalactic medium — large-scale structure of universe — quasars: absorption lines

1. Introduction

In the standard Big Bang model, the primordial gas becomes neutral at the recombination epoch, at redshift of $z \simeq 1100$. The absence of the hydrogen Gunn-Peterson trough (which would be expected in high redshift objects if hydrogen were neutral) implies that most of the hydrogen in the intergalactic medium (hereafter, IGM) was highly ionized at redshifts $z \lesssim 5$ (Gunn & Peterson 1965, Bahcall & Salpeter 1965, Scheuer 1965; see Schneider, Schmidt, & Gunn 1991, Dey et al. 1998, Weymann et al. 1998 for the most recent evidence at the highest redshifts). In the case of helium, only four quasars have been observed at the He II Lyα wavelength near a redshift $z = 3$ (Jakobsen et al. 1994, 1996; Davidsen, Kriss, & Zheng 1996; Hogan, Anderson, & Rugers 1997; Anderson et al. 1998). These four quasars show that a large fraction of the flux is absorbed (about 75%, much larger than for hydrogen at the same redshift), but with no evidence yet of a transition to a complete Gunn-Peterson trough. Reimers et al. (1997) have suggested that variations in the ratio of the He II to the H I optical depth is evidence that the overlap of He III regions was not yet complete at $z \simeq 3$, but this is still controversial (Miralda-Escudé 1998, Anderson et al. 1998).

The main source of opacity for ionizing photons is Lyman absorption in dense regions of the intergalactic medium. The mean free path of ionizing photons decreases strongly with redshift and becomes smaller than the Hubble radius at $z \gtrsim 2$. Therefore, not only the universe has to be reionized at some epoch, but a constant supply of ionizing photons is required to keep the Universe ionized and to maintain the observed UV background (e.g., Madau 1998 and references therein). Despite considerable observational and theoretical efforts, the sources of ionizing photons are still unclear. When investigating the role of potential sources of ionizing photons, it is obviously important to quantify how many photons are actually needed to first reionize the universe and then keep it ionized.

So far, most analytic treatments of the reionization process have assumed the IGM to be homogeneous (e.g., Arons & Wingert 1972; Shapiro & Giroux 1987). However, the IGM is well known to be highly inhomogeneous, owing to the non-linear collapse of structure. The inhomogeneity of the gas distribution affects the requirement on the number of photons needed for the reionization. The effect that has usually been considered is the simple increase in the number of recombinations that can take place due to the clumping factor of the gas (e.g., Madau, Haardt, & Rees 1998). In this paper, we address the effects of inhomogeneity in more detail. We shall see that another important factor is that reionization does not occur simultaneously throughout the universe. In the model we propose here, the higher the gas density, the later the gas will be ionized by the cosmic background. This can dramatically reduce the number of photons that need to be emitted before a large fraction of the volume in the IGM is reionized and the universe starts to become transparent.

The inhomogeneity of the gas distribution is also of primary importance to understand the absorption spectra of high redshift sources. The mean flux decrement of the photoionized intergalactic medium (or Lyα forest) increases with redshift. Eventually, as the epoch of reionization is approached, the Lyα forest should turn into a series of "clearings", or "gaps" of transmitted flux appearing through the otherwise completely opaque Gunn-Peterson trough, and the number of clearings should decline rapidly with increasing redshift. Quantitative predictions are needed to assess how this transition takes place in detail, depending on the spatial distribution of gas in the IGM and on the luminosity function and lifetime of the sources of ionizing photons.

In recent years, numerical simulations have led to considerable progress in the understanding of the density distribution and the ionization state of the IGM (Cen et al. 1994; Hernquist et al. 1996; Miralda-Escudé et al. 1996; Zhang, Anninos, & Norman 1995; Zhang et al. 1998). First attempts have also been made to incorporate ionizing sources and to investigate the corresponding evolution of the ionizing background during the epoch of reionization (Gnedin & Ostriker 1997). Here we use this knowledge of the density distribution of the IGM to make a simple model of the way reionization proceeds.

In §2 we define the concept of the global recombination rate, and show how it determines the advance of reionization. In §3 the evolution of the intensity of the background is expressed in terms of the mean-free path, and we discuss the effects that determine the number of gaps of transmitted flux visible in Lyα spectra as the epoch of reionization is approached. In §4 we study the conditions under which QSO's or star-forming galaxies can meet the requirements for reionization, and we discuss the question of the detectability of single H II regions around individual sources in the Lyα spectra. §5 discusses a number of special aspects of helium reionization, and §6 presents our conclusions. Unless otherwise specified, we use the cosmological model $H_0 = 65\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_0 = 0.3$, $\Lambda_0 = 0.7$.

2. Reionization of an Inhomogeneous Intergalactic Medium

The first discussion of the process of the reionization of hydrogen in the universe assumed an IGM with uniform density (Arons & Wingert 1972). It was immediately clear that the ionization would not take place homogeneously, but in individual cosmological H II regions growing around the first sources of ionizing photons. The mean free path of ionizing photons in the neutral IGM, λ, is very small [$\lambda = 0.28\,(\Omega_b h^2/0.02)^{-1}\,(1+z)^{-3}\,\mathrm{Mpc}$

at the threshold frequency, in proper units]. Unless the sources were very numerous and of low luminosity (with a number density as high as λ^{-3}), the thickness of the ionization fronts should be very small compared to the size of the H II regions, and the IGM should be divided into H II regions around every source and neutral regions in between. As the mean emissivity in the universe grows, the H II regions grow in size until they overlap. Two conditions need to be satisfied before the IGM is completely ionized: one photon for each atom in the IGM needs to be emitted, and the mean photon emission rate must exceed the recombination rate so that the hydrogen atoms are photoionized faster than they can recombine. If recombinations are never important, the H II regions will never reach their Strömgren radius before overlapping. If recombinations are important, then the H II regions will fill an increasing fraction of the IGM as the emissivity rises and the recombination rate is reduced due to the expansion of the universe. In both of these two cases (and assuming that the size of the H II regions is much smaller than the horizon), the model of the homogeneous IGM implies that there is a universal time when the last neutral regions are ionized. The mean free path of the ionizing photons and the intensity of the background increase very abruptly at this time. The reason for this sudden increase is that the universe is still very opaque to ionizing photons when only a small fraction of the atoms remain to be ionized, and it then takes a very short time to ionize these last atoms.

Based on this model, the argument has been put forward that there should be a well-defined "redshift of reionization" at which this sudden rise of intensity takes place (e.g. Haiman & Loeb 1998 and references therein). This conclusion is, however, an artifact of considering a model of the IGM with uniform density. Furthermore, even in a homogeneous model, only the initial increase of the ionizing flux will occur on a timescale short compared to the Hubble time. The build-up of the final flux and the decrease of the mean opacity to values of order unity still takes a fair fraction of the Hubble time for any realistic redshift evolution of the emissivity of ionizing photons.

We shall find in this paper that, when studying the problem of how the increasingly dense Lyα forest turns into a Gunn-Peterson trough at high redshift, the use of an IGM with homogeneous density is not just a poor approximation, but it results in a completely misleading picture of what actually happens. The fact that the IGM is highly inhomogeneous must be considered right from the start.

Let us therefore consider the volume-weighted probability distribution of the overdensity in the IGM, $P_V(\Delta)$, where $\Delta = \rho/\bar{\rho}$, ρ is the gas density and $\bar{\rho}$ is the mean density of baryons. We define also the cumulative distribution of gas with overdensity less than Δ,

$$F_V(\Delta) = \int_0^{\Delta} \mathrm{d}\Delta P(\Delta) \ . \tag{1}$$

A crucial concept in our analysis will be that of the *global* recombination rate. Imagine that all the gas at density $\Delta < \Delta_{\rm i}$ is ionized, while the higher density gas is neutral. Then, the mean number of recombinations per Hubble time that take place for each baryon in the universe is

$$R(\Delta_{\rm i}) = R_{\rm u} \int_0^{\Delta_{\rm i}} {\rm d}\Delta \, P(\Delta)\Delta^2 \ . \tag{2}$$

Here, $R_{\rm u}$ is the ratio of the recombination rate for a homogeneous universe to the Hubble constant, given by $R_{\rm u}(z) = 0.035 \, [\Omega_{\rm b} h (1 - 3Y/4)/0.025] \, (1+z)^{3/2}/\Omega_0^{1/2}$ for hydrogen (we use the high redshift approximation $\Omega(z) \simeq 1$, and a recombination coefficient $\alpha = 4 \times 10^{-13} \, {\rm cm^3 \, sec^{-1}}$, for temperature $T = 10^4$ K; the factor $1 - 3Y/4$ is the ratio of electron to baryon density, valid when helium is only once ionized). For $\Omega_0 = 0.3$ and $h = 0.65$, $R_{\rm u} = 1$ at $z \simeq 6$. We neglect here the effect of the mean variation of the temperature with the density, which can affect the recombination rate due to the temperature dependence of the recombination coefficient.

Now, consider what should take place in an inhomogeneous medium as the sources of ionizing photons appear. Clearly, the sources will be located in high density regions (galaxies), where stars or quasars can form. To affect the IGM, the radiation must first be able to escape from the dense gas. Sources that are too weak may never be able to ionize the surrounding dense gas, and will never contribute to the reionization. The radiation from the sources that do contribute will at some point ionize the surrounding gas, at least in some directions where the gas column density is lowest. Even if the source emits isotropically, the emission escaping into the IGM may be anisotropic due to the internal absorption. Once the photons have escaped, the cosmological H II region will expand fastest along the directions of lowest gas density, both because fewer atoms need to be ionized per unit volume, and because fewer recombinations take place. These H II regions will therefore expand very rapidly into the voids, but will be halted in the directions where they encounter a clump of dense gas. If recombinations are neglected, an isotropic source will have ionized a fixed column density in every direction at a given time; when recombinations are important, the advance of the H II region will be further slowed down along directions of high column density.

The picture of reionization that is implied, if the sources of radiation are numerous enough, is that the H II regions will generally expand along the voids and overlap through the sheets of the structure of the IGM formed by gravitational collapse, leaving the denser filaments and halos to be ionized later, when the intensity of radiation rises. This suggests that, as soon as most of the volume of the IGM is ionized, equation (2) provides a useful approximation to the mean number of recombinations per baryon taking place per Hubble time, with the parameter $\Delta_{\rm i}$ determining the stage of the ionization. In general, equation

(2) is a poor approximation before the epoch of overlap, because some low-density regions will remain neutral as long as the mean free path of ionizing photons is smaller than the mean separation between sources. If the sources are rare and highly luminous, overlap should occur later, and a large number of density structures need to be ionized by every source; this case will be discussed in more detail later.

We can now write an equation governing the reionization of this inhomogeneous IGM. Let $\epsilon(t)$ be the mean volume emissivity, measured in terms of the number of ionizing photons (with energy above 13.6 eV) emitted for each atom in the universe per Hubble time. We define also the ratio of the mean density of ionizing photons present in the cosmic background to the mean density of atoms, n_J. Then we have

$$\epsilon = \frac{\mathrm{d}n_J}{H\,\mathrm{d}t} + \frac{\mathrm{d}F_M(\Delta_\mathrm{i})}{H\,\mathrm{d}t} + R(\Delta_\mathrm{i}) + n_J\gamma \ , \tag{3}$$

where F_M is the fraction of mass with density $\Delta < \Delta_\mathrm{i}$. The last term includes the losses of ionizing photons due to redshift. The value of the constant γ depends on the spectrum of the background,

$$\gamma = \frac{I_\nu(\nu_T)}{\nu_T} \left(\int_{\nu_T}^{\infty} d\nu \frac{I_\nu}{\nu} \right)^{-1} \ , \tag{4}$$

where I_ν is the background intensity per unit frequency and ν_T is the threshold frequency for ionization. Equation 3 reflects the statement that for every ionizing photon that is emitted, there must be either a photon added to the existing ionizing background, or a new atom being ionized for the first time, or an atom recombining to compensate for the absorption of an ionizing photon, or a photon being redshifted below the ionization edge of hydrogen (we neglect here the fact that helium atoms have a different recombination coefficient than hydrogen atoms, as well as photons used to doubly ionize helium). To solve this equation, the function $\epsilon(t)$ must be specified from a model for the evolution of the sources of ionizing photons, and R and F_M are expressed as a function of Δ_i in equations (2) and (1) once a given model for the density distribution in the IGM is adopted. Thus, we have the two independent variables n_J and Δ_i, and an additional equation is required to solve for them. This second equation relates $\mathrm{d}F_M(\Delta_\mathrm{i})/\,\mathrm{d}t$ to the rates of ionizations and recombinations, and incorporates the mean free path of the photons, which depends on the spatial distribution of the gas clumps in the IGM.

However, in the limit where $n_J \ll \epsilon$, we can avoid this second equation altogether by neglecting n_J, and then equation (3) becomes

$$\epsilon = \mathrm{d}F_M(\Delta_\mathrm{i})/(H\,\mathrm{d}t) + R(\Delta_\mathrm{i}) \ . \tag{5}$$

This approximation is valid when the mean free path of the ionizing photons, λ_i, is much smaller than the horizon, and the emissivity can be assumed to be a smooth function of time without a very sudden universal increase. In this case, $n_J \simeq \epsilon\,(H\lambda_{\rm i}/c)$ depends only on the instantaneous value of the emissivity because all the photons are absorbed shortly after being emitted. We shall generally use this approximation in this paper.

Equation (2) is used here as an approximation for the case where the gas clumps with densities $\Delta > \Delta_{\rm i}$ are large enough to be self-shielding, so that the higher density gas is entirely neutral, and the ionized fraction declines very rapidly as the density approaches $\Delta_{\rm i}$ due to the reduction of the intensity caused by self-shielding. Of course, in practice the mean neutral fraction will vary smoothly with Δ, and one may define $\Delta_{\rm i}$ as the overdensity where the neutral fraction is equal to one half. An alternative model we shall use for the global recombination rate is to assume a uniform background intensity and gas temperature determining the neutral fraction everywhere by requiring ionization equilibrium (this is the limit when all the clumps are very small and optically thin). This results in

$$R(\Delta_{\rm i}) = R_{\rm u} \int_0^\infty \,{\rm d}\Delta\, P(\Delta) x_{\rm i}^2 \Delta^2 \ , \tag{6}$$

$$x_{\rm i} = \frac{\Delta_{\rm i}}{4\Delta} \left[\left(1 + \frac{8\Delta}{\Delta_{\rm i}} \right)^{1/2} - 1 \right] \ , \tag{7}$$

where $x_{\rm i}$ is the ionized fraction. The uniform photoionization rate implied in this model is $\Gamma = R_{\rm u}\Delta_{\rm i}/2$.

