Mon. Not. R. Astron. Soc. **000**, 1–11 (2000) Printed 8 February 2001 (MN LaTeX style file v1.4)

How big were the first cosmological objects?

Roger M. Hutchings[1], F. Santoro[1], P. A. Thomas[1⋆] & H. M. P. Couchman[2]
[1] *Astronomy Centre, CPES, University of Sussex, Falmer, Brighton, BN1 9QJ*
[2] *Department of Physics & Astronomy, McMaster University, Hamilton, Ontario, L8S 4M1, Canada*

ABSTRACT
We calculate the cooling times at constant density for halos with virial temperatures from 100 K to 1×10^5 K that originate from a 3σ fluctuation of a CDM power spectrum in three different cosmologies. Our intention is to determine the first objects that can cool to low temperatures, but not to follow their dynamical evolution. We identify two generations of halos: those with low virial temperatures, $T_{\rm vir} \lesssim 9000$ K that remain largely neutral, and those with larger virial temperatures that become reionized. The lower-temperature, lower-mass halos are the first to cool to 75 percent of their virial temperature. The precise temperature and mass of the first objects are dependent upon the molecular hydrogen (H_2) cooling function and the cosmological model. The higher-mass halos collapse later but cool much more efficiently once they have done so, first via electronic transitions and then via molecular cooling: in fact, a greater residual ionization once the halos cool below 9000 K results in an enhanced H_2 production and hence a higher cooling rate at low temperatures than for the lower-mass halos, so that within our constant-density model it is the former that are the first to cool to really low temperatures. In addition, the difference in collapse times between low- and high-mass halos is similar to the lifetime of massive stars. Further investigation is required to determine which generation of halos plays the dominant role in the formation of the first star clusters or whether both are equally important.

Key words: cosmology: theory — early Universe — galaxies: formation — molecular processes

1 INTRODUCTION

The study of the first generation of objects in the Universe that are able to cool sufficiently to collapse and form luminous objects is a well-defined problem. By definition, there are no stars or other sources of ionizing radiation, and one does not have to consider feedback from supernovae and enrichment of the Universe with metals. The first objects to form arise from the collapse of high-sigma fluctuations in the background density field. These peaks will virialize and begin to cool. Objects with virial temperatures $T \lesssim 9\,000$ K are cooled by H_2 molecules: the molecules are excited by collisions with neutral hydrogen and then spontaneously de-excite with the emission of a photon. In the absence of metals the dominant cooling mechanisms for $T \gtrsim 9\,000$ K are collisional excitation and ionization of neutral hydrogen and recombination of ionized hydrogen.

The question of what are the masses of the first objects to form in a standard CDM scenario has been tackled by Tegmark et al. (1997, hereafter T97) in a paper entitled "How small were the first cosmological objects?" They analytically tracked a top-hat collapse to the point of virialization, at which point the gas was cooled at constant density. They accepted an object as having cooled if it met the criterion $T(0.75 z_{\rm vir}) \leq 0.75\, T_{\rm vir}$, where $T_{\rm vir}$ is the virial temperature and $z_{\rm vir}$ the virialization redshift. They found that the first generation of objects that cooled in a standard CDM scenario, virialized at a redshift of 27 and had a baryonic mass of about 10^5 $M_\odot$. In a later paper, Abel et al. (1998) redid the calculation with a different H_2 cooling function and estimated a very similar virialization redshift but a smaller baryonic mass, 7×10^3 $M_\odot$.

In the present paper we adopt a similar approach to T97 but consider the collapse of not only small objects with virial temperatures $\leq 9\,000$ K, but also objects of virial temperatures up to 100 000 K. The motivation for this is that the high virial temperatures will partially re-ionize the gas. Since the gas will then cool more rapidly than it can recombine, the ionization level at temperatures $\leq 9\,000$ K will be greater than it otherwise would have been if no re-ionization had taken place. This in turn accelerates production of H_2, ultimately resulting in enhanced cooling at lower temperatures.

This means larger objects would cool to low temper-

⋆ p.a.thomas@sussex.ac.uk

© 2000 RAS

atures before smaller ones, assuming they virialized at the same time. In CDM cosmologies, the collapse of larger objects occurs after the collapse of smaller ones, but only slightly. The question as to whether larger objects can cool sufficiently more rapidly than smaller ones if the delay in collapse is taken into account is the main topic of this paper.

A second difference between our study and previous ones is that we consider more up-to-date cosmological models. In particular, we choose a density fluctuation spectrum with less power on small scales. In a critical density universe, this has the effect of delaying collapse until a redshift of 12. Higher redshifts can be recovered by the introduction of a cosmological constant.

We describe our chemical model in Section 2 and our numerical method in Section 3. The results are presented in Section 4 and discussed in Section 5. Finally, we summarize our conclusions in Section 6.

2 GAS CHEMISTRY AND COOLING FUNCTIONS

In this section we present a minimal model of the gas chemistry needed to accurately follow the temperature evolution of the gas. Our model is similar to those of Abel et al. (1997) and Fuller & Couchman (2000) but differs from theirs in the manner in which it handles the production and destruction of H^- and H_2^+. As a consequence, we end up with slightly different terms in our final equation for H_2 production. In principle this could be quite important; however it seems to make little difference to the results over the parameter ranges considered in this paper.

Table 1 lists the important reactions for the combinations of temperature and species abundances that we consider. The model is applicable to halos with temperatures between 100 K and 100 000 K and redshifts up to 50. The model does not include helium chemistry–the only effect this would have would be to introduce an error of up to 16 per cent in the ionization level at high temperatures. Since helium recombines more rapidly than hydrogen, any error will be reduced as the temperature drops. Nor do we include photo-ionization from cosmic microwave background photons which is only important at redshifts in excess of 100.

2.1 H_2 production and destruction

The equation for the rate of change of H_2 abundance is derived in the appendix:

$$\frac{dn_{\rm H_2}}{dt} = n_{\rm H}{}^2 \left[n_{\rm e^-} \frac{R_3 R_4}{R_{\rm H^-}} + n_{\rm H^+} \frac{R_6 R_7}{R_{\rm H_2^+}} \right] - n_{\rm H_2} \left[n_{\rm H^+} n_{\rm e^-} \frac{R_9 R_8}{R_{\rm H_2^+}} + n_{\rm H} R_{10} + n_{\rm e^-} R_{11} \right]. \quad (1)$$

We next discuss the regimes under which each of these terms is important.

2.1.1 *Formation of H_2*

The H^- channel is the dominant H_2 formation path. The H_2^+ channel is important only at high redshifts ($z \gtrsim 200$ when the H^- channel is suppressed by photo-destruction of H^- that we have omitted here) or high temperatures. However, at high temperatures, the H_2 is rapidly destroyed and for all the models that we consider in this paper, omitting the H_2^+ formation channel makes no difference to the final abundance of H_2.

2.1.2 *Destruction of H_2*

It is the process of H_2 destruction where the present work differs from that of Abel et al. (1997) and Fuller & Couchman (2000). In these two papers, the term for destruction by H^+ is simply included as

$$n_{\rm H_2} n_{\rm H^+} R_9, \quad (2)$$

whereas we have

$$n_{\rm H_2} n_{\rm H^+} R_9 \frac{n_{\rm e^-} R_8}{R_{\rm H_2^+}} = n_{\rm H_2} n_{\rm H^+} R_9 \left(1 - \frac{n_{\rm H} R_7}{R_{\rm H_2^+}} \right). \quad (3)$$

This takes into account that, at low ionization levels, much of the H_2^+ that is produced will be immediately converted back to H_2 so that there will be no net destruction. In some of the models that we consider, this can make transitory differences of a factor of 10 or more in the H_2 abundance. However, as H_2 production takes over from destruction at low temperatures, the effect on the final H_2 abundance is quite small (at most a few percent).

