Starlight in the Universe [1]

Piero Madau

Space Telescope Science Institute, Baltimore, MD 21218

Abstract. There has been remarkable progress recently in both observational and theoretical studies of galaxy formation and evolution. Largely due to a combination of deep *Hubble Space Telescope* (*HST*) imaging, Keck spectroscopy, and *COBE* far-IR background measurements, new constraints have emerged on the emission history of the galaxy population as a whole. The global ultraviolet, optical, near- and far-IR photometric properties of the universe as a function of cosmic time cannot be reproduced by a simple stellar evolution model defined by a constant (comoving) star-formation density and a universal (Salpeter) initial mass function, and require instead a substantial increase in the stellar birthrate with lookback time. While the bulk of the stars present today appears to have formed relatively recently, the existence of a decline in the star-formation density above $z \approx 2$ remains uncertain. The study of the transition from the cosmic 'dark age' to an ionized universe populated with luminous sources can shed new light on the star formation activity at high redshifts, and promises answers to some fundamental questions on the formation of cosmic structures. If stellar sources are responsible for photoionizing the intergalactic medium at $z \approx 5$, the rate of star formation at this epoch must be comparable or greater than the one inferred from optical observations of galaxies at $z \approx 3$. A population of quasars at $z \lesssim 2$ could make a significant contribution to the extragalactic background light if dust-obscured accretion onto supermassive black holes is an efficient process.

INTRODUCTION

There is little doubt that the last few years have been very exciting times in galaxy formation and evolution studies. The remarkable progress in our understanding of faint galaxy data made possible by the combination of *Hubble Space Telescope* (*HST*) deep imaging and ground-based spectroscopy has permitted to shed new light on the evolution of the stellar birthrate in the universe, to identify the epoch $1 \lesssim z \lesssim 2$ where most of the optical extragalactic background light was produced, and to set important contraints on galaxy evolution scenarios. The explosion in the quantity of information available on the high-redshift universe at optical wavelengths has been complemented by the detection of the far-IR/sub-mm background by DIRBE and FIRAS onboard the *COBE* satellite, and by theoretical progress made in understanding how cosmic structure forms from initial density fluctuations [41]. The IR data have revealed the 'optically-hidden' side of galaxy formation, and shown that a significant fraction of the energy released by stellar nucleosynthesis is re-emitted as thermal radiation by dust. The underlying goal of all these efforts is to understand the growth of cosmic structures, the internal properties of galaxies and their evolution, the mechanisms that shaped Hubble's morphological sequence, and ultimately to map the transition from the cosmic 'dark age' to a ionized universe populated with luminous sources. While one of the important questions recently emerged is the nature (starbursts

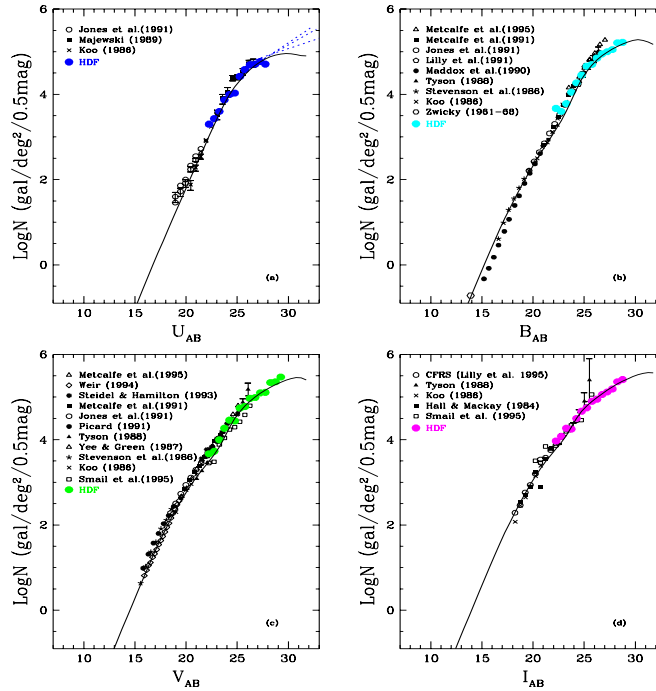arXiv:astro-ph/9902228v1 16 Feb 1999

FIGURE 1. Differential galaxy number counts per square degree as a function of apparent magnitude in four bandpasses from near-UV to near-IR. The sources of the data points are indicated in each panel (see [44] for the complete reference list). Note the decrease of the logarithmic slope $d \log N/dm$ at faint magnitudes.

or active galactic nuclei?) and redshift distribution of the ultraluminous sub-mm sources discovered by *SCUBA*, of perhaps equal interest is the possible existence of a large population of faint galaxies still undetected at high redshifts, as the color-selected ground-based and *Hubble Deep Field* (HDF) samples include only the brightest and bluest star-forming objects. In any hierarchical clustering ('bottom-up') scenario (the cold dark matter model being the best studied example), subgalactic structures are the first non-linearities to form. High-z dwarf galaxies and/or mini-quasars (i.e. an early generation of stars and accreting black holes in dark matter halos with circular velocities $v_c \sim 50 \text{ km s}^{-1}$) may then be one of the main source of UV photons and heavy elements at early epochs.