Armed with equation (3), we can now see how the reionization should proceed. Initially, photons will be used up to reionize the low density gas, and recombinations will be negligible (unless $R_{\rm u} \gg 1$). Therefore, the term ${\rm d}F_M(\Delta_{\rm i})/(H\,{\rm d}t)$ dominates. As the emissivity increases, denser gas is ionized (i.e., $\Delta_{\rm i}$ increases), the mean free path of the photons increases, and the global recombination rate increases. At some point, the global recombination rate should approach the value of the emissivity, and then the term $R(\Delta_{\rm i})$ dominates. Recombinations will eventually dominate as long as $R(\Delta_{\rm i})$ keeps increasing with $\Delta_{\rm i}$, and as long as the mean free path is still smaller than the horizon. When this last condition fails, then the terms accounting for the change in the intensity and the redshift will dominate in equation (3). Observations of the abundance of Lyman limit systems have shown that the latter stage is reached in the universe at $z \sim 2$ (e.g., Haardt & Madau 1996 and references therein).

Our discussion here has assumed that the sources are numerous enough, so that the H II regions they create will not reach sizes much larger than the scale of the non-linear structure in the IGM before they overlap. The main difference in the case where the reionizing sources are highly luminous and sparse is that the epoch when the intensity of the background fluctuates by large factors is extended up to a later stage. We can model this case by incorporating the new variable $Q_{\rm i}$, the fraction of the volume in the IGM subject to the radiation from any source, or the fraction of the volume filled by H II regions. In the volume fraction $Q_{\rm i}$, some gas may still be neutral because of its high density, whereas in the fraction $1 - Q_{\rm i}$ all the gas is neutral. Just as before, we can use the approximation where gas with density $\Delta < \Delta_{\rm i}$ is ionized in the fraction $Q_{\rm i}$ of the volume, and the higher gas density is neutral. Equation (5) is then modified to

$$\frac{\epsilon}{Q_{\rm i}} = \frac{F_M}{Q_{\rm i}} \frac{{\rm d}Q_{\rm i}}{H\,{\rm d}t} + {\rm d}F_M(\Delta_{\rm i})/(H\,{\rm d}t) + R(\Delta_{\rm i}) \ . \tag{8}$$

Notice that $\epsilon/Q_{\rm i}$ is simply the emissivity of sources averaged over the fraction $Q_{\rm i}$ of the volume only. As before, there is a first epoch where the two terms involving the variation of $Q_{\rm i}$ and F_M are dominant, and at some point recombinations start dominating globally to balance the emission. When $Q_{\rm i}$ is still less than unity at this point, we can justify the use of equation (2) for $R(\Delta_{\rm i})$ in the following way. In every H II region around individual sources, the radiation along every direction ionizes gas up to a point of high enough density where atoms recombine at the same rate at which the photons are arriving. If $R(\Delta_{\rm i})$ continuously increases with $\Delta_{\rm i}$, most of the recombinations take place at high densities, and therefore most of the photons are absorbed in a dense region at the end of every beam. These regions should have a typical density $\Delta_{\rm i}$ which is still determined by a balance between the global rates of recombination and emission. There should of course be a significant range of densities of the gas in the boundary of the H II regions, since some of the gas with $\Delta > \Delta_{\rm i}$ will be ionized when it is close to a luminous source; this implies that the model for the global recombination rate in equation (2) can only considered as a crude approximation, especially before the epoch of overlap. Nevertheless, equation (8) still makes apparent the fact that there should be no sudden transition when $Q_{\rm i}$ reaches unity because the mean free path of the photons is always controlled by the parameter $\Delta_{\rm i}$.

2.1. The Gas Density Distribution

To examine more quantitatively the effects discussed above, we shall use a model of the gas density distribution in the IGM based on the results obtained from hydrodynamic numerical simulations, which were found to be in good agreement with observations of the distribution of the transmitted flux in the Lyα forest (Rauch et al. 1997). Here, we shall obtain a simple analytical fit to the numerical results of the density distribution in the IGM at different redshifts, which provides an interpretation of the results and greatly facilitates the use of our model.

The formula we use to fit the result of numerical simulations is motivated by a simple approximation of the evolution of the density in voids. In the limit of low densities, we assume the gas in voids to be expanding at a constant velocity, as should be the case if tidal forces are negligible. Then, the proper density in the void decreases as $\Delta a^{-3} \propto t^{-3} \propto a^{-9/2}$, so $\Delta \propto a^{-3/2}$. Assuming that Δ starts decreasing according to this law when the linear overdensity δ reaches a fixed critical value, and using $\delta \propto a$, we obtain $\Delta \propto (-\delta)^{-3/2}$. The linear density δ is assumed to be smoothed on the Jeans length of the photoionized gas. For Gaussian initial conditions, where $P_V(\delta) \propto \exp[-\delta^2/(2\sigma^2)]$ (with σ being the rms density fluctuation), this yields $P_V(\Delta) \propto \exp(-C\Delta^{-4/3})\, \Delta^{-8/3}\, \mathrm{d}\Delta$, where C is a constant. We shall use the following fitting formula, which approaches this limit for $\Delta \ll 1$, and is also applicable in the linear regime when $|\Delta - 1| \ll 1$,

$$P_V(\Delta)\, \mathrm{d}\Delta = A \exp\left[-\frac{\left(\Delta^{-2/3} - C_0\right)^2}{2\,(2\delta_0/3)^2} \right]\, \Delta^{-\beta}\, \mathrm{d}\Delta\ . \tag{9}$$

When $\delta_0 \ll 1$ and $C_0 = 1$, the distribution approaches a Gaussian in $\Delta - 1$ with dispersion δ_0. When $\delta_0 \gg 1$, the distribution goes to the previous approximation for voids, with the peak at a density $\Delta \sim \delta_0^{-3/2}$. Thus, if the median density in voids evolves as $\Delta \propto a^{-3/2}$, then δ_0 should grow proportionally to a both in the linear and the highly non-linear regime. At $\Delta \gg 1$ (and for $\delta_0 \gg 1$), the distribution is a power-law, corresponding to power-law density profiles $\Delta \propto r^{-3/(\beta-1)}$ in collapsed objects (where r is the radius).

We show in Figure 1 the gas density distribution of the L10 simulation described in Miralda-Escudé et al. (1996), together with the fits obtained to equation (9), at the three redshifts $z = 2$, 3 and 4. The two functions shown at each redshift, for the simulations and the result of the fit, are $\Delta P_V(\Delta)$ and $\Delta^2 P_V(\Delta)$, which are the volume-weighted and mass-weighted probability density of Δ per unit $\log \Delta$, respectively. This density distribution was previously shown in Figure 7 of Rauch et al. (1997) at $z = 2$ only, together with the result from a simulation by Hernquist et al. (1996). As seen in that paper, the density distribution in these two simulations is very similar, since they used models with very similar power spectra. The density distribution should depend mostly on a single parameter, the amplitude of the initial density fluctuations smoothed on the Jeans scale of the photoionized gas. Table 1 gives the values of the four parameters of the fits at each redshift. The parameter δ_0 is indeed approximately proportional to the scale factor.

We have extrapolated the density distribution to higher redshifts by using $\delta_0 = 7.61/(1+z)$ and $\beta = 2.5$, the latter corresponding to an isothermal slope for high density halos. This reproduces the fits in Table 1 to better than 1%. The change of β with redshift is more uncertain, since we do not have a clear model to predict its evolution (the form of the high-density tail of the distribution in the simulations may also be affected by the resolution). The parameters A and Δ_0 can then be fixed by requiring the total volume and mass to be normalized to unity.

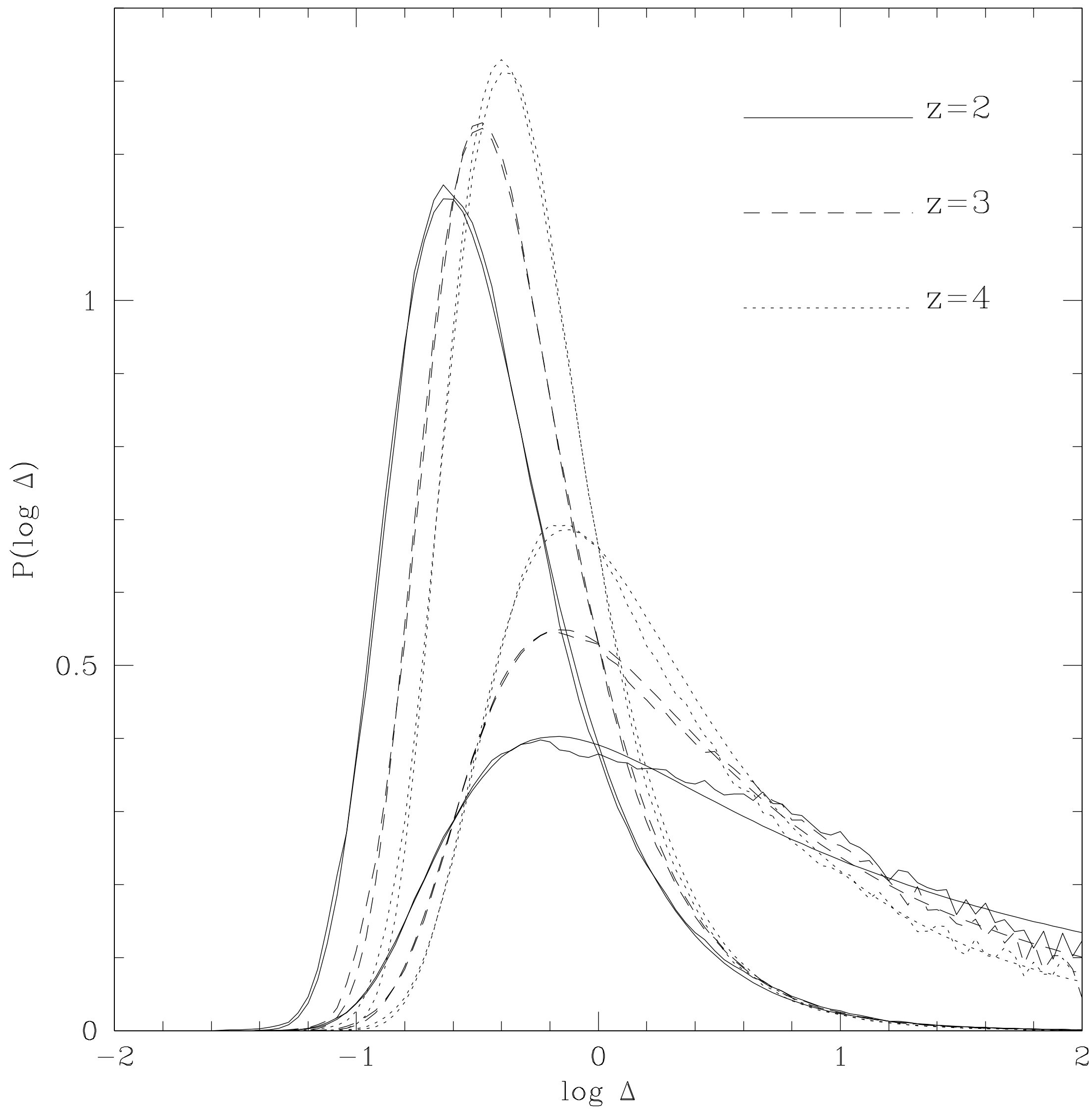

Fig. 1.— Mass and volume-weighted differential density distribution of the IGM, at the three indicated redshits. The lines with noise are from the numerical simulation in MCOR, and the smooth lines are the fit we have obtained with the analytical model in equation (9).

We note here the possibility that a substantial fraction of all baryons is contained in small halos with virial temperatures $\sim 10^4$ K, which collapsed to high densities before reionization when the Jeans mass was very small (Abel & Mo 1998). These halos could survive for a long time after reionization occurs if neither star formation nor tidal and ram-pressure stripping due to mergers into larger structures are able to destroy them. In this case, the fraction of baryons at high densities could be higher than in our model.

2.2. The Global Recombination Rate

In Figures 2(a,b,c), the solid and dotted lines show the cumulative probability distribution of the gas density weighted by mass and volume, respectively, at redshifts $z = 2, 3, 4$. These are obtained from equation (9), with the fit parameters in Table 1. Figure 2d shows the same cumulative distributions obtained at $z = 6$, according to the prescription described for choosing the parameters in equation (9); the values of these parameters at $z = 6$ are also given in Table 1. The dashed and dash-dot lines show the ratio $R/R_{\rm u}$ as obtained from equations (2) and (6), respectively (the former is the case of a sudden transition from fully ionized to fully neutral gas at $\Delta_{\rm i}$, while the latter assumes a uniform intensity of the ionizing background). For the density distribution we have adopted, R can grow to very large values as $\Delta_{\rm i}$ increases; this is of course true only as long as $P_V(\Delta)$ is less steep than Δ^{-3}.