With the correct rate for the H_2^+ destruction channel, there is a small parameter range for which H10 becomes the dominant destruction process. However, this is so fleeting that it makes a negligible difference to the results and can safely be omitted.

To be fair on the previous papers, they consider only low temperatures, $T < 6\,000$ K for which H_2 destruction is relatively unimportant compared to its production, and the error in their results is negligible. Nevertheless, a minimal model that includes H_2 destruction should use our Equation 1.

2.2 Ionization level

In principle, the Equation for the rate of change of ionization is every bit as complicated as that for H_2 abundance. However, it turns out in practice that there are only two important terms:

$$\frac{dn_{\rm e^-}}{dt} = n_{\rm e^-} n_{\rm H} R_1 - n_{\rm e^-} n_{\rm H^+} R_2. \quad (4)$$

2.3 Cooling Terms

We use the molecular hydrogen cooling rate from Galli & Palla (1998) as summarised by Fuller & Couchman (2000). This gives a fit to the low-density limit of the calculations of Martin et al. (1996) and Forrey et al. (1997) which together cover a wide temperature range:

$$\log_{10} \left(\frac{\Lambda_{\rm H_2}(T)}{\rm erg\,cm^3\,s^{-1}} \right) = -103.0 + 97.59\, T_{\rm log} - 48.05\, T_{\rm log}^2$$

Table 1. This table summarizes the important reactions needed in order to calculate accurately the abundance of H_2. References are: HTL Haiman, Thoul & Loeb (1996); GP Galli & Palla (1998); FC Fuller & Couchman (2000); SLD Stancil, Lepp & Dalgarno (1998); AAZN Abel, Anninos, Zhang & Norman (1997).

	Reaction	Rate/cm^{3}s^{-1}	Reference
H1	$H + e^- \longmapsto H^+ + 2e^-$	$5.9 \times 10^{-11}\, T_0^{0.5}(1+T_5^{0.5})^{-1} \exp(-1.578/T_5)$	HTL
H2	$H^+ + e^- \longmapsto H + \gamma$	$3.3 \times 10^{-10}\, T_0^{-0.7}(1+T_6^{0.7})^{-1}$	HTL
H3	$H + e^- \longmapsto H^- + \gamma$	$1.4 \times 10^{-18}\, T_0^{0.928} \exp(-T_4/1.62)$	GP
H4	$H^- + H \longmapsto H_2 + e^-$	1.3×10^{-9}	FC
H5	$H^- + H^+ \longmapsto 2H$	$4.0 \times 10^{-6}\, T_0^{-0.5}$	FC
H6	$H + H^+ \longmapsto H_2^+ + \gamma$	$2.1 \times 10^{-23}\, T_0^{1.8} \exp(-2/T_1)$	SLD
H7	$H_2^+ + H \longmapsto H_2 + H^+$	6.4×10^{-10}	GP
H8	$H_2^+ + e^- \longmapsto 2H$	$1.2 \times 10^{-7}\, T_0^{-0.4}$	SLD
H9	$H_2 + H^+ \longmapsto H_2^+ + H$	$3.0 \times 10^{-10} \exp(-2.105/T_4),\ T < 10^4$K $1.5 \times 10^{-10} \exp(-1.400/T_4),\ T > 10^4$K	GP
H10	$H_2 + H \longmapsto 3H$	$1.1 \times 10^{-10}\, T_0^{2.012}(1+0.247\,T_0)^{-3.512} \exp(-3.846/T_4)$	AAZN
H11	$H_2 + e^- \longmapsto 2H + e^-$	$4.4 \times 10^{-10}\, T_0^{0.35} \exp(-1.02/T_5)$	GP

$$+10.80\, T_{\log}^3 - 0.9032\, T_{\log}^4, \qquad (5)$$

where $T_{\log} = \log_{10}(T/\mathrm{K})$.

In addition, we include the following cooling terms for collisional ionization and excitation, and recombination, taken from Haiman, Thoul & Loeb (1996):

$$\frac{\Lambda_{\rm ci}(T)}{\mathrm{erg\,cm^3\,s^{-1}}} = 4.02 \times 10^{-19} \frac{T_5^{\frac{1}{2}}}{1+T_5^{\frac{1}{2}}}\, e^{\frac{-1.578}{T_5}}\, n_{\rm e^-}\, n_{\rm H} \qquad (6)$$

$$\frac{\Lambda_{\rm ce}(T)}{\mathrm{erg\,cm^3\,s^{-1}}} = 7.50 \times 10^{-19} \frac{1}{1+T_5^{\frac{1}{2}}}\, e^{\frac{-1.183}{T_5}}\, n_{\rm e^-}\, n_{\rm H} \qquad (7)$$

$$\frac{\Lambda_{\rm rec}(T)}{\mathrm{erg\,cm^3\,s^{-1}}} = 2.19 \times 10^{-24} \frac{T_6^{0.3}}{1+T_6^{0.7}}\, n_{\rm e^-}\, n_{\rm H^+} \qquad (8)$$

where $T_{\rm n}$ is the temperature in units of $10^{\rm n}$ K.

3 NUMERICAL PROCEDURE

3.1 General approach

The approach that we have adopted for this paper is to specify an initial temperature, ionization level and H_2 fraction and to follow their evolution at constant density, using the relevant cooling terms and reaction rates discussed in the previous section. The assumption of constant density is only valid for halos whose cooling times are much longer than their dynamical times. It is adequate for the purposes of this paper which is to determine the mass of the *first* halos that can cool, but not to follow the collapse of halos once they do cool.

The equations are integrated using the RK4 integrator from Press et al. (1992) modifed to use an adaptive timestep that allowed neither the abundances nor the temperature to vary by more than 0.1 per cent during a timestep. We also tried a Bulirsch-Stoer integrator which gave identical results for a slightly poorer performance. For tests, we integrated simplified networks of equations for which a solution can be obtained analytically: for example, Equations 15 and 16 of T97.

The initial halo parameters are picked to represent the conditions of a virialized object which has collapsed from a 3σ peak in a CDM scenario, with virial temperatures, $T_{\rm vir}$ in the range 100 K to 100 000 K. We consider two different measures of cooling. Firstly, we measure how long it takes halos to cool to $T_{0.75} = 0.75\, T_{\rm vir}$—T97 define an object to have cooled if its temperature decreases by 25 per cent or more in the time that redshift does likewise. We find two distinct populations of clouds: low-mass ones that cool via molecular hydrogen and high-mass ones that cool via electronic transitions. Secondly, we look at the amount by which an object can cool in one dynamical time. Clouds with virial temperatures greater than 9 000 K (i.e. those that have re-ionized) form later but cool much more effectively than lower-mass clouds and have a much greater H_2 fraction once they have cooled to low temperatures.

3.2 Cosmological Models

We compute the collapse redshift of objects arising from 3σ peaks of a CDM power spectrum. The spectrum was calculated using a real-space top-hat window function, the transfer function of Bond & Efstathiou (1984) for scales above $1\,h^{-1}$Mpc and the transfer function of Bardeen et al. (1986), or BBKS, for smaller scales. We choose this combination as the Bond & Efstathiou transfer function is more accurate than BBKS but makes no attempt to accurately calculate the function on scales below $1\,h^{-1}$Mpc.

