Hannah Krug
ASTR 688R
Spring 2008
Final Project
Due 5/13/08

# Reionization of the Intergalactic Medium: What Is it and When Did it Occur?

In the time following the Big Bang, there are two epochs which astronomers can consider arguably the most important cosmologically: the epoch of recombination and the epoch of reionization. But while recombination, the time when free protons and electrons coupled to form neutral hydrogen throughout the universe, has been well studied and pinpointed by WMAP (as we have discussed thoroughly in class), determining when reionization, the period in which the universe switched from neutral hydrogen back to an ionized plasma, occurred has proven much more elusive. Much work is currently being done to try to solve this mystery, including here at Maryland; the NEWFIRM instrument was developed in part for this purpose and I plan to make use of it for my PhD thesis. In this paper, I will outline the importance of the reionization epoch and the basic work being done to attempt to determine when this important epoch occurred.

In current cold dark matter cosmology, the universe is believed to have formed in a "bottom-up" hierarchy: the smallest objects formed first and large scale structures such as galaxy clusters and filaments formed at later times, with all structures stemming originally from density perturbations. Once the original stars formed, it is believed that those objects (or others, as will be described) provided the energy which ionized neutral hydrogen and allowed the universe to become transparent to photons once again, thus ending the cosmological "dark ages" (Loeb & Barkana 2001 reionization review, hereafter LB). The starting point to probing this period of reionization comes from the fact that the ionization energy of hydrogen is 13.6 electron volts and that the nuclear fusion reactions that take place in the center of stars release on the order of one hundred thousand times that energy amount; the primary assumption is that stars alone could have been responsible for this reionization epoch.

Had reionization never occurred and space stayed opaque to photons, the universe would be a very different place and large scale structure may never have been able to develop, and thus pinpointing when this epoch took place is of great interest to astronomers. In 2001, astronomers believed reionization to have taken place sometime between redshifts 30 and 7, a wide range, one which more recent efforts have attempted to narrow (LB). One result that has helped to narrow down the predicted reionization redshift range comes from the important Wilkinson Microwave Anisotropy Probe, or WMAP. Not only did WMAP constrain important cosmological parameters (the omega values which we have used in class and which are used in all the studies which will be described here), but from its measurements of the cosmic microwave background, or CMB, a neutral universe is inferred for redshifts approximately greater than or equal to 14 (Cantalupo, Porciani, & Lilly 2007). A rough lower limit can be determined by applying the Gunn-Peterson test to quasars. With the Lyman-α atomic transition, even a region with a small neutral hydrogen number density can show a strong resonant scattering feature which simulates absorption at the transition line. One would normally expect a sharp extinction drop between the long-wavelength and short-wavelength sides of the Lyman-α emission line, but due to this resonant scattering feature, we do not observe that. The Gunn-Peterson test can thus be used to provide an upper limit for the neutral hydrogen density (Coles & Lucchin 2002). Applying the Gunn-Peterson test to quasars at around redshift 6 has shown a very low neutral hydrogen fraction at that time, implying that the universe has already been ionized (Cantalupo et al. 2007).

Knowledge of how exactly reionization occurred can help cosmologists more accurately pinpoint when it occurred. In standard reionization theory, an "intergalactic ionizing radiation field" is assumed to have spread throughout the universe, ionizing hydrogen as it propagated. Astronomers are thus faced with a major problem from the start, because this ionizing field is made up of ionizing radiation which has escaped from stars and quasars, but determining the fraction of ionizing radiation that escapes from these stars and quasars is very difficult and has not yet been well constrained (LB). Attempts to constrain this escape fraction have included looking at theoretical models of the Milky Way's own O and B stars (since the hottest stars would have been those which produced the ionizing radiation), giving 3-14% escape fraction, measuring H-α lines in the Magellanic Stream, giving 6% escape fraction, looking at nearby starburst galaxies, giving a range between 3 and 57 percent, and making the assumption that galaxies are isotropic point sources, which yields a 5-60% escape fraction (LB). This extremely wide range, 3-60%, is a problem in and of itself, but becomes even worse due to the fact that the distribution of gas in these galaxies is thought to play a major role in determining the escape fraction. High redshift stars and galaxies that would be producing the ionization fronts may have very different gas distributions than galaxies in our local universe, so the escape fraction could be much lower (LB). The escape fraction used in reionization models plays a major role in the calculation of the reionization redshift, so better constraints on that fraction are crucial to a theoretical understanding of the reionization epoch. More recent studies have limited this escape fraction to under 20%, based on Lyman-α opacity evolution and the amount of ultraviolet radiation found in the cosmic background (Ciardi, Ferrara, & White 2003).