In the case where the sources of radiation are sufficiently numerous, the H II regions should overlap soon after all the voids, with $\Delta < 1$, have been ionized. As an example, at the time when $\Delta_{\rm i} = 3$, only about half the baryons need to have been ionized, but 95% of the volume is ionized. These fractions do not change very much at the different redshifts used in Figures 2(a,b,c,d). The dashed curves also show that $R/R_{\rm u} = 1$ at $\Delta_{\rm i} \sim 3$. If this level of ionization is reached at $z = 6$, where $R_{\rm u} \simeq 1$, then sources need to be emitting about one ionizing photon per Hubble time in order to compensate for the recombinations and continue to increase $\Delta_{\rm i}$. This simple case exemplifies that the clumpiness of the gas

Table 1

Redshift	A	δ_0	β	C_0
2	0.406	2.54	2.23	0.558
3	0.558	1.89	2.35	0.599
4	0.711	1.53	2.48	0.611
6	0.864	1.09	2.50	0.880

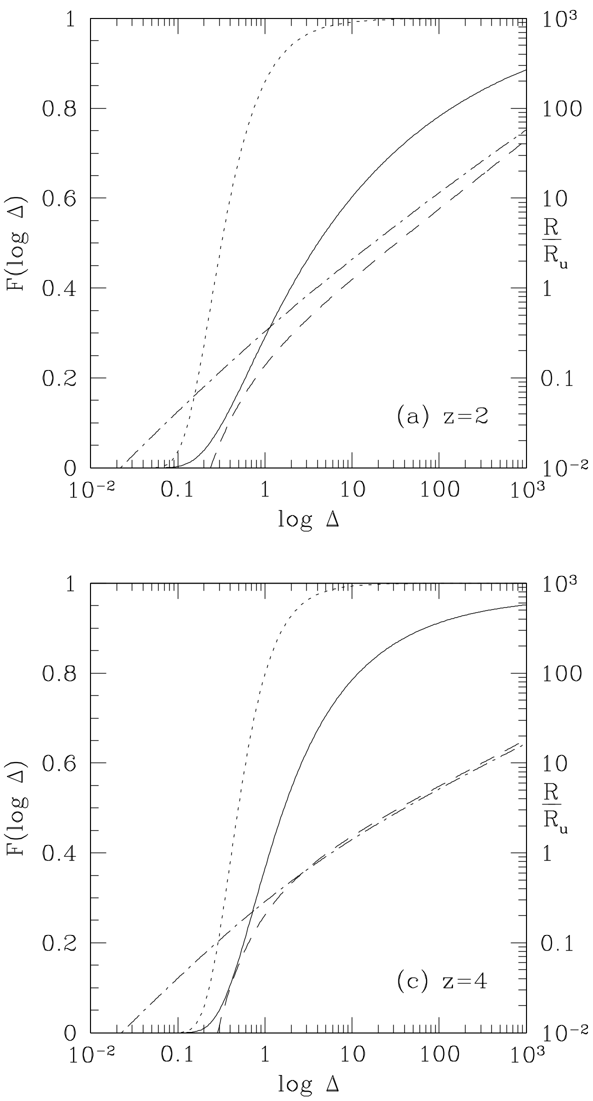

Fig. 2.— Solid and dotted lines are the cumulative density distribution weighted by mass and volume, respectively. The dashed line gives the global recombination rate R when all the gas is assumed to be neutral at overdensities greater than Δ, and ionized at lower overdensities. The dash-dot line is the global recombination rate when ionization equilibrium with a uniform ionizing background is assumed, with the neutral fraction being one half at overdensity Δ. The values of $R_{\rm u}$ at the redshifts of the four panels, $z = 2, 3, 4, 6$, are $R_{\rm u} = 0.32$, 0.50, 0.71, 1.18 for hydrogen, and $R_{\rm u} = 1.89$, 2.98, 4.19, 6.99 for helium (a gas temperature of 10^4 K has been assumed).

does not necessarily increase the number of photons needed to reionize the universe. In fact, when most of the gas is in very dense clumps with a small covering factor over a path of the order of the mean separation between sources, the required number of photons *decreases* because the highly dense clumps can stay neutral until the mean free path increases to a value much larger than the size of the H II regions before overlap.

3. The Lyman Alpha Flux Decrement

We now address the question of the flux decrement that should be observed to the blue of the Lyα wavelength, as the redshift increases toward the epoch of reionization. The optical depth of a uniform, completely neutral IGM, τ_0, is extremely large,

$$\tau_0 = 2.6 \times 10^5 \, [\Omega_{\rm b} h(1-Y)/0.03] \, [H_0(1+z)^{3/2}/H(z)] \, [(1+z)/7]^{3/2} \ . \tag{10}$$

Because of this, a small neutral fraction is sufficient to yield a high enough optical depth to make the transmitted flux undetectable, so the Lyα spectrum might already be completely blocked at redshifts well after the epoch when most of the volume in the IGM becomes ionized.

In the reionization model described in the previous section, regions of higher gas density become progressively ionized as the emissivity increases. In order to calculate the Lyα optical depth through an ionized region, we need to know the intensity of the ionizing radiation. The intensity increases as the reionization proceeds owing to the increase of the emissivity, and also to the increase of the mean free path of the photons. When the mean free path $\lambda_{\rm i}$ is much smaller than the horizon, the intensity depends only on the instantaneous emissivity:

$$n_J = \epsilon \, \frac{H \, \lambda_{\rm i}}{c} \ . \tag{11}$$

In this approximation, we have $n_J \ll \epsilon$, and equation (5) is valid.

3.1. The mean free path of the ionizing photons

In the simple model where all the gas at densities $\Delta > \Delta_{\rm i}$ is still neutral, and all the gas at lower densities is ionized, a photon is absorbed whenever it enters a region with $\Delta > \Delta_{\rm i}$, assuming that these regions are sufficiently large to be optically thick. The mean free path $\lambda_{\rm i}$ is then simply the mean length of regions with $\Delta < \Delta_{\rm i}$ along random lines of sight. Of course, in reality there should be a gradual change of ionization, and differences in the geometrical shape of the structures will cause varying degrees of self-shielding at a given gas density, so photons will be absorbed over a range of densities. Nevertheless, since most of the recombinations take place at densities near $\Delta_{\rm i}$, most of the photons will be absorbed by gas near this density. The spacing between successive contours at $\Delta_{\rm i}$ should then provide a reasonable estimate of the mean free path. Notice that we assume that $\lambda_{\rm i}$ is larger than the mean free path for the case of a uniform IGM, $\lambda_{\rm u} = [n_{\rm a}\bar{\sigma}(1-F_M)]^{-1}$, where $n_{\rm a}$ is the density of atoms, $\bar{\sigma}$ is a frequency-averaged photoionization cross section, and F_M is the fraction of mass in gas with $\Delta < \Delta_{\rm i}$. If $\lambda_{\rm u} > \lambda_{\rm i}$, then the neutral clumps are optically thin and the mean free path is close to $\lambda_{\rm u}$ independently of the density structure of the IGM. We also emphasize that $\lambda_{\rm i}$ is only the mean free path for photons that escape into the IGM. Sources of photons will generally appear in regions of high density, and some fraction of the photons will be absorbed locally; these photons are not counted in the mean emissivity ϵ in equations (3) and (5).

Here, we shall use a very simple model for the mean free path of ionizing photons: we assume

$$\lambda_{\rm i} = \lambda_0\,[1 - F_V(\Delta_{\rm i})]^{-2/3} \ . \tag{12}$$

In the limit of high densities, when $1 - F_V \ll 1$, this is valid for any population of absorbers where the number density and shape of isolated density contours remains constant as $\Delta_{\rm i}$ is varied: the fraction of volume filled by the high density regions is $1 - F_V$, so their size is proportional to $(1 - F_V)^{1/3}$, and the separation between them along a random line of sight is proportional to $(1 - F_V)^{-2/3}$. The shape of the absorbers in large-scale structure theories is actually highly complex, with structures varying from ellipsoidal to filamentary to sheet-like as $\Delta_{\rm i}$ is reduced. However, we have checked that the proportionality $\lambda_i \propto (1 - F_V)^{-2/3}$ is still approximately obeyed in the results of numerical simulations with photoionized gas dynamics, for $\Delta_{\rm i} \gtrsim 1$ (see e.g. Figure 3 of MCOR).

In this paper, we shall use the scale $\lambda_0\,H = 60\,{\rm km\,s}^{-1}$, which reproduces the scales of the Lyα forest structures in the simulation of the cold dark matter model with a cosmological constant at redshift $z = 3$ in MCOR (designated as L10 simulation in that paper). Although the scale λ_0 should vary with redshift and with the cosmological model, the quantity $\lambda_0\,H$ stays roughly constant in this simulation. In fact, $\lambda_0\,H$ is basically determined by the Jeans length of photoionized gas at high redshift, when the amplitude of density fluctuations reaches non-linearity at the Jeans scale, and it should probably increase at low redshifts as structures collapse on larger scales. This increase might not be fully reflected in numerical simulations like those in MCOR owing to the small size of the simulated box.

3.2. Optical depth of the ionized IGM

The next step now is to compute the optical depth to Lyα scattering through the ionized IGM, $\tau_{\rm i}$, in a region of density Δ. This is equal to $\tau_0 \Delta$ (eq. 10), times the neutral fraction, which we assume to be in ionization equilibrium. For a constant gas temperature, this optical depth is

$$\tau_{\rm i} = \tau_0 \frac{\alpha}{\bar\sigma c n_J} \frac{\Delta^2}{f_J} ~, \tag{13}$$

where α is the recombination coefficient (we assume full ionization and neglect the effect of double helium ionization on the electron density), $\bar\sigma$ is the frequency-averaged photoionization cross section, and we have defined f_J to be the ratio of the local photon density of the ionizing background at a given point in the IGM to the mean photon density n_J, given by equation (11). The local intensity fluctuates due to the discreteness of the sources that can illuminate a given point. Before the H II regions around individual sources overlap, the intensity fluctuations are of course very large, but they are rapidly reduced after the mean free path increases to values larger than the typical separation between neighboring sources (Zuo 1992).

Substituting equations (11) and (12) into (13), we obtain for the optical depth of an ionized region,

$$\tau_{\rm i} = \tau_0 \, \frac{\alpha}{\bar\sigma c} \, \frac{c}{H \lambda_0} \, (1 - F_V)^{2/3} \, \frac{\Delta^2}{\epsilon f_J} ~. \tag{14}$$

Notice that F_V depends on $\Delta_{\rm i}$, determining the maximum densities that are ionized, while Δ is the density in the region yielding an optical depth $\tau_{\rm i}$. The mean emissivity ϵ can be obtained from equation (5). The term $\mathrm{d}F_M/(H\,\mathrm{d}t)$ in equation (5) obviously depends on the model for the emitting sources. However, if the sources of ionizing photons do not evolve in a synchronous fashion on timescales much shorter than the Hubble time, then this term should be of order $\mathrm{d}F_M/(H\ \mathrm{d}t) \sim 1$. Here we shall simply use the expression $\epsilon = 1 + R$; the important point is that $\epsilon = R$ when recombination dominates, and $\epsilon \sim 1$ when most of the baryons are being ionized if recombinations are not important. Thus, we have

$$\tau_{\rm i} = 1.73 \, \frac{\Omega_{\rm b} h (1-Y)}{0.03} \, \frac{H_0 (1+z)^{3/2}}{H(z)} \left(\frac{1+z}{7} \right)^{3/2} \frac{c}{H\lambda_0} \frac{(1-F_V)^{2/3}}{1 - F_M + R} \, \frac{\Delta^2}{f_J} \equiv \tau_{\rm u} \, \frac{\Delta^2}{f_J} ~. \tag{15}$$

We have used a value $\bar\sigma = 2 \times 10^{-18}\,\mathrm{cm}^{-2}$, which should be approximately valid for the spectra emitted by quasars or star-forming galaxies. The above equation can be reexpressed in terms of the recombination rate for a uniform IGM, $R_{\rm u}$,

$$\tau_{\rm u} = \frac{\tau_0}{\bar\sigma \, n_{\rm e} \, \lambda_0} \, \frac{(1 - F_V)^{2/3}}{(1+R)/R_{\rm u}} = 1.14 \, \frac{c}{H \lambda_0} \, \frac{(1 - F_V)^{2/3}}{(1+R)/R_{\rm u}} ~, \tag{16}$$

where n_e is the electron density.

The value of τ_u as a function of Δ_i is shown in Figure 3, at redshifts $z = 2, 3$ and 4. We have used $c/(H\lambda_0) = 5000$, and the cosmological model mentioned in the introduction. As Δ_i increases, the larger mean free path of the photons results in an increasing intensity of the ionizing background, and that decreases the optical depth of the ionized regions. The variation of $\tau_u(\Delta_i)$ with redshift can be understood as follows: if we ignore the variation of the density distribution F_V with redshift, and for large Δ_i, we have $R \gg 1$, and τ_u is independent of redshift from equation (16). For small Δ_i, τ_u increases with redshift since R is small compared to unity. The small variation of τ_u with redshift at large Δ_i is due to the variation of the density distribution.

If the IGM were homogeneous, the mean flux decrement due to Lyα scattering would simply be $e^{-\tau_u}$, but for the real IGM the dependence of the flux decrement on τ_u is of course more complicated. The mean flux decrement of the Lyα forest obtained from the simulation in MCOR is shown as a function of τ_u in Figure 4, at redshifts $z = 2, 3, 4$. These curves were computed directly from the simulation, and therefore include the effects of thermal broadening and peculiar velocity; they were shown also in Fig. 19 in MCOR (although with a different scaling in the horizontal axis). Each curve starts at the point where the flux decrement is equal to the observed value (see Press, Rybicki, & Schneider 1993; Rauch et al. 1997). Therefore, in the MCOR model, if all the baryons were spread out uniformly in the IGM keeping the intensity of the ionizing background fixed, the optical depth to Lyα scattering would be $\tau_u = (0.3, 1, 4)$ at $z = (2, 3, 4)$. The curves show how the flux decrement would increase at a fixed redshift if the intensity of the background was reduced, increasing the value of τ_u. At fixed τ_u, the flux decrement increases slightly with redshift. The reason is that voids become more underdense with time, and therefore the amount of flux that is transmitted through the most underdense regions for high τ_u decreases.