We present results for three different cosmological models, as listed in Table 2. For comparision with previous work, we first choose the standard, cold dark matter cosmology (SCDM) with power spectrum shape parameter, $\Gamma = 0.43$. When normalised to the COBE results, this power spectrum is now known to have too much power on small scales and so the other two models that we consider use a smaller value, $\Gamma = 0.21$.

The ΛCDM model is our best-guess at the most favourable CDM model. The combination of Ω_0, Ω_{b0} and h naturally produces a power spectrum with the correct value of Γ, and the Hubble parameter and baryon fraction both lie close to currently preferred values (see, for example, Freedman et al. 2001 for the former and Ettori & Fabian 1999 for the latter). The normalisation of the power spectrum, σ_8,

Table 2. Cosmological parameters for the three models: model name; density parameter; cosmological constant in units of $\lambda_0 = \Lambda/3H_0^2$; Hubble parameter in units of $h = H_0/100\,\mathrm{km\,s^{-1}Mpc^{-1}}$; root-mean-square dispersion of the density within spheres of radius $8\,h^{-1}$Mpc; initial ionization fraction; initial H_2 abundance.

Name	Ω_0	λ_0	Ω_{b0}	h	Γ	σ_8	$f_{e,\mathrm{init}}$	$f_{\mathrm{H}_2,\mathrm{res}}$
SCDM	1.0	0.0	0.076	0.5	0.43	0.60	1.58×10^{-4}	1.1×10^{-6}
τCDM	1.0	0.0	0.184	0.5	0.21	0.60	6.52×10^{-5}	1.1×10^{-6}
ΛCDM	0.35	0.65	0.038	0.7	0.21	0.90	1.33×10^{-4}	1.1×10^{-6}

is chosen to reproduce the correct local abundance of rich clusters (Viana & Liddle 1996).

The τCDM model is an attempt to salvage a critical density model in the post-COBE epoch. It has some motivation in decaying neutrino scenarios (e.g. Hansen & Villante 2000) but is mostly phenomenological in nature. For this reason, we have fixed Γ to have the same value as in the ΛCDM model, but have used a lower value of σ_8 to again reproduce the rich cluster abundance. In addition, we have had to take a lower value of the Hubble parameter in order to overcome the age problem, and a baryon density considerably higher than that predicted by primordial nucleosynthesis to give the correct baryon fraction in clusters.

3.3 Initial halo properties

We calculate the initial properties of halos by assuming that they have settled down into virial equilibrium. This will be a good approximation for halos whose cooling time greatly exceeds their dynamical time and should therefore enable us to identify the first objects that are able to cool with reasonable accuracy. To correctly model the properties of halos whose cooling time is shorter than their dynamical time would require a more sophisticated model.

In an isothermal sphere, the density is strongly peaked towards the centre of the halo. Here we follow the evolution of gas which has a density fixed at the mean value within the virialised halo,

$$\rho_{\mathrm{vir}} = \left(\frac{\Delta_c}{\Omega}\right)\rho_{b0}(1+z_{\mathrm{vir}})^3 \tag{9}$$

where z_{vir} is the virialization redshift; ρ_{b0} is the current mean density of baryons in the Universe; and Δ_c is the mean overdensity relative to the critical density within the virialized halo which we take to be $18\pi^2$ for the SCDM and τCDM cosmologies, and $18\pi^2\Omega^{0.45}$ for the ΛCDM cosmology, where Ω is the density parameter at the time of virialization (Eke, Navarro & Frenk, 1998). We take a hydrogen mass fraction (in all its forms: neutral, ionized and molecular) of $X = 0.755$.

For the τCDM and ΛCDM cosmologies, the residual ionization level from the early universe is taken from Peebles (1968, 1993) but divided by two as an attempt to compensate for the neglect of stimulated recombination as suggested by the results of Sasaki & Takahara (1993) and Grachev & Dubrovich (1991). Although dividing by two is somewhat arbitrary, it does result in a more accurate estimate of the initial ionization level than many previous papers without explicitly solving the time-dependent evolution of the ionization level with the inclusion of stimulated recombination. We take

$$f_{e,\mathrm{res}} = 6 \times 10^{-6}\,\Omega_0^{\frac{1}{2}}\,(\Omega_{b0}h)^{-1} \tag{10}$$

which gives the initial ionization levels listed in Table 2.

For halos hotter than about 8 000–9 000 K, the gas will be ionized above the residual value by the process of virialization. For these we use the equilibrium value determined by solving Equation 4 for $\mathrm{d}n_{\mathrm{e}^-}/\mathrm{d}t = 0$:

$$f_{e,\mathrm{eq}} \equiv \frac{n_{\mathrm{e}^-}}{n_{\mathrm{H}} + n_{\mathrm{H}^+}} = \frac{R_1}{R_1 + R_2}. \tag{11}$$

Halos above 28 000 K are almost completely ionized. In this calculation we have neglected the effect of helium: at very high temperatures this would contribute extra electrons so that we would underestimate the electron abundance slightly (by at most 16 per cent).

The background level of H_2 in the Universe is taken to be $f_{\mathrm{H}_2,\mathrm{res}} = 1.1 \times 10^{-6}$ as calculated by Galli & Palla (1998).

We take the initial temperature to be

$$\begin{aligned} T_{\mathrm{vir}} &= \frac{\mu m_H}{k_B}\frac{GM_{\mathrm{tot}}}{2r_{\mathrm{vir}}} \qquad (12)\\ &\approx 40.8\,\frac{\mu}{1.225}\,(1+z_{\mathrm{vir}})\left(\frac{\Delta_c h^2}{18\pi^2\Omega_0}\right)^{\frac{1}{3}}\left(\frac{M_{\mathrm{tot}}}{10^5\,\mathrm{M}_\odot}\right)^{\frac{2}{3}}\mathrm{K}. \end{aligned}$$

Here M_{tot} is the total mass (dark plus baryonic) which we assume is distributed as an isothermal sphere within the virial radius, r_{vir}; m_H is the mass of a hydrogen atom; k_B is the Boltzmann constant; G is the gravitational constant; and μ is the mean mass of particles in units of m_H. We calculate this assuming the ionization level given by Equation 11,

$$\mu = \frac{4\,(R_1 + R_2)}{R_1 + R_2 + X\,(7R_1 + 3R_2)}. \tag{13}$$

μ varies between 1.225 at low temperatures, $T \lesssim 10\,000$ K, and 0.636 at high temperatures, $T \gtrsim 30\,000$ K†.

We shall use the halo virialization redshift, z_{vir}, as our ordinate in most of the plots that follow. This can be converted to temperature or mass in a cosmology-dependent way. We plot these relations in Figure 1 and list examples of the conversion between temperature and other quantities in Table 3. The kink in the temperature profiles correspond to the changing value of μ from Equation 13. Note that halos of a given temperature have similar total masses in each cosmology, although they may virialize at very different redshifts.

† assuming that helium is neutral: for fully-ionized helium μ is lowered to 0.590 so that the error is at most 8 per cent

Table 3. Conversion between virialization temperature, redshift and mass for the three models: temperature; model; virialization redshift; total mass; baryonic mass.