The reionization epoch is not believed to have occurred all at once, but rather believed to have progressed through a series of stages. In the beginning, individual sources would begin producing ionization radiation during the "pre-overlap" phase, and it is thought that these first sources would come from regions with the highest density and most massive dark matter halos, according to Gnedin as cited in Loeb & Barkana. The ionization fronts move slowly at first, as they are originating in high-density environments, and then progress faster once they reach regions of low density. The IGM is considered to be a "two-phase medium" at this point because it is composed of ionized hydrogen and regions of neutral, high-density hydrogen. This is enhanced by the fact that ionization does not occur at the same rate in every region, since ionization fronts are being produced by sources with different temperatures and luminosities. During this phase, each ionization front produces an H II "bubble" surrounding the galaxy from which the ionization front originated and with a radius determined by the mass of the interior halo (LB). Next comes the "overlap" phase, which occurs when these H II bubbles begin to cross each other and nearly every point in the universe becomes subjected to ionizing radiation from multiple fronts, and most of the neutral, high-density regions become ionized. Since the hydrogen in the universe has become much lower density by this point, the ionization fronts can move faster and with higher intensity, so this phase is believed to happen on much shorter timescales than the first phase, perhaps taking less than a Hubble time. The universe would also have become much more transparent, so photons from various galaxies can travel farther, begin reaching each other, and increase the rate of galaxy formation throughout the universe (LB). According to the Loeb and Barkana review, the actual "moment of reionization" occurs when the overlap phase ends; the low density IGM has all been highly ionized and the only remaining neutral locations are regions of gas located inside of high density clouds. This moment of reionization is what astronomers are striving to determine. The "post-overlap" phase follows this and continues for a great amount of time. During post-overlap, there is gradual ionization of the neutral high density regions, and it takes so long because neutral gas can be found in all regions of density perturbation collapse – the post-overlap phase thus continues today. One other important event, "breakthrough," happens at a redshift of approximately 1.6 according to Madau, when all ionization sources can see all other ionization sources (LB).

There are several parameters which must be pinned down to properly describe these phases in theoretical models. First is the "clumping factor," which describes the ability of free protons and electrons to recombine (as we see at the time of the CMB formation) and thus enabling the hydrogen to stay neutral. If this clumping factor is high, as is often the case in high density regions, then the ionization fronts will not propagate as quickly and the moment of reionization will happen later. The escape fraction is also key, as has already been described, as is the formation efficiency of the ionizing sources. If stars and quasars form slowly, that could again slow the ionization front propagation (LB). If one assumes the simplest possible ionization scenario of instant IGM reionization throughout the entire universe, with optimal escape fraction and low clumping factor, and combines that with the WMAP result of an optical depth of scattering of tau equal to 0.17 +/- 0.04, then reionization must have occurred at a redshift of 17 +/- 5 (Benson, Sugiyama, Nusser, & Lacey 2006). Although it is fairly obvious that the intergalactic medium of the entire universe could not have become reionized simultaneously, this provides an updated upper limit on the epoch redshift, more accurate than the upper limit of 30 predicted back in 2001.

Modeling the IGM ionization fronts also depends greatly on what objects are producing the ionizing radiation. Much of what has already been discussed assumes that it is stars that are primarily responsible for this radiation. If it is quasars that are doing most of the radiation, however, some different effects will be observed. Rather than the ionizing radiation being spread throughout the host galaxy, it will emerge from just one point and can thus escape from the galaxy more quickly in an "H II funnel." This might lead to a more efficiently propagating ionization front. Additionally, the ionizing photons from quasars are harder than those from stars and so they can drive farther through the neutral medium and create an ionization front that is much thicker than a solely stellar one (LB).