3.3. What is the redshift where the Gunn-Peterson trough is reached?

As sources at progressively higher redshift are being discovered, one of the most interesting questions that arise is: what is the redshift at which a complete Gunn-Peterson trough will first appear? It is often assumed that the Gunn-Peterson trough should appear when the epoch of reionization is reached, and a region of the IGM that is still mostly neutral lies on the line of sight to a source. Actually, the only requirement to produce a complete Gunn-Peterson trough is that the IGM is opaque to Lyα photons everywhere, and given the enormous value of the optical depth for a neutral medium (eq. 10), a very small neutral fraction can reduce the transmitted flux to negligible levels. In ionization

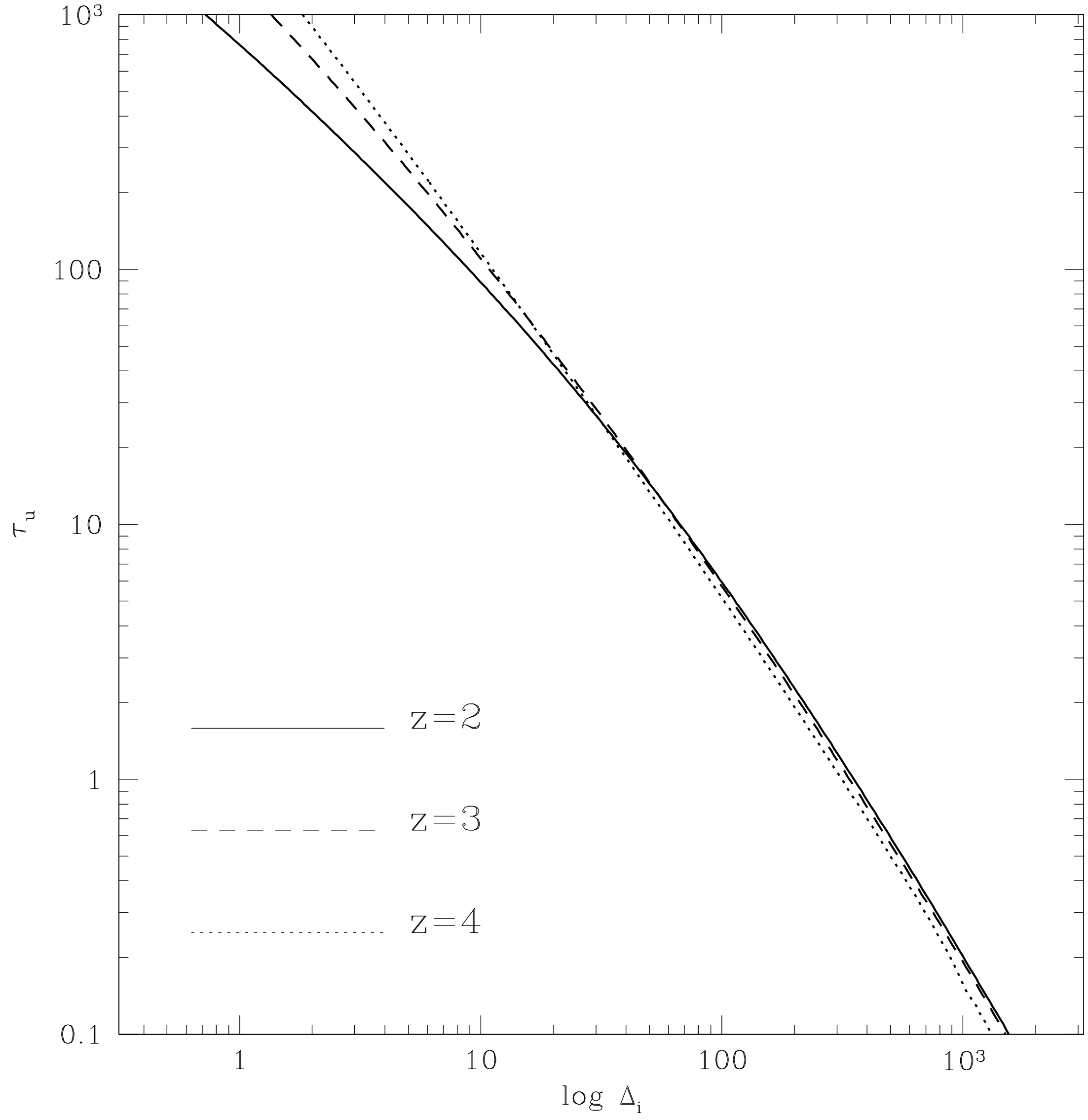

Fig. 3.— Optical depth τ_u that a uniform IGM would yield for the intensity of the ionizing background derived when the emissivity and the mean free path are given by the global recombination rate and our mean free path model, as a function of the overdensity to which the inhomogeneous IGM is ionized.

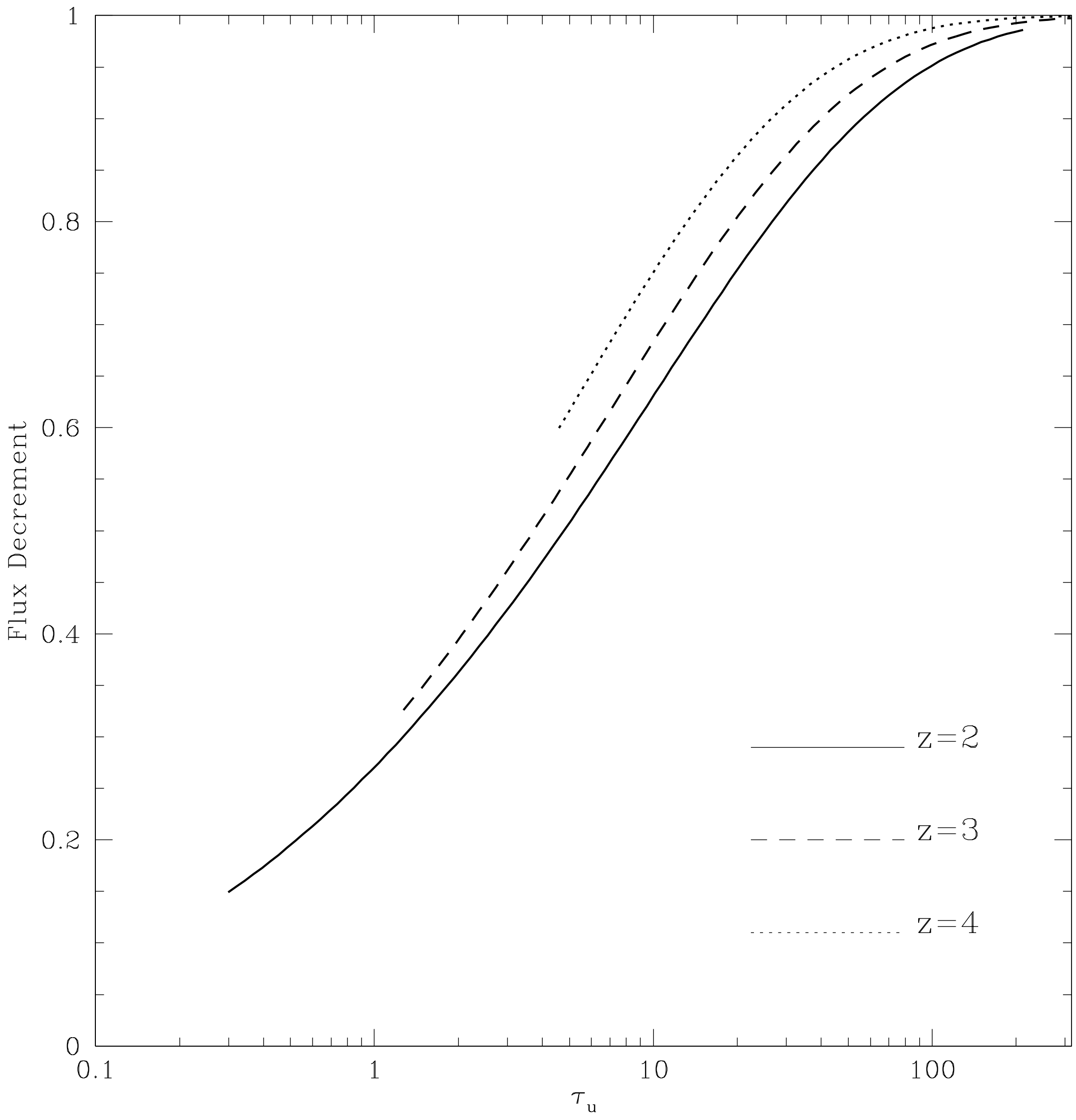

Fig. 4.— Mean flux decrement as a function of the optical depth that a uniform IGM with the same mean density and intensity of the ionizing background would have.

equilibrium with a uniform background, the most underdense voids will have the lowest optical depths. For the gas density distribution used in §2, the lowest densities are $\Delta \simeq 0.1$ at $z = 2$ and $\Delta \simeq 0.25$ at $z = 6$. Since the optical depth is proportional to Δ^2, the transmitted flux ought to decrease very rapidly to negligible levels when $\tau_{\rm u}$ reaches the inverse square of the lowest underdensities in voids, as the curves in Figure 4 show (the peculiar velocity effect in voids helps to allow for some transmitted flux at slightly higher values of $\tau_{\rm u}$).

The distribution of $\tau_{\rm i}/\tau_{\rm u}$, where $\tau_{\rm i}$ is the optical depth at a pixel in a well resolved Lyα spectrum, was plotted in Figure 11 of MCOR (assuming a uniform intensity of the background radiation). The lowest value of this ratio is indeed of order the inverse square of the lowest underdensities. Thus, in the model of MCOR, at $z = 3$ only 2% of the pixels in a well resolved Lyα spectrum should have $\tau_{\rm i}/\tau_{\rm u} < 0.01$, and at $z = 4$ only 2% have $\tau_{\rm i}/\tau_{\rm u} < 0.02$, reflecting the decrease of the underdensities in voids with time. For example, for the value $\tau_{\rm u} = 4$ at $z = 4$ that agrees with the observed flux decrement, the lowest optical depths in a Lyα spectrum are predicted to be about 0.08.

Figure 5 shows the same flux decrement as in Figure 4, plotted as a function of $\Delta_{\rm i}$. The large filled squares are the values of the observed transmitted flux, according to Rauch et al. (1997). The figure therefore provides a prediction for $\Delta_{\rm i}$ as a function of redshift, given our assumed model for the density distribution: the gas in the Lyman limit systems at column densities $\sim 10^{18}\,{\rm cm}^{-2}$, where the transition from ionized to neutral gas takes place due to self-shielding, should have typical overdensities of order $\Delta_{\rm i} \sim (700, 300, 100)$ at $z = (2, 3, 4)$. Notice, however, that the approximation of neglecting redshift effects in equation (3) is starting to fail at $z = 2$, so $\Delta_{\rm i}$ should be somewhat underestimated.

It is clear from Figure 5 that, if the trend of decreasing $\Delta_{\rm i}$ with redshift continues at $z > 4$, then the Gunn-Peterson trough should appear at $z \simeq 6$. There are two main reasons for the increase of the flux decrement with redshift. The most important is the decrease of the density $\Delta_{\rm i}$ to which the universe is ionized. The second one is that the flux decrement increases for fixed $\Delta_{\rm i}$ owing to the lower underdensities of voids at high redshift; thus, at $z \simeq 6$ the transmitted flux should drop to 1% at $\Delta_{\rm i} \simeq 15$ (notice from Figure 3 that $\tau_{\rm u}$ remains approximately constant with redshift at fixed $\Delta_{\rm i}$).

Figures 2(a,b,c) shows that the global recombination rate corresponding to the densities of Lyman limit systems mentioned above, $\Delta_{\rm i}$, at $z = (2, 3, 4)$, are $R/R_{\rm u} = (40, 15, 6)$, corresponding to an emissivity of $\epsilon = R = (14, 8.1, 4.3)$ photons per baryon per Hubble time. If the trend of decreasing emissivity with redshift continues, clearly the Gunn-Peterson trough will be reached at $z \simeq 6$. In fact, in order to have more than a few percent of transmitted flux at $z = 6$, according to Figures 5 and 2d, we need $\Delta_{\rm i} \gtrsim 20$ and $\epsilon > 2.5$.

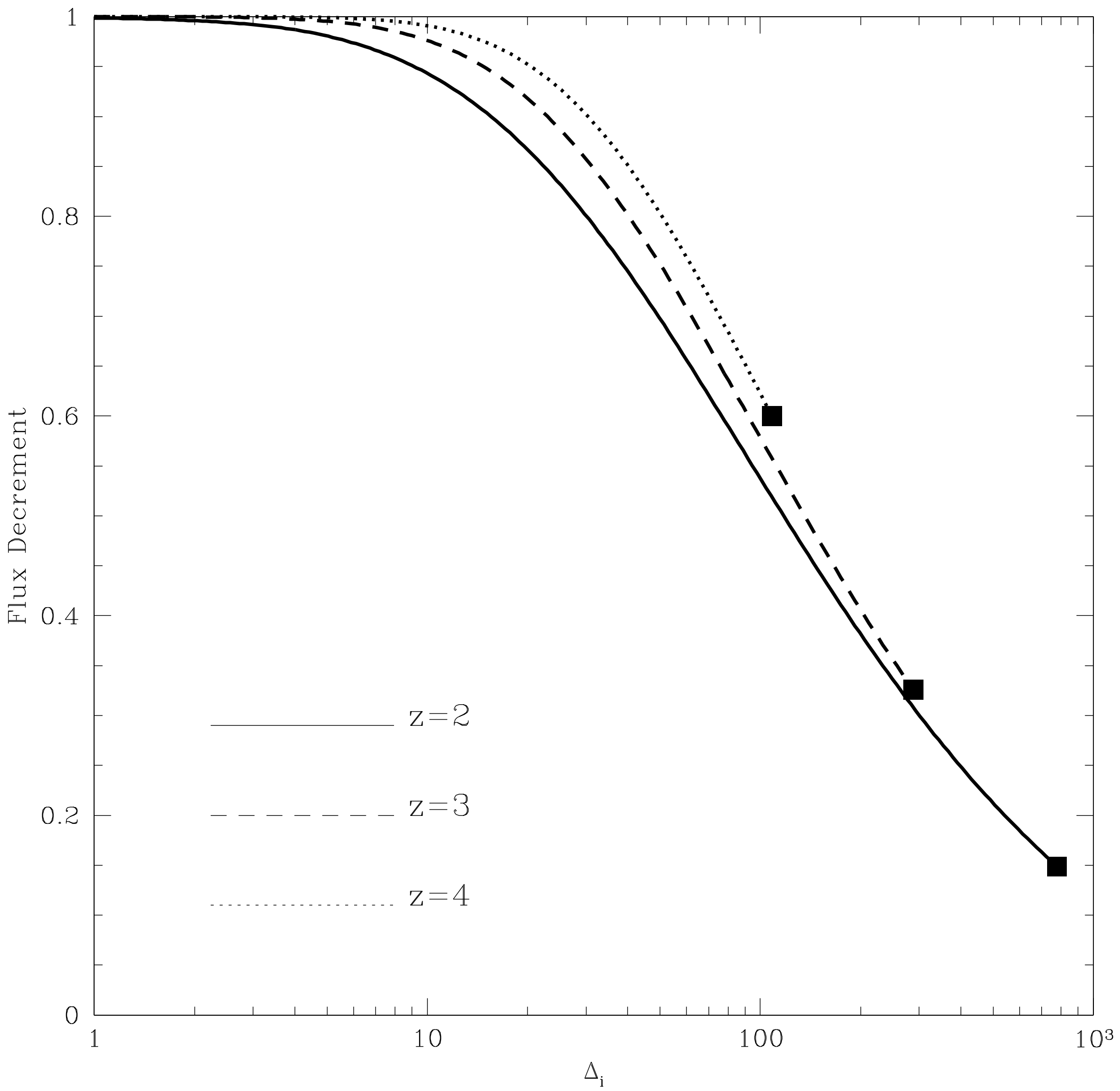

Fig. 5.— Mean flux decrement as a function of the overdensity up to which the gas is ionized. Solid squares indicate the observed value of the mean flux decrement according to Rauch et al. (1997) at $z = 2, 3, 4$.