T/K	Model	$z_{\rm vir}$	$M_{\rm tot}/\,{\rm M}_\odot$	$M_{\rm bary}/\,{\rm M}_\odot$
100	SCDM	31.6	4.2×10^3	3.2×10^2
	τCDM	18.0	9.2×10^3	1.7×10^3
	ΛCDM	34.0	2.7×10^3	1.0×10^2
1 000	SCDM	25.0	1.8×10^5	1.4×10^4
	τCDM	14.3	4.1×10^5	7.5×10^4
	ΛCDM	27.3	1.1×10^5	4.2×10^3
10 000	SCDM	18.9	8.7×10^6	6.6×10^5
	τCDM	10.9	1.8×10^7	3.3×10^6
	ΛCDM	20.9	5.5×10^6	2.1×10^5
100 000	SCDM	11.7	1.5×10^9	1.1×10^8
	τCDM	6.8	3.0×10^9	5.5×10^8
	ΛCDM	13.5	8.4×10^8	3.2×10^7

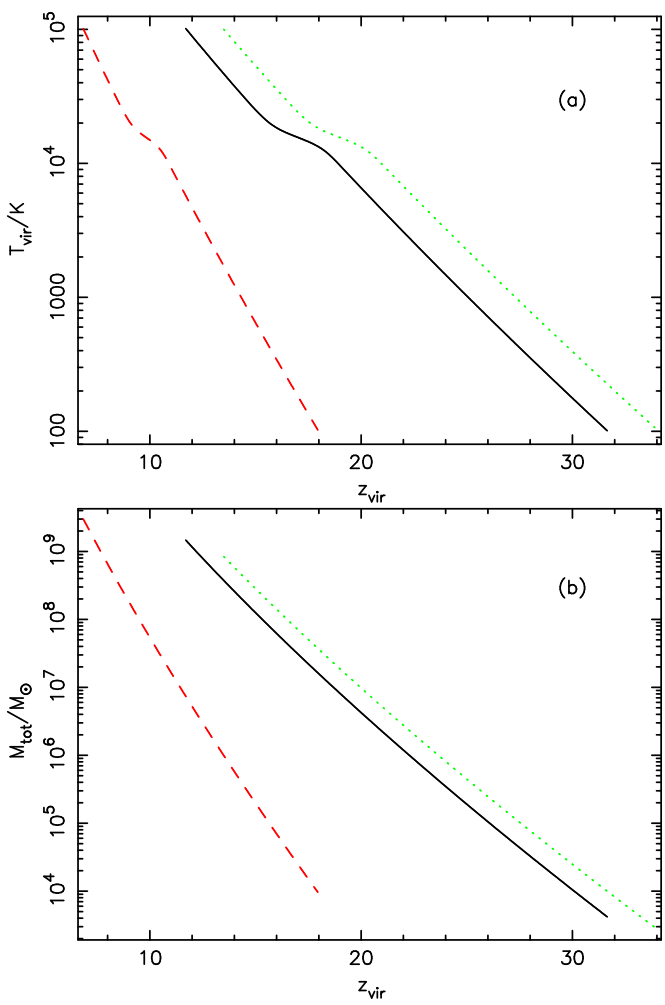

Figure 1. This figure shows the relation between virialization redshift and (a) temperature and (b) total mass for the three cosmologies: SCDM (solid line), τCDM (dashed line), and ΛCDM(dotted line).

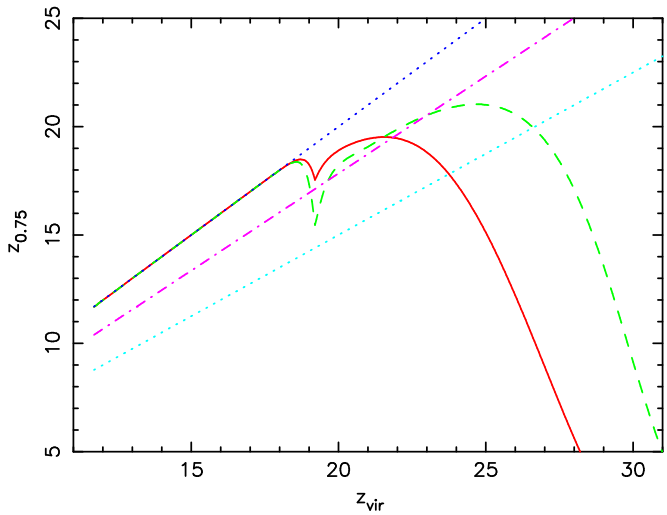

Figure 2. Redshift at which a gas halo will cool to 75 per cent of its initial temperature, $z_{0.75}$, versus virialization redshift, $z_{\rm vir}$, for cosmology SCDMfor two different cooling functions: dashed curve, LS84; solid curve, GP98. The upper and lower dotted lines show the relations $z_{0.75} = z_{\rm vir}$ and $z_{0.75} = 0.75\,z_{\rm vir}$, respectively; the dash-dotted line shows the redshift one dynamical time after virialization.

4 RESULTS

4.1 SCDM

To compare with previous work, we first consider the collapse of halos in the SCDM cosmology. We do this for two different forms of the H_2 cooling function. The first, described above in Equation 5, is from Galli & Palla (1998, hereafter GP98); the second is from Lepp & Shull (1984; hereafter LS84). This latter form is the one used by Abel et al. (1998); it gives an order of magnitude more cooling at temperatures below 1000 K and so favours the collapse of low-mass, low-virial temperature objects.

Initially the gas within the halos is assumed to be shocked to the virial temperature and to be pressure supported. If it can cool significantly within one dynamical time then it will contract towards the centre of the halo in order to maintain pressure support. Accordingly, we plot in Figure 2 the redshift, $z_{0.75}$, at which the gas (if it maintains constant density) will lose 25 percent of its initial energy. The solid curve was generated using the GP98 cooling function and the dashed curve using that of LS84. The lower dotted line shows the relation $z_{0.75} = 0.75\,z_{\rm vir}$ which is the condition used by T97 to separate clouds which can cool from those which cannot. The dashed line shows the redshift one dynamical time, $t_{\rm dyn}$, after the time of virialization, $t_{\rm vir}$, where we set $t_{\rm dyn} = t_{\rm vir}/4\sqrt{2}$.

We will first concentrate on the GP98 curve. This intersects the line $z_{0.75} = 0.75\,z_{\rm vir}$ at $z_{\rm vir} \approx 23.8$ which means that the smallest clouds that can cool according to the T97 criterion will have virial temperatures of $T_{\rm vir}$=1 600 K and masses of $4\times10^5\,{\rm M}_\odot$. The most important quantity for comparison with the results of T97 is the temperature which agrees very well with that given in their Figure 5. For a virialization redshift of 23.8, they require a slightly higher

temperature of 2 500 K for collapse, but this difference is attributable to our having a slightly higher baryon density and H_2 cooling rate.

T97 assume that the smallest halos that can cool will be the first to do so, but that is not the case. As one moves to lower $z_{\rm vir}$ and higher $T_{\rm vir}$, the efficiency of H_2 production increases and the cooling time decreases. Consequently, the first objects to cool to 75 per cent of their initial temperature virialise later at $z_{\rm vir} \approx 21.6$, and have higher virial tempertures (3 600 K) and masses ($1.6 \times 10^6\,{\rm M}_\odot$). This maximum in $z_{0.75}$ corresponds roughly to the redshift one dynamical time after virialisation.

At higher temperatures the efficiency of H_2 cooling continues to increase but cannot compensate for the later virialization redshift and so $z_{0.75}$ decreases once more. Note the cusp in the curve at $z_{\rm vir} \approx 19.2$, corresponding to $T_{\rm vir} \approx 9\,000\,K$. As this temperature is approached, H_2 starts to be destroyed and so its abundance plummets and the cooling time increases. However, at higher temperatures collisional excitation and ionization take over and these are much more efficient cooling processes so the cooling time drops once more. In fact for temperatures above 11 000 K the cooling is essentially instantaneous (i.e. the cooling time much less than the dynamical time).