Determining whether the reionization of the intergalactic medium was a result of stars, quasars, or a mix of the two can have ramifications not only on the redshift when the moment of reionization occurred but also on galaxy formation after this epoch. The propagating radiation fronts not only ionized the hydrogen in the universe but in doing so, the radiation is believed to have heated the temperature of the IGM to greater than 10000 degrees Kelvin. This heating would have led to a dramatic increase in the Jeans mass throughout the universe – several orders of magnitude according to Loeb and Barkana's review – and and thus influenced the minimum mass of new galaxies (raising it!), dropping the overall collapsed mass by as much as 50%. If it is quasars that are doing the ionization and thus the heating, however, the new intergalactic temperature could be double that of the new temperature caused solely by stars, again modifying the Jeans mass.

In addition, this intergalactic medium reheating is thought to suppress star formation. This leads to one important observational test for determining the moment of reionization – plotting the faint end of the luminosity function at high redshift. The suppression of star formation in small halos is thought to decrease star formation in the universe as a whole, and thus the number of galaxies with low luminosity should drop as well. Looking for this drop in the faint end of the luminosity function could help pinpoint that reionization redshift (LB). Probing the faint end of this function in Lyman-$\alpha$ emission is one of the goals of the NEWFIRM survey which I will be participating in. The number of quasars is also expected to change, with the minimum circular velocity for quasars to form in a host galaxy increasing by a factor of roughly five after reionization (LB).

One of the most commonly used observational probes of reionization has already been alluded to, and it involves the detection of Lyman-$\alpha$ photons. Lyman-$\alpha$ is not only the most prominent of all hydrogen emission lines but its production comes from hydrogen ionization; Lyman-$\alpha$ photons are created via recombination and/or by free electrons causing collisional excitations (Cantalupo et al. 2008). In ionization fronts generated by quasars, the conditions are ideal for collisional excitations to produce Lyman-$\alpha$ photons – the temperature is about 10000 degrees Kelvin and neutral hydrogen fraction is about one-half – so according to the Cantalupo work, one should measure Lyman-$\alpha$ emission that is proportional to the square of the neutral hydrogen fraction before the ionizing radiation reached the emission region. Thus the amount of ionized vs neutral hydrogen can be measured in this fashion, and it can in principle be determined whether or not the universe has been ionized at that redshift (though this is harder in practice to measure). Additionally, halos of Lyman-$\alpha$ emission can sometimes be detected around high redshift sources (the halos being generated by a Doppler shifting of photons due to the Hubble flow of the universe) and these halos help to trace the pre-reionization neutral hydrogen distribution (LB).

In principle, the 21-cm emission line produced by the hydrogen spin-flip transition is also a good tracer of neutral hydrogen, as if this emission is observed to be strong in high redshift sources, then it can be surmised that reionization has not yet occurred. This emission line should disappear throughout the universe for a short time at the overlap phase of reionization, so if that were able to be observed, it would be ideal (LB). However, there are great difficulties in measuring this line. Primarily, measurements are taken by measuring the difference in brightness-temperature of the area in question from that of the cosmic microwave background. However, at such high redshifts, the 21-cm signal is approximately one-ten-thousandth that of Milky Way emission, so very sensitive instruments would be needed to detect this 21-cm emission with any kind of accuracy (Cantalupo et al 2008). Thus the Lyman-α line is often preferred to the 21-cm line, even though it does not pinpoint reionization as well.

The anisotropies of the cosmic microwave background, related to the optical depth of the last scattering surface, are another interesting, albeit practically difficult, reionization probe. According to the Loeb and Barkana review, reionization can modify the anisotropies seen in the CMB, damping anisotropies that were initially there and adding new ones. The earlier reionization occurs, the easier scattering occurs, and the larger the angular scale perturbations that are damped. So the reionization epoch redshift can be measured by determining the optical depth of last scattering. However, this is a difficult measurement since the CMB anisotropies depend on other parameters as well, and so it can only really provide a maximum redshift limit for instantaneous reionization, as has already been mentioned (LB).

A great deal of ongoing work is being conducted to try to make use of these methods and predictions and thus determine when and how the epoch of reionization actually occurred, using both theory and observation. While it would be impossible to list all work being done in such a paper, a short description of a few ongoing surveys will be provided here.