Since we defined ϵ as the number of photons emitted per Hubble time, then as long as the emissivity per physical time is not larger at $z = 6$ compared to $z = 4$, the transmitted flux at $z = 6$ should be less than $\sim 3\%$. We emphasize that the reason is not because the epoch of reionization is reached at this redshift, but because even the most underdense voids have a large enough density of neutral hydrogen to scatter essentially all the Lyα photons from a background source.

Nevertheless, the redshift at which the Gunn-Peterson trough is first seen might be higher if more sources were present at high redshift. For example, if the emissivity remained at $\epsilon \simeq 5$, then $\Delta_{\rm i} \simeq 100$ at $z = 6$ and the flux decrement would not decline very fast at $z > 4$. In fact, because less mass has collapsed to high densities at high redshift, the value of $\Delta_{\rm i}$ (and therefore the flux decrement) becomes extremely sensitive to R (Fig. 2d). However, our model of the density distribution does not take into account that dense gas could be present in low-mass halos that collapsed before the IGM was ionized, on scales below the Jeans scale of the photoionized gas.

The observations of the highest redshift sources (Dey et al. 1998, Weymann et al. 1998) suggest so far that the flux decrement rapidly declines, indicating that the decrease of $\Delta_{\rm i}$ with redshift probably continues along the same trend shown in Figure 5 at higher redshifts.

3.4. The mean free path and the abundance of Lyman limit systems

Figure 6 shows the mean free path λ_i as a function of $\Delta_{\rm i}$ that we derive in our model. For the values of $\Delta_{\rm i}$ obtained previously to match the observed flux decrement, the mean free path at $z = (2, 3, 4)$ is $\lambda_i = (50, 33, 18) \times 10^3\,{\rm km\,s^{-1}}$, which is in good agreement with the number of observed Lyman limit systems, $\sim 2[(1 + z)/4]^{1.5}$ per unit redshift (e.g., Storrie-Lombardi et al. 1996). We notice that in the MCOR simulation, the number of Lyman limit systems was much less than observed. The reason we obtain a good match here is because we have fitted the high-density tail of the density distribution with a power-law that is close to an isothermal slope for gas distributed in halos, and we have also used a power-law dependence of λ_i on $\Delta_{\rm i}$ that assumes the presence of halos with density cusps. The numerical simulation in MCOR contains instead halos of gas with core sizes limited by the resolution, which explains why the number of Lyman limit systems in the simulations was smaller than in our model.

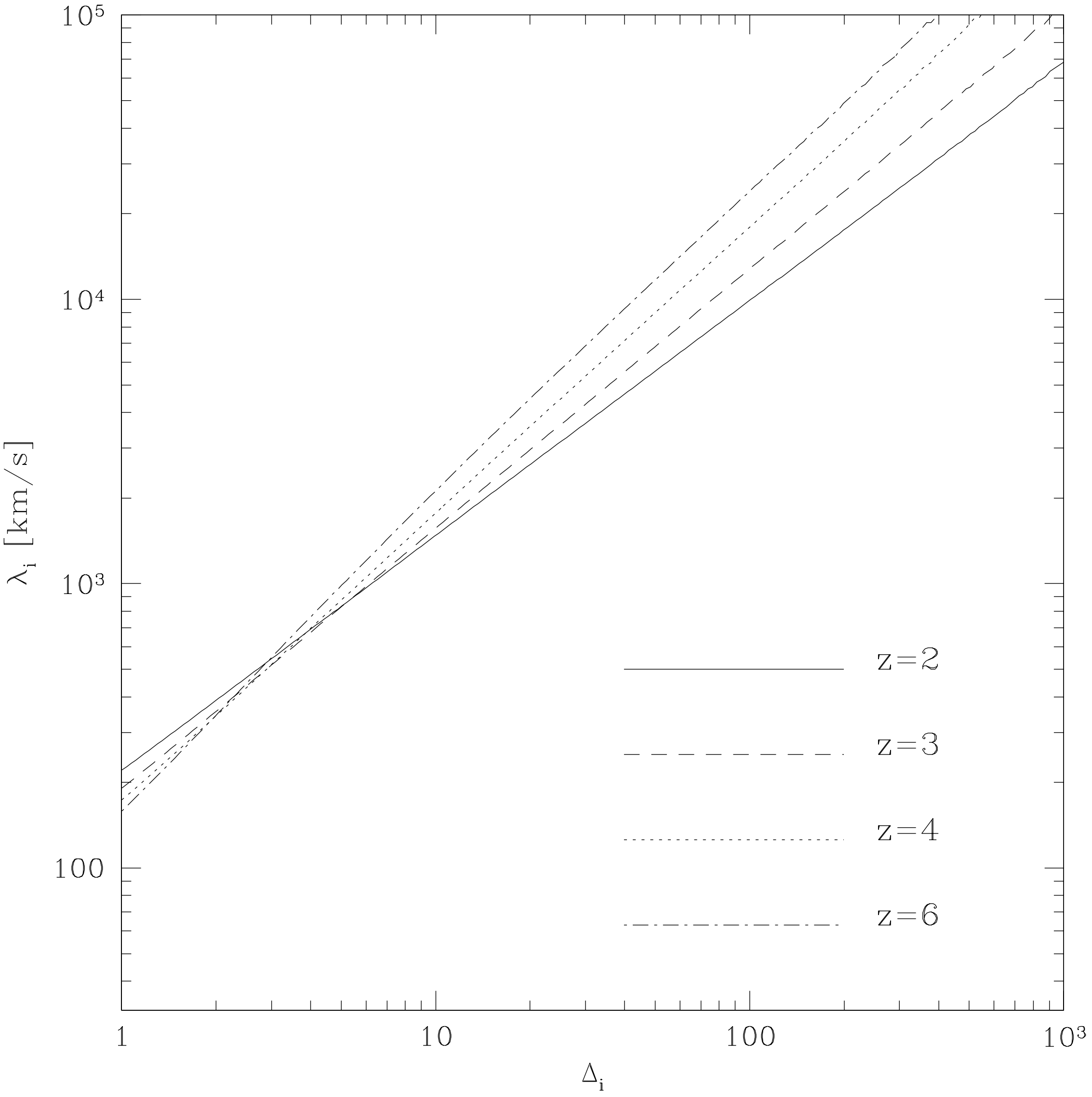

Fig. 6.— Mean free path of ionizing photons as a function of the overdensity up to which the gas is ionized, according to equation (12), at the indicated redshifts.

3.5. Gaps in the Gunn-Peterson trough due to individual H II regions

Is it possible that some of the last gaps of transmitted flux appearing in Lyα spectra will correspond to individual H II regions around the sources that reionized the IGM before the epoch of overlap? As pointed out in Miralda-Escudé (1998, hereafter M98), there are two conditions that need to be satisfied for an ionized region to transmit flux. The first condition is the one we have also discussed here: the optical depth through the ionized region should not be much higher than unity. Given the discussion at the end of §2, the formalism developed here to compute the typical value of $\tau_{\rm i}$ from equation (15) should be equally valid before the epoch of overlap, in H II regions around individual sources, taking into account the following differences: (a) The mean emissivity is now substituted by $\epsilon/Q_{\rm i}$, averaging only over the ionized volume of the universe. (b) The mean free path λ_i should be identified with the radius of the H II region. (c) The fluctuations of the intensity should be much higher than after the epoch of overlap, so when the line of sight passes near a luminous source τ_i may be decreased owing to a high value of f_J in equation (15).

From Figure 5, we see that the transmitted flux becomes undetectably small when the optical of a uniform IGM under the same intensity of the ionizing background is $\tau_{\rm u} \gtrsim 50$. For our model of the mean free path, this corresponds to $\Delta_{\rm i} \lesssim 20$, and $H(z)\lambda_i \lesssim 4000\,{\rm km\,s^{-1}}$ near $z = 6$. This condition is equally applicable to individual H II regions: in order for an H II region to be visible in the Lyα spectrum as a gap in the Gunn-Peterson trough, the mean intensity of radiation in the H II region should be such that $\tau_{\rm u} \lesssim 50$, which would allow the H II region to reach a "Strömgren radius" equal to $\lambda_{\rm i}$, ionizing the gas up to overdensities $\Delta_{\rm i} \sim 20$ for our value of the mean free path. A gap in the transmitted flux will then be observed when the line of sight crosses a void in the H II region with a typical background intensity given by the luminosity of the source and the size of the ionized region. Smaller H II regions could be seen only rarely if the line of sight happens to pass through a void close to the source (so that $f_J \gg 1$). If the individual H II regions around the sources of radiation that reionized the IGM had radii $H(z)\lambda_i \lesssim f_J^{-1} 4000\,{\rm km\,s^{-1}}$, then the epoch of reionization would not be directly observed in Lyα spectra because the intensity of ionizing radiation in the H II regions should be too weak to reduce the neutral density to sufficiently low values.

The second condition for an isolated H II region to transmit any Lyα photons arises from the fact that if a large fraction of the IGM is still neutral, then the damped absorption profiles of the neutral gas in front and behind the H II region will overlap, scattering the remaining photons that could have crossed an underdense, ionized region. It was shown in M98 that an H II region must have a proper size larger than $1h^{-1}[\Omega_{\rm b}h(1-Y)/0.03]$ Mpc to prevent this overlap of damping wings, and that the optical depth in an H II region of this size is $\tau_u = 475(T/10^4\mathrm{K})^{-0.7}\,\epsilon^{-1}$ (see eq. 4 in M98; the quantity N_r in M98 is $2n_b/3n_e$ times our ϵ, where $n_e/n_b = 1 - Y/2$), independently of the redshift or any other parameters. This condition is generally not as restrictive as the first one, and therefore an H II region large enough to have a low optical depth in voids will also not have its gap through the Gunn-Peterson trough obstructed by damping wings. Thus, the condition $\Delta_i > 20$ implies $\tau_u < 50$ (Figure 3), a stronger requirement than the above one for $\epsilon \simeq R \simeq 2$ at $z = 6$. However, for small H II regions, this second condition implies that even if the gap of transmitted flux could be visible due to $f_J \gg 1$, the flux might still be removed due to the damping wings.

We therefore conclude that whether Lyα photons can be transmitted before the overlap of H II regions is complete depends mainly on the typical luminosity of the sources of ionizing photons. We shall discuss plausible sizes of the H II regions in §4.

This section has presented a general formalism to calculate the evolution of τ_u with redshift, the epoch when the Lyα forest should become a complete Gunn-Peterson trough, and the condition on the size of the individual H II regions required for Lyα flux to be transmitted through them. However, our numerical results depend on the density distribution $P_V(\Delta)$ and the scale of the density structures in the IGM, λ_0, which we have taken as constant in reshift space here. These quantities need to be calculated separately for every cosmological model as a function of redshift.

4. Hydrogen Reionization

4.1. Sources of Ionizing Radiation : QSO's and Galaxies

Among known, observed objects, QSO's and galaxies are the two candidates for the sources of ionizing photons that reionized the IGM. At $z \simeq 3$ the total number of UV photons emitted by galaxies longward of the Lyman break outnumbers those produced by optically bright QSO's by a factor of approximately ten to thirty (Boyle & Terlevich 1998; Haehnelt, Natarajan, & Rees 1998). For photons capable of ionizing hydrogen, the balance is less certain. Observed bright, high-redshift QSO's show no Lyman break, whereas galaxies have a strong intrinsic Lyman break arising from the stellar atmospheres (of a factor ~ 5 for the usual mass function of stars observed in galaxies), and a very uncertain escape fraction of photons beyond the Lyman edge. For starbursts at low redshift, the escape fraction is reported to be of order 10% (Leitherer et al. 1995; Hurwitz, Jelinski & Van Dyke Dixon 1997). A moderately higher value would therefore be required for UV emission by star-forming galaxies to rival that of QSO's at $z = 3$ in terms of ionizing photons.

The solid curves in Figure 7 show a fit to the luminosity function of QSO's at $z = 4$ obtained by Pei (1995): $\mathrm{d}\Phi = \Phi_*/((L/L_*)^{\beta_\mathrm{l}} + (L/L_*)^{\beta_\mathrm{h}})\, \mathrm{d}L/L_*$, with $\Phi_* = 8.9 \times 10^{-7}\,\mathrm{Mpc}^{-3}$, $L_* = 7.8 \times 10^{45} \mathrm{erg\,s}^{-1}$ (B-band) and $\beta_\mathrm{h} = 3.52$. The luminosity function is plotted as follows: on the horizontal axis, instead of the luminosity L, we use the variable $R_\mathrm{HII} = [3N_\mathrm{phot}/(4\pi n_\mathrm{H})]^{1/3}$, where N_phot is the total number of ionizing photons emitted by the source for an assumed lifetime t_Q, n_H is the comoving number density of hydrogen, and R_HII is the comoving radius of the H II region that a source of luminosity L would produce in a completely neutral homogeneous medium if there were no recombinations. We obtain the number of photons from the relation $N_\mathrm{phot} = Lt_\mathrm{Q}/E_\mathrm{ion}$, where E_ion is the mean energy per photon. We have assumed that the QSO's emit twice the luminosity in ionizing photons as in the B-band (a reasonable approximation for the typical observed quasar spectra), and have used $E_\mathrm{ion} = 20\,\mathrm{eV}$. Also shown in the upper horizontal axis is the corresponding expansion velocity at $z = 4$, $H(z = 4)\, R_\mathrm{HII}/(1 + 4)$.

On the vertical axis, the mean comoving volume emissivity of QSOs per unit natural logarithm of R_HII is plotted, in terms of the number of photons emitted per hydrogen atom per Hubble time. The curves in Figure 7 are an "energy function" (or the distribution of total energy emitted by sources), rather than a luminosity function. The energy function depends of course on the assumed lifetime, and we have plotted it for two values: $t_\mathrm{Q} = 3 \times 10^7$ years and $t_\mathrm{Q} = 8 \times 10^8$ years. For other values of t_Q, the curves simply shift horizontally by a factor proportional to $t_\mathrm{Q}^{1/3}$. The time 8×10^8 years is equal to half the age of the universe at $z = 4$; this can be considered as an upper limit for t_Q, because the number of quasars of given luminosity should be rapidly increasing with time at high redshift. All values are computed for our adopted cosmological model mentioned in the introduction.