Thus the two peaks in Figure 2 correspond to two different classes of object: the first to collapse, at $z_{0.75} = 19.5$, are dominated by molecular cooling and the second, at $z_{0.75} = 18.5$ by electronic transitions. As we are interested in the first objects to cool, it might be thought that the smaller halos are the more important, but this is not entirely clear because the T97 criterion does not trace the cooling down to very low temperatures. In fact the more massive halos have a large residual ionization which results in a greater H_2 production once they have cooled below 9 000 K, and so they may be the first objects to cool to really low temperatures. We will discuss this further in Section 4.2, below.

Turning now to the LS84 curve, we see that this is qualitatively similar. However, it predicts shorter cooling times for low-mass clouds because the H_2 cooling rate is much higher. Hence the smallest objects that can collapse have a lower mass and temperature than for the GP98 cooling function. The lowest-mass objects that satisfy the T97 criterion virialize at a redshift of 26.6 and have virial temperatures of 570 K—this seems to agree reasonably well with the predictions from Figure 12 of Abel et al. (1998). The peak of the curve has moved to $z_{\rm vir} \approx 24.7$ and the temperature and mass of these first objects are significantly lower, 1 100 K and $2.3 \times 10^5\,{\rm M}_\odot$, respectively. By contrast, the properties of the halos corresponding to the higher-mass peak are only slightly modified.

We will use the more up-to-date GP98 cooling function throughout the rest of this paper.

4.2 τCDM

We now turn to the τCDM cosmology. Like the SCDM cosmology, this has a critical density of matter and a high baryon density, but it has a spectral shape that gives less power on small scales and hence the first objects collapse at much lower redshift. This is shown in Figure 3 which is the analogue of Figure 2 but for the τCDM cosmology. We can see that gas halos first manage to cool to $0.75\,T_{\rm vir}$ only

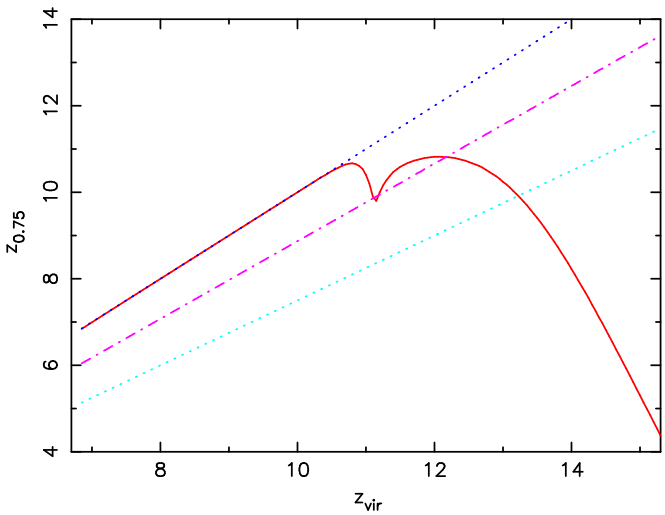

Figure 3. The solid line shows the redshift at which a gas halo will cool to 75 per cent of its initial temperature, $z_{0.75}$, versus virialization redhsift, $z_{\rm vir}$, for cosmology τCDM using the GP98 cooling function. The upper and lower dotted lines show the relations $z_{0.75} = z_{\rm vir}$ and $z_{0.75} = 0.75\,z_{\rm vir}$, respectively; the dash-dotted line shows the redshift one dynamical time after virialization.

at redshift $z_{0.75} = 12.1$: these have virial temperatures of 4 500 K and total masses of $5.0 \times 10^6\,{\rm M}_\odot$.

In order to form stars, the gas has to cool to much less than $T_{0.75}$. Therefore we consider a second measure of cooling: Figure 4 shows the final temperature of the halo after one dynamical time (once again assuming constant density—this will be a good assumption only for clouds which cool by a small amount in a dynamical time; for others it will underestimate their cooling rate). The coldest clouds are those to the right-hand-side of the Figure, but this is only because they were born with low temperatures: they have cooled very little. The final temperature is an increasing function of virial temperaure up to about 9 000 K when it suddenly plummets, so that somewhat surprisingly, the larger, higher virial temperature halos, have cooled to a lower temperature than the smaller ones. The reason for this is shown in Figure 5 which shows that the H_2 abundance after one dynamical time is larger in high-mass halos than low-mass ones. This is because they become highly ionized and their residual ionization once they have cooled back down below 9 000 K is greater than that of low-mass halos, which in turn results in a greater production of H_2. Note that the cooling rate of these high-mass halos will continue to be greater than that of low-mass halos even after one dynamical time. It is therefore conceivable that they will be the first to cool to really low temperatures. Because the gas is free to move around on this timescale, however, realistic dynamical simulations are required to model the problem accurately.

4.3 ΛCDM

This cosmology is currently the most popular CDM model. Halos virialize at higher redshift than in τCDM but qualitatively, their behaviour is just the same. Figures 6–8 mimic

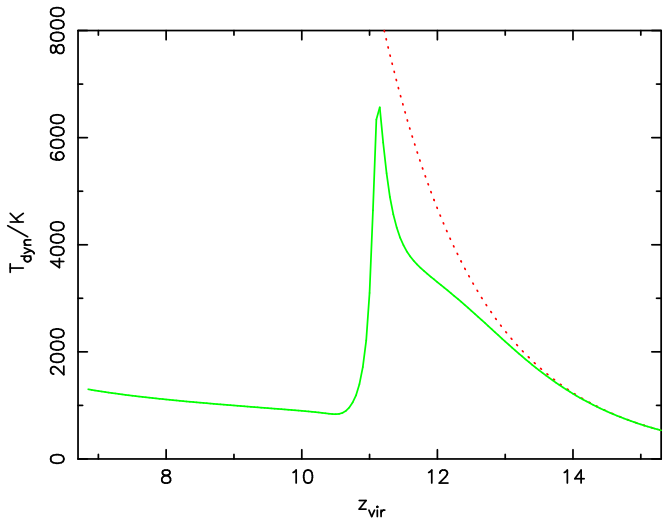

Figure 4. Temperature obtained after one dynamical time versus virialization redshift, for cosmology τCDM. The dotted line shows the initial, virial temperature of the halo. The cusp in the curve corresponds to a virial temperature of 9 000 K.

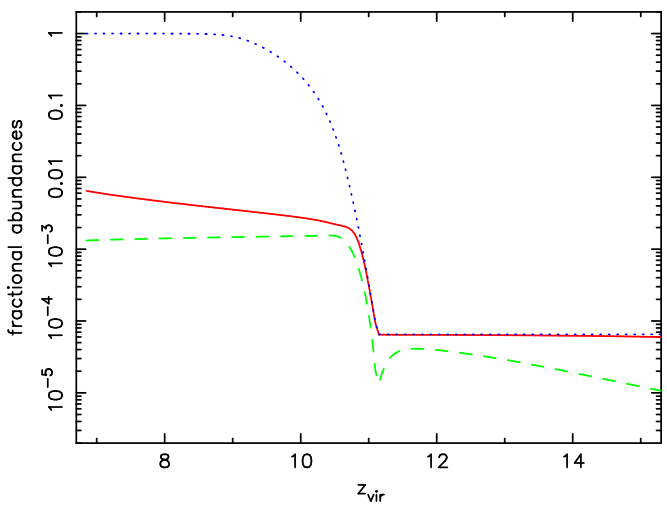

Figure 5. Ionization fraction (solid line) and H_2 abundance (dashed line) after one dynamical time, for cosmology τCDM. The dotted line shows the maximum ionization level of the halo.