One study by Tittley and Meiksin (2007) is focusing on the temperature of the intergalactic medium following full reionization to try to figure out what sources provided the radiation for the ionization fronts, since low density regions should keep evidence of post-ionization temperatures. They make use of a simplified assumption involving one ionization front with radiative transfer of energy, and thus they are modeling the pre-overlap phase of reionization. This study uses an N-body computer simulation and determines the post-ionization temperatures due to ionization fronts produced by quasars, "miniquasars" (active intermediate black holes left behind by the earliest massive stars, as described by Madau), stars, and combinations thereof. While the authors stress that there are many uncertainties that need to be addressed (via more accurate hydrodynamic and radiative transfer calculations, for example), they found that the model that begins out with stars dominating the ionizing radiation front and then becomes dominated by quasars most closely reproduced temperature values determined by Doppler measurements. Though this is far from being conclusive, it is further evidence that points toward ionization fronts being powered by a mixture of stars and quasars (Tittley & Meiksin 2007).

Ciardi, Ferrara, and White (2003) have also made use of N-body simulations and radiative transfer codes to try to discover whether the discrepancy between models which predict reionization redshifts of 8 to 10 and observations like those of optical depth by WMAP which predict redshifts of 12 to 22 is a real one and whether it can be reconciled by more updated theory. They point out that due to the accuracy of the WMAP measurements, there must be physical processes missing from the models; those processes could include other ionizing radiation sources like population III stars (found in low mass regions that are not yet galaxies) or the Madau miniquasars, or, as they conclude is most likely, very efficient photon production and a high escape fraction (a constraint of 20% as mentioned earlier). They also note that the initial mass function that is used in the model can play a major role, and find that the model that results in optical depth values closest to those measured by WMAP is one with a top-heavy IMF and a 20% escape fraction, which reionizes the universe by a redshift of approximately 13. Additionally, it could be possible for the universe to have reionized twice; the efficiency of ionization fronts could have a large drop caused by a change in escape fraction, IMF, or an increase in the metallicity of stars and thus the universe could become neutral following an initial reionization at around a redshift of 9 and then become ionized again. The authors do believe this idea is highly unlikely and would require a great deal of work to provide evidence for, but it is interesting to note the possibility (Ciardi et al. 2003) .

One observational result of interest is a purported discovery by Pello et al. in 2004 of a redshift 10 galaxy with a strong Lyman-α emission line. If reionization of the IGM had not yet occurred by z = 10, then this line would be damped by the Lyman-α absorption line, unless this galaxy was a source of ionizing radiation (assuming that their calculations are correct to begin with and the galaxy is indeed at such a high redshift!). Loeb, Barkana, and Hernquist examined this galaxy in 2005 in attempt to determine whether reionization had occurred universally at this point, or whether it was simply evidence of an individual H II region. They examined the physical properties of hydrogen ionization (again with an uncertainty in the escape fraction) and of individual ionizing sources, which produce a region whose maximum size can be found by assuming negligible recombination and a 100% escape fraction. According to their models, if it is assumed that the Lyman-α optical depth at high redshift matches that determined by Sloan Digital Sky Survey quasars, then it is believable that this galaxy could be located in a region which has already been ionized. However, it is difficult to determine whether or not the entire IGM was reionized at this time without a wider survey of the z = 10 universe (Loeb, Barkana, & Hernquist 2005).

While theoretical studies and observational discoveries of the last seven years have helped to constrain the possible redshift range better than it was constrained in 2001 (the upper limit of 30 no longer seems plausible), a great deal of uncertainties still remain – escape fraction, ionization source type, luminosity function, etc – that plague those who have tried to study the reionization epoch, not to mention the technological limitations on observers searching for high redshift objects. Theory has helped to explain the why of reionization, and data provided by WMAP have helped to at least somewhat explain the when, but a fully constrained model for reionization remains elusive. I hope to make some contributions to this ever-expanding field of knowledge during my graduate research, but reionization studies are most likely far from over.

# REFERENCES

## Papers:

Benson, A.J., Sugiyama, N., Nusser, A., Lacey, C.G., 2006, MNRAS, 369, 1055.

Cantalupo, S., Porciani, C., Lilly, S.J., 2008, ApJ, 672, 48.

Ciardi, B., Ferrara, A., White, S.D.M., 2003, MNRAS, 344, L7.

Loeb, A., Barkana, R., 2001, ARAA, 39, 19.

Loeb, A., Barkana, R., Hernquist, L., 2005, ApJ, 620, 553.

Tittley, E.R., Meiksin, A., 2007, MNRAS, 380, 1369.

## Books:

Coles, P., Lucchin, F., 2002, "Cosmology: The Origin and Evolution of Cosmic Structure", 2nd Ed.