The faint end slope of the luminosity function is still rather uncertain, so we have plotted three curves for the values $\beta_\mathrm{l} = (1.5, 1.65, 1.8)$, where $\beta_\mathrm{l} = 1.64$ is the value preferred by Pei. The value of the lowest luminosity for which the abundance of quasars has been determined observationally at $z = 4$, $10^{46} \mathrm{erg\,s}^{-1}$, is indicated with an arrow. For lower luminosities, the model assumes that the luminosity function has the same shape as at lower redshift.

The diamond symbols in Figure 7 show in the same way the mean comoving volume emissivity of UV-emitting galaxies. We use a Schechter function fit to the luminosity function at $z = 3$ (Dickinson et al. 1998), $\mathrm{d}\Phi = \Phi_*\,(L/L_*)^\beta \exp{(-L/L_*)}\, \mathrm{d}L/L_*$, with $\Phi_* = 1.6 \times 10^{-3}\,\mathrm{Mpc}^{-3}$, $L_* = 1.4 \times 10^{44} \mathrm{erg\,s}^{-1}$ (defined as $\nu L_{*\nu}$ at 1500 Å), and $\beta = -1.5$. The luminosity function of UV-emitting galaxies does not evolve strongly between $z = 3$ and $z = 4$ (Steidel et al. 1998). The ionizing luminosity is assumed to be a fraction 0.06 of νL_ν at $1500\mathring{A}$, due to an intrinsic Lyman break of a factor 5, and assuming a fraction of 30

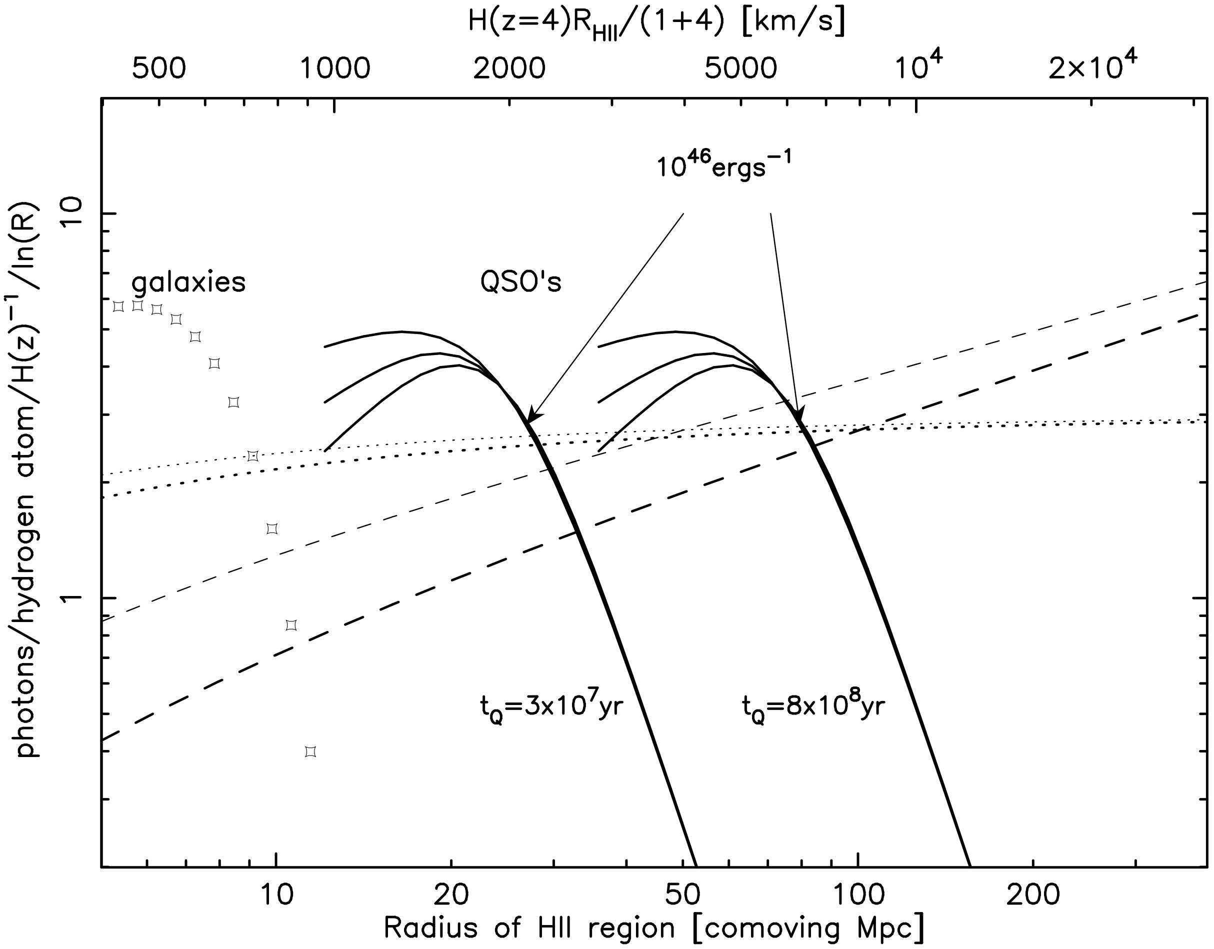

Fig. 7.— The dashed lines show the global number of recombinations taking place per Hubble time and per hydrogen atom as a function of the mean free path at redshifts $z = 4$ and $z = 6$. The recombination rate was calculated as described in §3 for the density distribution shown in Fig. 1. The dotted curves give the mass fraction required to ionize up to the overdensity corresponding to that mean free path, times a factor 3. The solid curves show the mean comoving volume emissivity from QSOs in terms of the number of ionizing photons emitted per hydrogen atom per Hubble time (at $z = 5$) as a function of the comoving radius of the reionized region created by a quasar over a lifetime $t_{\rm Q}$, if recombinations are ignored. The QSO luminosity function at $z = 4$ of Pei (1995) is used. The three curves in each set are for faint end slopes $\beta_{\rm l} = (1.5, 1.65, 1.8)$. The open diamond symbols show the mean comoving volume emissivity of star-forming galaxies at $z = 4$ calculated from the luminosity fuction according to Dickinson et al. (1998). An intrinsic Lyman break of a factor five and an escape fraction of ionizing photons of 30 percent was used to extrapolate from the flux at 1500 Å and a lifetime $t_{\rm Q} = 8 \times 10^8\,$yr was used. The arrows denote the radii which correspond to a $10^{46}{\rm erg\,s^{-1}}$ QSO

% of emitted ionizing photons to escape into the IGM. The age of the galaxies is set equal to 8×10^8 years, or half the age of the universe at $z = 4$.

The integrals under the curves of the energy functions give us the total emissivity, in terms of number of emitted photons per hydrogen atom per Hubble time. For both QSO's and galaxies, their total emissivity at $z = 4$ is $\epsilon \sim 5$. The radii of the H II regions when the lifetime is set to $t_S = (\epsilon H)^{-1}$ have a special significance: over the time $(\epsilon H)^{-1}$, one ionizing photon is emitted for each atom in the universe, and therefore the H II regions around each source are just large enough for overlapping. Therefore, for this lifetime, the variable in the horizontal axis is equal to the radius that the ionized regions around a source of a given luminosity must have grown to at the epoch of overlap, when reionization ends. The curve corresponding to this lifetime would be shifted to the left from the curve for $t_Q = 8 \times 10^8$ yr by a factor $(5/3)^{1/3} = 1.19$.

The requirement for having completed the reionization with a given population of sources can now be easily visualized from Figure 7. We shall consider first the case of long-lived sources that remain active with an approximately constant luminosity over the epoch of reionization; the differences in the case of short-lived sources will be discussed later. In the previous sections we computed the global recombination rate in terms of the maximum overdensity to which the universe is reionized, Δ_i, which in turn can be expressed in terms of the mean free path λ_i (using equations [1], [2], [9], and [12]). The dashed lines in Figure 7 give the number of recombinations taking place per Hubble time as a function of λ_i. The dotted curves give the mass fraction F_M required to ionize the IGM up to the overdensity corresponding to the mean free path λ_i, times a factor 3. These curves are shown at $z = 4$ (thick lines) and $z = 6$ (thin lines). At high redshift, the emissivity from quasars could of course be different, depending on how they evolve. At the epoch of overlap, the radii of the H II regions must have grown to the values derived for the lifetime $t_S = (\epsilon H)^{-1}$. The mean free path of the ionizing photons at this time will be of order the radius of the H II regions (see the discussion at the end of §2). The first condition necessary to reach the epoch of overlap is that $\epsilon t_R H > F_M$, where t_R is the time during which sources have been producing the emissivity ϵ, and $\epsilon t_R H$ is the total number of photons emitted. Choosing $t_R = (3H)^{-1}$ (i.e., half the age of the universe at the epoch of overlap), this condition becomes $\epsilon > 3F_M$, roughly equivalent to requiring that the solid line for a lifetime t_S should be above the dotted line in Figure 7. Obviously, if reionization occurs with a large degree of synchronization ($t_R \ll H^{-1}$), a greater emissivity is required. The second condition is that the global recombination rate when the mean free path is equal to the size of the H II regions that must be reached for overlap is not greater than the emissivity. This condition roughly implies that the solid line must be above the dashed line.

As the emissivity of the sources increases, the epoch of overlap is reached when the mean emissivity of the sources rises above both the dashed and dotted lines. This shows that *the importance of recombinations, and the required emissivity to reionize the IGM, increases not only with redshift but also with the luminosity of the sources.*

For sources with luminosities typical of high-redshift starburst galaxies, the fact that the dashed line is below the dotted line implies that recombinations should not be very important in the case of late reionization by galaxies. Even if very luminous QSO's reionized the universe, recombinations increase the required number of emitted photons by a moderate factor only, because the IGM does not need to be ionized up to a very high overdensity to increase the mean free path up to the size of the H II regions produced by these sources. Recombinations can of course be more important as the redshift of reionization is increased.

Now, we consider the case of short-lived sources, with lifetimes much less than t_R. As the reionization proceeds, the IGM should contain active H II regions, which continue to expand as more radiation is emitted by the source, and fossil H II regions, which started to recombine after their central source turned off, but are still mostly ionized (if an H II region becomes mostly neutral a long time after the source turns off, its presence has no effect on the subsequent development of reionization and can therefore be considered as "disappeared"). This implies that in the case of short-lived sources, there are two different epochs of overlap, a first one for the overlap of fossil H II regions, and a second one for the overlap of active H II regions. The first epoch of overlap of fossil H II regions is determined by the energy function of sources. This occurs when the emissivity rises above the dashed and dotted lines at the value of the mean free path equal to the radius of fossil H II regions, determined by the lifetime $t_{\rm Q}$. The second epoch of overlap is determined by the luminosity function, and it generally occurs significantly later and requires a higher emissivity, determined by the same condition as for long-lived sources obtained when the solid lines are shifted to the right to the lifetime $t_{\rm S} = (\epsilon H)^{-1}$. Between these two epochs of overlap, the baryons that have recombined in fossil H II regions need to be ionized again, and additional gas in more overdense regions needs to be ionized in order to increase the mean free path to the separation between active H II regions. The shorter the lifetime of the sources, the more widely separated the two epochs of overlap become.

In the example of a lifetime $t_Q = 3 \times 10^7$ years, fossil H II regions overlap when a typical quasar with luminosity $10^{46} \mathrm{erg\,s^{-1}}$ has produced an H II region of comoving radius 25 Mpc. The overlap of active H II regions would be reached only when the radius of H II regions around a quasar of this luminosity increases up to a radius of 70 Mpc, requiring a higher emissivity to balance the global recombination rate.

It is clear from Figure 7 that for both QSO's and galaxies, the emissivity is sufficient to have reionized the universe by $z = 4$. If we use the approximation $\epsilon = R$, our model predicts that at $z = 4$, $\Delta_{\rm i} = 100$ and $R \simeq 6 R_{\rm u} \simeq 4$ (see Figs. 2c, 5), so the combined emissivity of quasars and galaxies can actually be slightly higher than what is needed to match the observed flux decrement. At higher redshifts, the dashed line indicating the recombination rate increases and also becomes shallower, and the epoch of reionization is reached when the dashed line is equal to the emissivity at the mean separation between the sources. If the emissivity from low-luminosity galaxies remains at a constant level, this may not happen until $z \simeq 10$ or higher. To make further progress, a determination of the escape fraction of ionizing photons from high-redshift galaxies and of the number densities of faint quasars and galaxies at high redshift will be required.

4.2. The Size of H II Regions at the Epoch of Reionization

We now come back to the question of whether individual H II regions might be seen in the Lyα spectra of high-redshift sources before the hydrogen reionization is completed. In §3 we showed that gaps of transmitted flux through the Gunn-Peterson trough will generally be produced by H II regions if $\tau_{\rm u} \lesssim 50$, which for our value of the mean free path corresponds to a size of the H II region of $H\,\lambda_{\rm i} \simeq 4000\,{\rm km\,s^{-1}}$, or $(1+z)\lambda_{\rm i} \simeq 40$ Mpc, at $z \simeq 6$ (the lowest redshift where a patchy reionized hydrogen in the IGM might be present). Smaller H II regions can still produce a gap when $f_J > 1$, if the line of sight passes through a void close to the central source, where the intensity of radiation is higher than average. At higher redshifts, the required size of the H II regions to produce a gap of transmitted flux increases, both due to the lower underdensities of voids, and to the longer mean free path at fixed $\tau_{\rm u}$ (see Figures 3, 5 and 6).

Figure 7 shows that QSO's of sufficiently high luminosity can meet this requirement. In the case of short-lived, highly luminous QSO's, the H II regions created by the QSO's would be smaller at the first epoch of overlap of fossil regions. However, when the active H II regions overlap, their size obviously depends only on the observed luminosity function of active sources, and not on their lifetimes. Star forming galaxies do not reach the required size, and therefore their associated H II regions can only be visible in cases where f_J is large.

The size of H II regions at overlap has strong implications for the effect of reionization on the fluctuations of the CMB background (e.g. Gruzinov & Hu 1998). Unfortunately the constraints are extremely weak. The typical size could range from less than 1 Mpc if faint star-forming galaxies were the ionizing sources to several tens of Mpc if reionization were caused by bright QSO's.