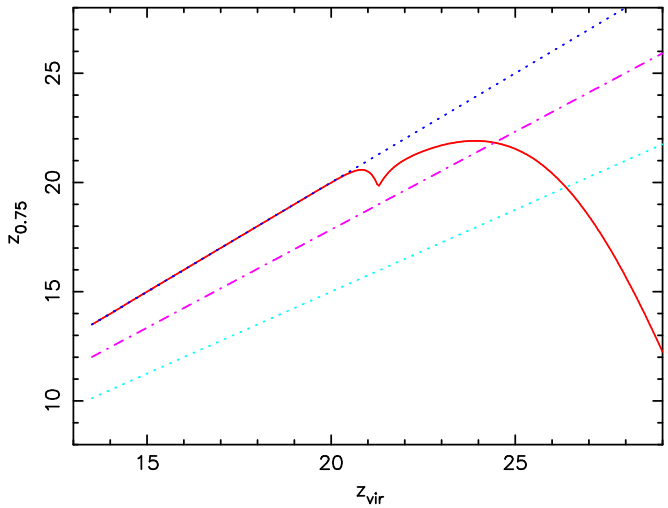

Figure 6. The solid line shows the redshift at which a gas halo will cool to 75 per cent of its initial temperature, $z_{0.75}$, versus virialization redhsift, $z_{\rm vir}$, for cosmology ΛCDMusing the GP98 cooling function. The upper and lower dotted lines show the relations $z_{0.75} = z_{\rm vir}$ and $z_{0.75} = 0.75\, z_{\rm vir}$, respectively; the dash-dotted line shows the redshift one dynamical time after virialization.

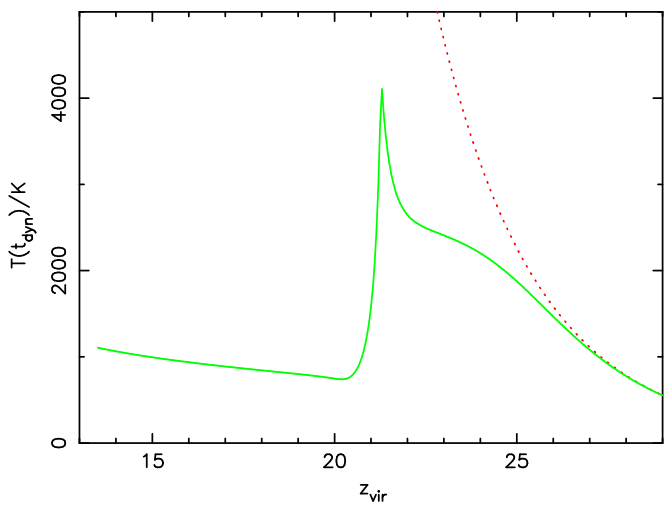

Figure 7. Temperature obtained after one dynamical time versus virialization redshift, for cosmology ΛCDM.

those of Figures 3–5 and show all the same features. In this cosmology the first halos to cool to $T_{0.75}$ have virial temperatures of 3 400 K, masses of 8.6×10^5 $M_\odot$ and cool to 75 per cent of their virial temperature at $z_{0.75} \approx 21.9$; the second generation of higher mass halos cool at $z \approx 20.6$ and have virial temperatures and masses of 10 400 K and 5.7×10^6 $M_\odot$.

5 DISCUSSION

The properties of the halos corresponding to the two peaks in Figures 2, 3 and 6 are summarized in Table 4.

Note that, although the redshift at which the first objects cool is quite different in different cosmologies, their virial temperatures are very similar, reflecting the similar chemical processes that are occurring within them.

5.1 The effect of cosmology

Although it is often not stated explicitly, previous work seems to concentrate on the SCDM cosmology which is now known to be a poor model of the Universe: the current favoured model being ΛCDM. Coincidentally these two give similar formation redshifts for the two generations of halos.

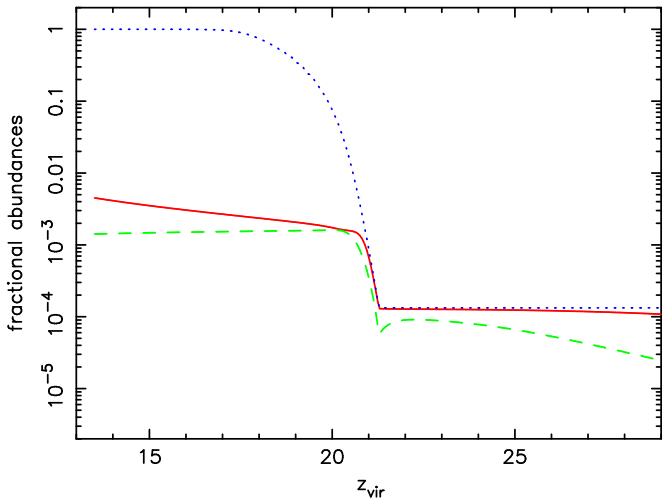

Figure 8. Ionization fraction (solid line) and H_2 abundance (dashed line) after one dynamical time, for cosmology ΛCDM. The dotted line shows the maximum ionization level of the halo.

Table 4. Properties of the two generations of halos to form in each of the cosmologies (corresponding to the molecular hydrogen and electron-cooling peaks of the curves in Figures 2, 3 and 6): cosmological model; redshift at which the halo cools to 75 per cent of the virial temperature; virial temperature; total mass of the halo; baryonic mass of the halo.

Model	$z_{0.75}$	$T_{\rm vir}$/K	$M_{\rm tot}/\,{\rm M}_\odot$	$M_{\rm bary}/\,{\rm M}_\odot$
Generation 1				
SCDM	19.5	3 600	1.6×10^6	1.2×10^5
τCDM	10.8	4 500	5.0×10^6	9.2×10^5
ΛCDM	21.9	3 400	8.6×10^5	3.3×10^4
Generation 2				
SCDM	18.5	10 800	9.9×10^6	7.5×10^5
τCDM	10.7	10 600	2.1×10^7	3.9×10^6
ΛCDM	20.6	10 400	5.7×10^6	2.2×10^5

The phenomenologically-based critical-density τCDM cosmology, on the other hand, forms its first objects at much lower redshifts, but nevertheless still comfortably above the observed redshifts of the most distant objects yet observed.

The baryonic masses of these first cooled halos are a factor of 30 times larger in the τCDM than in the ΛCDM cosmology—in the former they correspond roughly to the mass of globular clusters. Note, however, that the difference in cooling redshift for halos differing by a factor of 10 in mass is not huge and so we would expect a wide spread in masses for these first objects.

5.2 The H_2 mass fraction required for collapse

We have shown in Section 4.1, that the virial temperature required for collapse agrees with the results from T97 and Abel et al. (1998). These two papers both claim that the H_2 fraction required for collapse of the first halos is 5×10^{-4}, a result confirmed by Fuller & Couchman (2001). However, the H_2 fractions shown in Figures 4 and 7 are much lower, about 1×10^{-4}. There is no contradiction here, however. The Figures show the H_2 fraction after one dynamical time, whereas T97 give the asymptotic H_2 fraction at late times—for which we get a similar value of 4.4×10^{-4} for the SCDM cosmology (similarly 4.2×10^{-4} for ΛCDM and a slightly lower value of 3.2×10^{-4} for τCDM).

Table 5. The redshift, $z_{0.75}$ and age, $t_{0.75}$, of the universe at the time at which the first halos in each generation have cooled to 75 per cent of the virial temperature. The final column gives the difference in age between the two generations.

Model	$z_{0.75}$	$t_{0.75}$/Myr	$z_{0.75}$	$t_{0.75}$/Myr	Δt/Myr
	Generation 1		Generation 2		
SCDM	19.5	158	18.5	170	12
τCDM	10.8	361	10.7	368	7
ΛCDM	21.9	143	20.6	157	14

5.3 The relative roles of the two generations of halos

It is clear from Figures 4 and 7 that the second generation of halos, whose virial temperatures exceed about 9 000 K, cool much more efficiently down to low temperatures than the first generation. This is because they have an excess of free electrons which leads to an enhanced production of H_2—as shown in Figures 5 and 8. Nevertheless, it is the first objects to cool that we are interested in and so should we not identify these with the first generation of halos? In addition, it must not be forgotten that we have presented results for $3\,\sigma$ density fluctuations. In reality there will be fluctations with a continuous range of overdensities. In particular, a $3\,\sigma$, high-mass object will contain even larger fluctuations on smaller scales. The inclusion of density profiles may raise the core density of even small halos to a point where free-fall collapse could proceed.

Despite the above arguments, there are two reasons why we think that larger objects may have an important role: (i) for the masses we are dealing with, the power spectrum is almost flat—in this situation Press-Schecter theory (Press & Schechter 1974) predicts that there is as much collapsed mass in each decade of mass as any other, and (ii) the difference in collapse time between a big object (Generation 2) and a small object (Generation 1) is similar to the lifetime of the most massive stars (see Table 5—a 15 solar mass star has a lifetime of 10 Myr). This means that even if small objects are the first to collapse, larger objects may collapse around them and form stars before the first generation of supernovae explode (this neglects any consequences of ionizing radiation). This first of these points can be addressed using a statistical model of the collapse of halos but the second requires numerical simulations of the dynamics of the collapse.

5.4 Towards more realistic models

Our model does not attempt to follow the increasing density in objects that do cool on less than a dynamical time. To do so requires the use of some form of hydrodynamics code. Several groups have attempted this. Haiman, Thoul & Loeb (1996), Omukai & Nishi (1998) and Nakamura &

Umemura (1999) have all simulated spherically-symmetric collapses; Fuller & Couchman (2000) go further and simulate a top-hat collapse with and without substructure using a three-dimensional code. In general, these results confirm the simple analytic predictions.

One of the most important simplifications of our model is that it ignores substructure. Any halo that can cool and collapse on a dynamical time will contain smaller fluctuations of even higher overdensity. This will result in fragmentation of the cloud as the collapse proceeds. This has been investigated by Abel et al. (1998) and Abel, Bryan & Norman (2000) who have used a hierarchical grid code to perform very high-resolution simulations of the collapse and fragmentation of the first objects, down to scales of just 1 pc. They find a very filamentary structure develops with the first stars forming from small knots at the intersection of filaments. Only a small fraction of the gas reaches sufficiently high densities to allow star-formation. If these first stars were to feed back energy into the surrounding cloud and disrupt it, then that would suggest that the size of the first star clusters may be much smaller than the size of the cloud from which they formed. However, the masses of these star clusters in the simulations is small and may not produce many high-mass stars. Also, the simulations have not yet been run for long enough to see whether the surrounding cloud will collapse before feedback becomes effective, as we have suggested in Section 5.3. It should be noted that very different results were obtained by Bromm, Coppi & Larson (1999) using a particle-based hydrodynamics, but with much poorer resolution. They found that the cloud collapsed to a rotationally-supported disk which then broke up into very massive star clusters. It will be a while, we suspect, before any concensus emerges.

Fuller & Couchman (2000) simulated a cubical region of side $25\,h^{-1}$kpc in the SCDM cosmology with several different realisations of the density fluctuation spectrum, using an N-body, hydrodynamics code. This random realisation is just what is required to look at the relative importance of halos of different mass. They found that the most massive objects that collapsed within the region did so at a wide range of redshifts, 15–30, and had a similar large spread in mass. Unfortunately, their simulation volume was not large enough to sample our second-generation halos.

6 CONCLUSION

In this paper we have considered the cooling of gas within spherical, virialized halos in the high-redshift Universe. Our technique is similar to that used by T97, but with a more up-date cooling function and cosmological model. In addition, we have investigated halos with a wider range of virial temperatures, in the range of 100 K to 100 000 K.

We have followed the abundances of H_2 and H^+/e^-without including Helium chemistry but using a more complete destruction term for H_2, taking into account the relative importance of the reverse reaction H7 which other authors have ignored. This makes a difference of a few per cent to the final H_2 abundances for halos which cool from above 9 000 K.

The main coolant for temperatures below 9 000 K is molecular hydrogen. Unfortunately, the cooling rate seems poorly-known and a recent determination by Galli and Palla (1998) gives values at temperatures below 1 000 K over an order of magnitude lower than those of Lepp and Shull (1984). Consequently the first objects to cool in the former do so later and have much higher masses and virial temperatures.

We follow T97 by defining clouds to have cooled if they lose 25 per cent or more of their energy in the time that the redshift has decreased by 25 per cent. We obtain similar virial temperatures for the smallest halos that can cool, but note that the *first* objects to cool in each cosmology are more massive and have higher virial temperatures of about 4 000 K. In the SCDM and ΛCDM cosmologies the formation redshift is $z_{0.75} \approx 20$; for the τCDM cosmology it is much lower, $z_{0.75} \approx 11$.

We identify a second generation of halos that cool about 10 Myr after the first one. These are halos with virial temperatures in excess of 9 000 K for which there is a significant fraction of free electrons. The cooling is dominated by electronic transitions at high temperatures and is almost instantaneous, occurring on much less than a dynamical time. Just as significant, however, is the fact that the residual ionization is greater than in low-mass halos and the production of H_2 is much greater. Consequently, they cool to much lower temperatures in a dynamical time than do the first generation of halos.

Our model suffers from three deficiencies: we consider only 3σ density fluctuations, we ignore substructure, and we cannot follow the collapse of halos whose cooling times are shorter than their dynamical times. More sophisticated studies are required to determine whether the first or second generation halos are more important for determining the mass of the first star-clusters, or whether they both have a role to play. We hope to report on these in future papers.

7 ACKNOWLEDGEMENTS

PAT is a PPARC Lecturer Fellow. We are grateful to Todd Fuller for helpful discussions. FS thanks his parents for their support.