5. Helium Reionization

The double reionization of helium can be treated following the same method we have used for hydrogen. The He III recombination rate is 5.5 times faster than that of hydrogen (this is increased to 5.9 times faster when including the increase in the electron density due to the ionization of helium), $R_{\rm HeIII,u}(z) = 0.21\,[\Omega_{\rm b}h(1-Y/2)/0.027]\,(1+z)^{3/2}/\Omega_0^{1/2}$. If the number of photons emitted above the He II ionization threshold is not higher than $5.9 n_{\rm He}/n_{\rm H} = 0.46$ times the number of photons emitted above the hydrogen ionzation edge (which is the case for all existing candidates of the sources of the ionizing background), and if recombinations are the dominant balance to the emissivity, then the He II should be reionized at a later epoch than the hydrogen.

Equations (2), (5), (11), and (12) are equally applicable for He II , once we define the new quantities $\epsilon_{\rm He}$ and $n_{\rm J,He}$ as the emissivity per Hubble time and the number density of photons above the He II ionization edge for each helium atom, and $\Delta_{\rm HeIII}$ as the density below which helium starts being doubly ionized. Equation (16) is also equally valid, except that the numerical factor changes from 1.14 to 1.23 (due to the ratio of the number of hydrogen atoms to the total number of hydrogen and helium atoms that is included in this equation for the case of hydrogen). The basic reason is that the ratio of the integrated optical depths to Lyα line scattering and to continuum ionizing radiation is always the same for an H II or a He III region.

Figure 8 shows $\tau_{\rm u}$ versus $\Delta_{\rm HeIII}$, and Figure 9 shows the flux decrement, at redshifts $z = 2, 3, 4$ (the latter is obtained directly from the numerical simulation in MCOR [see their Fig. 19], and it includes the effect of thermal broadening of helium). The solid triangle indicates the observed mean flux decrement in the He II Lyα forest at redshift near $z = 3$ (Davidsen et al. 1996; Reimers et al. 1997; Hogan et al. 1997; Anderson et al. 1998). In order to match this observation of the flux decrement in our model for the gas density distribution and the mean free path, helium should be doubly ionized up to overdensities $\Delta_{\rm HeIII} \simeq 50$ at $z = 3$, with a corresponding mean free path $H\lambda_{\rm He} \simeq 7000\,{\rm km\,s^{-1}}$. As in the case of hydrogen, we see that even if the helium in the IGM is doubly ionized over most of the volume at higher redshifts, the transmitted flux should be negligibly small when $\Delta_{\rm HeIII} \lesssim 15$.

Figure 10 is analogous to Figure 7, showing the emissivity of He II -ionizing photons per helium atom from quasars as a function of the comoving radii of their He III regions, at $z = 3$, plotted for two assumed lifetimes: 3×10^7 and 10^9 years (the latter is $\sim$ half the age of the universe at $z = 3$). The top horizontal axis gives the Hubble velocity corresponding to the radius of the He III region at $z = 3$. We use again the parameterization of the QSO luminosity function given by Pei (1995). As before the dashed line shows the

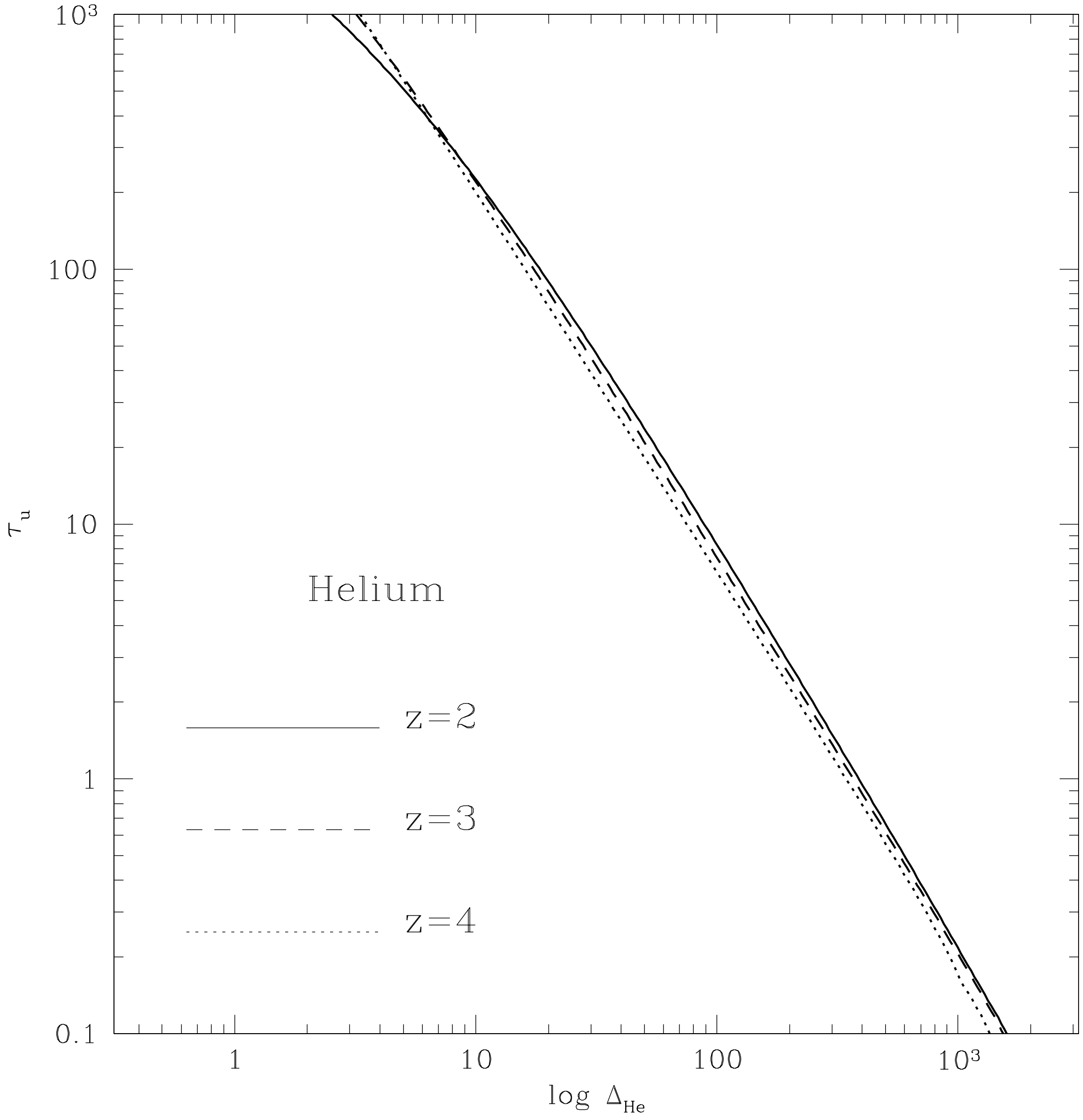

Fig. 8.— Same as Fig. 3, for the case of helium.

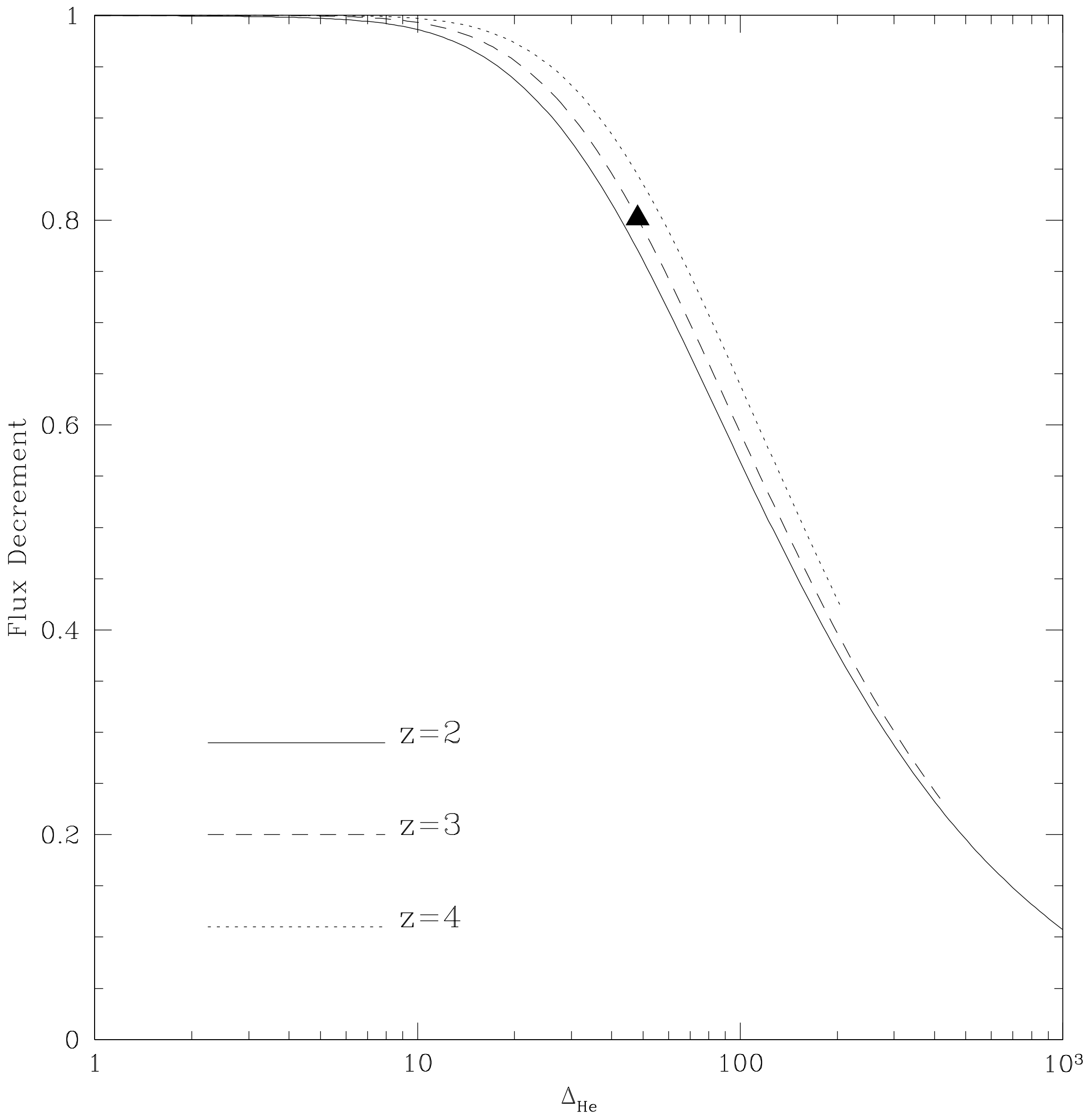

Fig. 9.— Mean flux decrement as a function of the overdensity up to which the gas is ionized, for the case of helium. Solid triangle indicates the observed mean flux decrement at $z = 3$.

global recombination rate per helium atom. We have assumed a spectral index of $\alpha = 1.8$ ($f_\nu \propto \nu^{-\alpha}$) to extrapolate from the hydrogen to the helium Lyman continuum (Zheng et al. 1997). The total emissivity of QSO's is sufficient to completely reionize He II before $z = 3$ according to the model by Pei, since the emissivity is above the global recombination rate for the lifetime $t_S = (\epsilon_{He} H)^{-1} \simeq 1.7 \times 10^8$ yr.

Notice that, in the case of helium, recombinations are much more important than the first ionization of the helium for luminous sources. Because the recombination rate grows rather steeply with the mean free path, low-luminosity sources are very likely to have ionized the He II entirely at $z > 3$. If the emissivity from sources with the abundance of galaxies in Figure 7 is more than 10% of the emissivity from luminous quasars, then galaxies should have reionized the He II first. The emissivity from galaxies is not likely to be much smaller, because the cooling radiation from supernova remnants and hot gas in the halo produces photons of the right energy for the ionization of helium. Therefore, unless luminous quasars have always dominated by a large factor the emissivity of He II -ionizing photons, the He II will be ionized by lower luminosity sources and the epoch of overlap of the He III regions should be substantially earlier than $z = 3$.

However, the emissivity from quasars for the lifetime t_S is barely above the dashed line, and the mean free path of helium Lyman continuum photons is similar to the separation between sources with luminosities greater than $\sim 2 \times 10^{46} \mathrm{erg\,s^{-1}}$, which still contribute substantially to the emissivity. The He III ionizing flux should therefore fluctuate substantially because, in a random region of space, it is dominated by only a few active sources. This can produce optical depth fluctuations in the He II forest on scales comparable to the mean free path, $\sim 7000 \,\mathrm{km\,s^{-1}}$. These fluctuations may also be due, however, to the random fluctuations in the number of underdense voids in the IGM crossed by the line of sight. A strong proximity effect is also expected, and is observed in the spectra of two of the four quasars for which the He II Lyα forest has been observed (Anderson et al. 1998 and references therein).