REFERENCES

Abel T., Anninos P., Zhang Y., Norman M. L., 1997, NewA, 2, 181
Abel T., Anninos P., Norman M. L., Zang Y., 1998, ApJ, 508, 518
Abel T., Bryan G., Norman M., 2000, ApJ, 540, 39
Bardeen J. M., Bond J. R., Kaiser N., Szalay, A. S., 1986, ApJ, 304, 15
Bond J. R., Efstathiou G. P., 1984, ApJ, 285, L45
Bromm V., Coppi P. S., Larson R. B., 1999, ApJ, 527, 5
Nakamura F., Umemura M., 1999, ApJ, 515, 239
Eke V. R., Navarro J. F., Frenk C. S., 1998, ApJ, 503, 569
Ettori S., Fabian A. C., 1999, MNRAS, 305, 834
Forrey R. C., Balakrishnan N., Dalgarno A., Lepp S., 1997, ApJ, 489, 100
Freedman W. L., et al., 2001, ApJ, in press (astro-ph/0012376)
Fuller T. & Couchman H. M. P., 2000, ApJ, 544, 6
Galli D. & Palla F., 1998, AA, 335, 403
Grachev S. I., Dubrovich V. K., 1991, Astrophysics, 34, 124
Haiman Z., Thoul A. A., Loeb A., 1996, ApJ, 464, 523
Hansen S. H., Villante F. L., 2000, PhysLettB, 486, 1

Lepp S., Shull J. M., 1984, ApJ, 280, 465
Martin P. G., Schwarz D. H. & Mandy M. E., 1996, ApJ, 461, 265
Omukai K., Nishi R., 1998, ApJ, 508, 141
Peebles P. E. J., 1968, ApJ, 153, 1
Peebles P. E. J., 1993, Principles of physical cosmology, Princeton UP
Press W. H., Schechter P. G., 1974, ApJ, 187, 425
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recepies in Fortran 2nd edition, CUP
Sasaki S., Takahara F., 1993, PASJ, 45, 655
Stancil P. C., Lepp S., Dalgarno A., 1998, ApJ, 509, 1
Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1 (T97)
Viana P. T. P., Liddle A. R., 1996, MNRAS, 281, 323

The appendices describes two different ways of deriving Equation 1 for the rate of change of abundance of H_2. We do this because nowhere in the literature does it seem to be spelt out in detail and because our equation differs slightly from those used previously, as described in Section 2.1.

APPENDIX A: DERIVATION VIA REMOVAL OF H^- AND H_2^+ FROM THE REACTION NETWORK

The reaction rates for destruction of H^- and H_2^+ are much greater than for formation or destruction of H, H^+, e^- and H_2. Hence we can regard H^- and H_2^+ as short-lived species that are destroyed the instant that they are produced. This has the advantage that we can eliminate them from the reaction network, as described below.

Consider first H^-. This has two destruction channels, Reactions H4 and H5 in Table 1, whose combined rate is

$$R_{H^-} = n_H R_4 + n_{H^+} R_5, \tag{A1}$$

where R_4 and R_5 are the reaction rates as listed in the Table (at redshifts greater than 110 photo-destruction of H^- by cosmic microwave background photons is also important, but we do not consider such high redshifts in this paper). Hence the fractions of H^- that decay via reactions H4 and H5 are $n_H R_4/R_{H^-}$ and $n_{H^+}\ R_5/R_{H^-}$, respectively. Reaction H4 is always important, whereas Reaction H5 is significant only at high ionization levels.

Now Reaction H3 for H^- production can instead be rewritten as two different reaction chains:

$$\text{H3H4} \qquad 2H + e^- \longmapsto H_2 + e^- + \gamma \tag{A2}$$

$$\text{H3H5} \qquad H + e^- + H^+ \longmapsto 2H + \gamma \tag{A3}$$

that occur at rates ${n_H}^2 n_{e^-} R_3 R_4/R_{H^-}$ and $n_H n_{e^-} n_{H^+} R_3 R_5/R_{H^-}$, respectively. The first of these reaction chains forms molecular hydrogen using electrons as catalysts whereas the second leads to a reduction in the ionization level.

We can treat H_2^+ in the same manner. Its decay channels are Reactions H7 and H8 which sum to a total destruction rate of

$$R_{H_2^+} = n_H R_7 + n_{e^-} R_8. \tag{A4}$$

Once again we have neglected photo-destruction which is important only at very high redshifts.

Then Reaction 6 becomes the two reaction chains

$$\text{H6H7} \qquad 2H + H^+ \longmapsto H_2 + H^+ + \gamma \tag{A5}$$

$$\text{H6H8} \qquad H + H^+ + e^- \longmapsto 2H + \gamma \tag{A6}$$

which are analogous to those for H^-. There is a second couplet starting with H_2:

$$\text{H9H7} \qquad H_2 + H^+ + H \longmapsto H_2 + H^+ + H \tag{A7}$$

$$\text{H9H8} \qquad H_2 + H^+ + e^- \longmapsto 3H. \tag{A8}$$

Of these, H9H7 is the more important at low ionization levels and will strongly suppress destruction of H_2 via collisions with protons.

Using the reactions listed in Table 1, but replacing Reactions H3 through H9 with the reaction chains derived above, we can write down an equation for the rate of change of H_2 abundance:

$$\frac{dn_{H_2}}{dt} = {n_H}^2 \left[n_{e^-} \frac{R_3 R_4}{R_{H^-}} + n_{H^+} \frac{R_6 R_7}{R_{H_2^+}} \right] - n_{H_2} \left[n_{H^+} n_{e^-} \frac{R_9 R_8}{R_{H_2^+}} + n_H R_{10} + n_{e^-} R_{11} \right]. \tag{A9}$$

APPENDIX B: DERIVATION USING EQUILIBRIUM VALUES OF H^- AND H_2^+

The second derivation uses the fact that the destruction rates for H^- and H_2^+ are high to derive equilibrium values for n_{H^-} and $n_{H_2^+}$ which can then be eliminated from the equations.

Consider first the equation for the rate of change of n_{H^-}:

$$\frac{dn_{H^-}}{dt} = n_H n_{e^-} R_3 - n_{H^-} \left[n_H R_4 + n_{H^+} R_5 \right]. \tag{B1}$$

The destruction rates within the square brackets are very large which means that n_{H^-} will rapidly evolve to the equilibrium value in which creation and destruction of n_{H^-} balance and $dn_{H^-}/dt \approx 0$. Then we have

$$n_{H^-} \approx \frac{n_H n_{e^-} R_3}{n_H R_4 + n_{H^+} R_5} \equiv \frac{n_H n_{e^-} R_3}{R_{H^-}}. \tag{B2}$$

Similarly for $n_{H_2^+}$ we have

$$n_{H_2^+} \approx \frac{n_H n_{H^+} R_6 + n_{H_2} n_{H^+} R_9}{n_H R_7 + n_{e^-} R_8} \equiv \frac{n_H n_{H^+} R_6 + n_{H_2} n_{H^+} R_9}{R_{H_2^+}}. \tag{B3}$$

We can now write down the equation for the rate of change of H_2, then substitute for n_{H^-} and $n_{H_2^+}$.

$$\begin{aligned} \frac{dn_{H_2}}{dt} &= n_H n_{H^-} R_4 + n_H n_{H_2^+} R_7 - n_{H_2} \left[n_{H^+} R_9 + n_H R_{10} + n_{e^-} R_{11} \right] \qquad \text{(B4)} \\ &= {n_H}^2 n_{e^-} \frac{R_3 R_4}{R_{H^-}} + {n_H}^2 n_{H^+} \frac{R_6 R_7}{R_{H_2^+}} + n_H n_{H_2} n_{H^+} \frac{R_9 R_7}{R_{H_2^+}} - n_{H_2} \left[n_{H^+} R_9 + n_H R_{10} + n_{e^-} R_{11} \right] \\ &= {n_H}^2 \left[n_{e^-} \frac{R_3 R_4}{R_{H^-}} + n_{H^+} \frac{R_6 R_7}{R_{H_2^+}} \right] \qquad \text{(B5)} \end{aligned}$$

$$- n_{\mathrm{H}_2} \left[n_{\mathrm{H}^+} n_{\mathrm{e}^-} \frac{R_9 R_8}{R_{\mathrm{H}_2^+}} + n_{\mathrm{H}} R_{10} + n_{\mathrm{e}^-} R_{11} \right].$$

Equation B5 is identical to Equation A9 derived above.