In the case of the quasar HE 2347-4342, observed by Reimers et al. (1997), it was claimed that the ratio of the hydrogen optical depth to the He II optical depth varied by large factors, which can only be explained by a large variation in the spectrum of the ionizing background. Two gaps of transmitted flux were observed in the He II Lyα spectrum, with widths of 500 to $1000 \,\mathrm{km\,s^{-1}}$ (see their Fig. 4). One of these two gaps seems to have detectable associated hydrogen Lyα absorption, whereas other regions in the spectrum having no evidence for hydrogen Lyα absorption show no tranmitted flux in the He II spectrum. Reimers et al. interpreted this as evidence of a patchy He II ionization in the IGM, with individual He III regions producing the gaps of transmitted flux in their He II Lyα

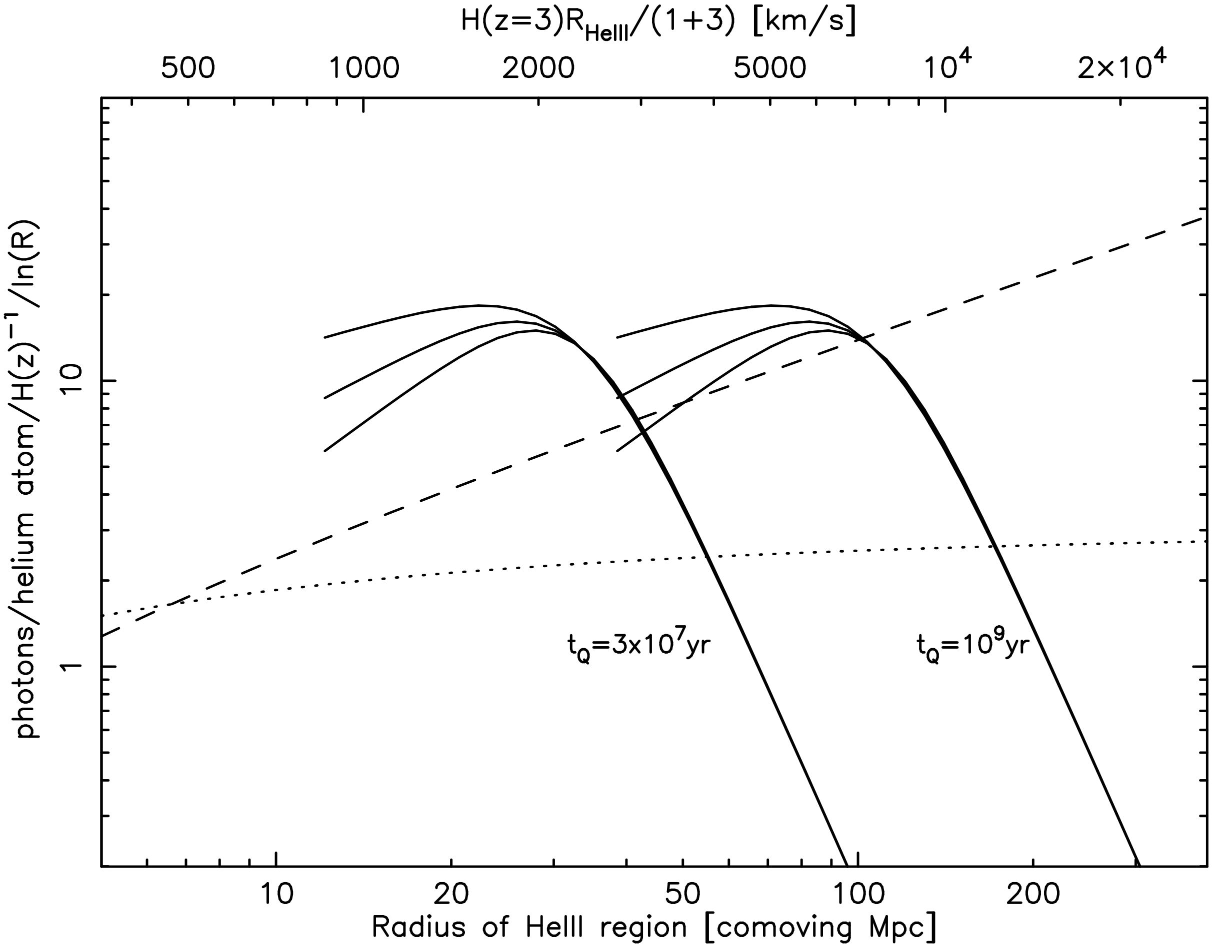

Fig. 10.— Same as Figure 7, for the case of helium, at $z = 3$. The emissivity from quasars is shown here in terms of number of photons emitted per Hubble time and per helium atom, as a function of the radius of the He III region that a quasar of lifetime t_Q would create in the absence of any recombinations.

spectrum, and He II patches yielding no observable He II flux in regions where the hydrogen optical depth is very small. However, it is not clear if these large fluctuations in the ratio of the hydrogen to He II optical depth are real, or could arise from errors in fitting the quasar continuum in the hydrogen Lyα spectrum or subtracting the background flux in the He II spectrum. As explained before, it is unlikely that a patchy ionization structure in the He II is present in the IGM at $z = 3$. This could only occur if very luminous quasars were the only sources of He II -ionizing photons. Therefore, the observed gaps in the spectrum of HE 2347-4342 are probably associated with underdense regions of the IGM ionized by a fluctuating ionizing background (see also the discussion in M98).

Songaila & Cowie (1996) have reported a decrease of the Si IV/C IV ratio with increasing redshift for intermediate column density hydrogen Lyα absorption lines (see also Boksenberg 1998). Songaila & Cowie interpreted this evolution as a softening of the spectrum with increasing redshift and argued for a rather sudden reionization of He III at $z \simeq 3$. As discussed in §3, reionization is expected to be gradual, and we should therefore expect a progressive decrease of the intensity of the He II -ionizing background with redshift, as the He II photon mean free path decreases. This is consistent with the more recent results of Boksenberg, Sargent, & Rauch (1998), where a smooth evolution of the ratios Si IV/C IV and C II/C IV is reported, and also with the findings of Davé et al. (1998) who used the observed O VI absorption to argue that at least 50 percent of the volume of the universe at redshift three are illuminated by a hard spectrum with significant flux well above the He II Lyman edge.

The reionization of He II can affect the evolution and spatial distribution of the temperature in the IGM (Miralda-Escudé & Rees 1994). Each helium atom yields an energy input of $\sim 54/(\alpha - 1)\,\mathrm{eV}$. For $\alpha = 1.8$, this results in a temperature increase of 20000 K. Since hydrogen is already fully ionized by the time the He II reionization occurs, line cooling processes are suppressed and the time for thermal evolution in the IGM becomes as long as the Hubble time. Thus, spatial fluctuations in the temperature introduced when He II is reionized, modulated on scales as large as the size of He III regions for the most luminous quasars (~ 100 comoving Mpc, as shown in Fig. 10), could affect the evolution of the IGM and the formation of low-mass galaxies, because the increased entropy of the ionized gas in regions heated by the hard spectrum of a luminous quasar could slow down the collapse of gas in halos of low velocity dispersion (see e.g., Efstathiou 1992, Thoul & Weinberg 1996). In the likely case where the ionizing radiation from a quasar is anisotropic, predominantly emitted along the jets, the present distribution of dwarf galaxies might show a large quadrupole moment around the locations of ancient quasars, probably occupied today by massive clusters of galaxies. These special galaxy fluctuations, which would constitute a new type of non-local galaxy bias unrelated to the primordial fluctuations, might be detected in future galaxy redshift surveys on the very large scales that He III regions can reach, where the amplitude of primordial fluctuations is greatly decreased.

6. Discussion and Conclusions

We have discussed in this paper how the clumpiness of the matter distribution and the discreteness of ionizing sources affects the reionization of hydrogen and helium and the appearance of the absorption spectra of high-redshift sources. In a clumpy medium, the underdense regions filling most of the volume of the universe will be reionized first, while the denser regions are gradually ionized later from the outside. As the emissivity increases, a balance is maintained with the global recombination rate by increasing the mean free path of ionizing photons, as the size of the neutral, self-shielded regions at high density (which are observed as Lyman limit systems) shrinks.

The first gaps in the hydrogen and helium Gunn-Peterson trough in the absorption spectra of high-redshift objects should be caused by the most underdense voids. As the redshift is increased, the transmitted flux in Lyα spectra decreases rapidly not only because the intensity of the ionizing background diminishes if the emissivity from sources is lower, but also because the global recombination rate increases, and the voids are less underdense as predicted by the gravitational evolution of large-scale structure. We showed that if the emissivity does not increase with redshift at $z > 4$, then the Gunn-Peterson trough should be reached at $z \simeq 6$. This redshift is determined by the fact that the most underdense voids become optically thick to Lyα photons. However, the low-density IGM is likely to have been ionized everywhere at this redshift already, and the epoch of overlap of the H II regions around individual sources could have occurred at a substantially higher redshift if the emissivity is dominated by sources of low luminosity.

Whether it will be possible to see gaps of transmitted flux in a Lyα spectrum across the Gunn-Peterson trough, produced by individual H II regions around ionizing sources, before the major fraction of the volume of the universe is reionized, depends on the luminosity of the source. For a QSO with luminosity $10^{46}\mathrm{erg\,s^{-1}}$ at $z = 6$, creating an H II region of radius ~ 70 comoving Mpc, the mean transmitted flux through the H II region would only be 10% (from Figures 5 and 6), and sources of lower luminosity would produce no detectable gap in the Gunn-Peterson trough. If luminous quasars exist but are short-lived, there should be a first epoch of overlap of fossil H II regions (where the size of the H II regions would be smaller by a factor $(t_{\mathrm{Q}}\epsilon H)^{1/3}$ relative to the case of long-lived sources), and a second epoch of overlap of active regions, having a size depending only on the luminosity of the quasars.

Smaller H II regions might on rare occasions produce gaps of transmitted flux when the intensity of radiation is much higher than average, when an underdense void is intersected by a line of sight that is in close proximity to the source, compared to the radius of the H II region. For an H II region with proper radius smaller than 1 Mpc, surrounded by neutral IGM, any residual flux that might be transmitted in these circumstances would still be absorbed in the damping wings of the absorption by the neutral IGM (see M98),

The reionization might be complete at a substantially earlier time than the redshift were the Gunn-Peterson trough is first encountered, if it is produced by low-luminosity sources. The fact that reionization occurs outside-in and the dense gas stays neutral until a late epoch, well after the overlap of H II regions, implies that the clumpiness of the IGM does not increase the mean number of times that a baryon will recombine during the reionization epoch. Thus, for the density distribution suggested by numerical simulations, the number of photons per hydrogen atom required for the reionization is of order unity. In fact, this number can even be less than unity because the baryons in high density regions do not need to be ionized in order for the mean free path of the photons to increase to values much larger than the mean separation between sources. Only at very high redshift, when $R_{\rm u} \gg 1$, recombinations imply a significant increase in the required emissivity for reionization. This modest requirement can be met by both faint QSO's and star-forming galaxies. The remaining uncertainties to determine which sources reionized the IGM are the space density of faint high-redshift QSO's and star-forming galaxies, and the escape fraction of Lyman continuum photons from high-redshift galaxies. The required number of photons for reionization could, however, be substantially higher if a significant fraction of the gas was efficiently locked up in small haloes of high density but large covering factor, which collapsed before reionization.

We would like to thank Tom Abel, Piero Madau, and Tom Theuns for helpful discussions. We also thank R. Cen and J. P. Ostriker for giving us permission to show in Figure 1 the density distribution obtained from their numerical simulation. JM thanks the Alfred P. Sloan Foundation for support.

REFERENCES

Abel, T., & Mo, H. J. 1998, ApJ, 494, L151

Anderson, S. F., Hogan, C. J., Williams, B. F., & Carswell, R. F. 1998, AJ, submitted (astr0-ph/9808105)

Arons J., & Wingert D. W. 1972, ApJ, 177,1

Bahcall, J. N., & Salpeter, E. E 1965, ApJ, 142, 1677

Boksenberg, A., 1997, in Structure and Evolution of the IGM, 13th IAP colloquium, eds. Petitjean P., Charlot S., p. 85

Boksenberg, A., Sargent, W. L. W., & Rauch, M. 1998, to appear in the Proceedings of the Xth Rencontre the Blois, The Birth of Galaxies (astro-ph/9810502)

Boyle B., & Terlevich, R. J., 1998, MNRAS, 293, 49p

Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9

Davé, R., Hellsten, U., Hernquist, L., Katz, N., & Weinberg, D. H. 1998, ApJ, submitted (astro-ph/9803257)

Davidsen, A. F., Kriss, G. A., & Zheng, W. 1996, Nature, 380, 47

Dey A., Spinrad H., Stern D., Graham J.R., & Chaffee F.H., 1998, ApJ, 498, L93

Dickinson, M. 1998, in STScI Symposium 1997 "The Hubble Deep Field", eds. M. Livio, S. M. Fall, P. Madau (astro-ph/9802064)

Efstathiou, G. 1992, MNRAS, 256, 43p

Gnedin, N. Y., & Ostriker, J. P. 1997, ApJ, 486, 581

Gruzinov, A., & Hu, W. 1998, ApJ, in press

Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633

Haardt, F., & Madau, P. 1996, ApJ, 461, 20

Haehnelt, M. G., Natarajan, P., & Rees, M. J. 1998, MNRAS, 300, 817

Haiman, Z., & Loeb, A. 1998, ApJ, submitted (astro-ph/9807070)

Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, ApJ, 457, L51

Hogan, C. J., Anderson, S. F., & Rugers, M. H. 1997, AJ, 113, 1495

Hurwitz, M., Jelinski, P., & van Dyke Dixon, W. 1997, ApJ, 481, L31

Jakobsen, P., et al. 1994, Nat, 370, 35

Jakobsen, P. , et. al., 1996, in Benvenuti P. Macchetto F.D., Schreier E.J., " Science with the Hubble Space Telescope-II, STScI, Baltimore, p. 153

Leitherer, C., Ferguson, H. C., Heckman, T. M., & Lowenthal, J. D. 1995, ApJ, 454, L19

Madau, P., in: Corbelli E., Galli D., Palla F., eds. Molecular Hydrogen in the Early Universe, Mem.S.A., in press, astro-ph/9804280

Madau, P., Haardt, F., & Rees, M. J. 1998, submitted to ApJ (astro-ph/9809058)

Miralda-Escudé J., 1998, ApJ, 501, 15 (M98)

Miralda-Escudé J., Rees M.J., 1994, MNRAS, 266, 343

Miralda-Escudé, J., Cen, R., Ostriker, J. P., Rauch, M., 1996, ApJ, 471, 582

Pei, Y. C. 1995, ApJ, 438, 623

Press, W., Rybicki, G. B., & Schneider, D. P. 1993 ApJ, 414, 64

Rauch, M., Miralda-Escudé, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, H., Katz, N., Cen, R., & Ostriker J. P. 1997, ApJ, 489, 7

Reimers D., Köhler S., Wisotzki L., Groote D., Rodriguez-Pascal P., Wamsteker W., 1997, AA, 327, 890

Scheuer, P., 1965, Nature, 207, 963

Schneider, D. P., Schmidt, M., & Gunn, J. E. 1991, AJ, 102, 837

Shapiro, P., & Giroux, M. 1987, ApJ, 321, L107

Songaila, A., & Cowie, L. L. 1996, AJ, 112, 335

Steidel, C. S., Adelberger, K. L, Giavalisco, M., Dickinson, M., Pettini, M. 1998, ApJ, submitted (astro-ph/9811399)

Storrie-Lombardi, L. J., McMahon, R. G., Irwin, M. J., & Hazard, C. 1994, Apj, 427, L13

Thoul, A., & Weinberg, D. H. 1996, ApJ, 465, 608

Weymann, R. J., Stern D., Bunker A., Spinrad H., Chafeee, F. H., Thompson R. I., & Storrie-Lombardi L. J. 1998, ApJ, 505, L95 (astro-ph/9807208)

Zhang, Y., Meiksin, A., Anninos, P., & Norman, M. L. 1998, ApJ, 495, 63

Zhang, Y., Anninos, P., & Norman, M. L. 1995, ApJ, 453, L57

Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, ApJ, 492, 855

Zuo L., 1992, MNRAS, 258, 36