In this talk I will focus on some of the open issues and controversies surrounding our present understanding of the history of the conversion of cold gas into stars within galaxies, and of the evolution of luminous sources in the universe. An Einstein-deSitter (EdS) universe ($\Omega_M = 1$, $\Omega_\Lambda = 0$) with $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1} = 0.5$ will be adopted in the following.

COUNTING GALAXIES

Much observing time has been devoted in the past few years to the problem of the detection of galaxies at high redshifts, as it was anticipated that any knowledge of their early luminosity and color evolution would set important constraints on the history of structure and star formation in the universe. As the best view to date of the optical sky at faint flux levels, the

HDF imaging survey has rapidly become a key testing ground for models of galaxy evolution. The field, an undistinguished portion of the northen sky at high galactic latitudes (the data from a southern deep field are being analyzed as we speak), is essentially a deep core sample of the universe, acquired with the *HST* in a 10-day exposure. With its depth – reaching 5-σ limiting AB magnitudes of roughly 27.7, 28.6, 29.0, and 28.4 in U, B, V, and I [2] – and four-filter strategy to provide constraints on the redshift and age distribution of galaxies in the image, the HDF has offered the astronomical community the opportunity to study the galaxy population in unprecedented detail [57].

There are about 3000 galaxies in the HDF, corresponding to 2×10^6 deg^{-2} down to the faint limit of the images. The galaxy counts are shown in Figure 1 in four bandpasses centered at roughly 300, 450, 600, and 800 nm. A compilation of existing ground-based data is also shown, together with the predictions of no-evolution models, i.e. models in which the absolute brightness, volume density, and spectra of galaxies do not change with time. In all four bands, the logarithmic slope α of the galaxy number-apparent magnitude counts, $\log N(m) = \alpha m$, flattens at faint magnitudes, e.g., from $\alpha = 0.45$ in the interval $21 < B < 25$ to $\alpha = 0.17$ for $25 < B < 29$. The slope of the galaxy counts is a simple cosmological probe of the early history of star formation. The flattening at faint apparent magnitudes cannot be due to the reddening of distant sources as their Lyman break gets redshifted into the blue passband,[3] since the fraction of Lyman-break galaxies at $B \sim 25$ is only of order 10%. Moreover, an absorption-induced loss of sources could not explain the similar flattening of the galaxy counts observed in the V and I bands. Rather, the change of slope suggests that the surface density of luminous galaxies declines beyond $z \sim 1.5$.

THE BRIGHTNESS OF THE NIGHT SKY

The extragalactic background light (EBL) is an indicator of the total luminosity of the universe. It provides unique information on the evolution of cosmic structures at all epochs, as the cumulative emission from galactic systems and active galactic nuclei (AGNs) is expected to be recorded in this background.

The contribution of known galaxies to the optical EBL can be calculated directly by integrating the emitted flux times the differential galaxy number counts down to the detection threshold. The leveling off of the counts is clearly seen in Figure 2, where the function $i_\nu = 10^{-0.4(m+48.6)} \times N(m)$ is plotted against apparent magnitude in all bands [44]. While counts having a logarithmic slope of $\alpha \geq 0.40$ continue to add to the EBL at the faintest magnitudes, it appears that the HDF survey has achieved the sensitivity to capture the bulk of the extragalactic light from discrete sources (an extrapolation of the observed counts to brighter and/or fainter magnitudes would typically increase the sky brightness by less than 20%). To $AB = 29$ mag, the sky brightness from resolved galaxies in the I-band is $\approx 2 \times 10^{-20}$ ergs cm^{-2} s^{-1} Hz^{-1} sr^{-1}, increasing roughly as λ^2 from 2000 to 8000 Å. The

---

[2] To get a feeling of the depth of this survey, note that $AB = 29$ mag corresponds to the flux at Earth from a 100 Watt light bulb at a distance of 10 million kilometers.

[3] For galaxies with $z > 2$ ($z > 3.5$), the H I Lyman edge shifts into the 300 (450) nm HDF bandpass. Neutral hydrogen, which is ubiquitous both within galaxies and in intergalactic space, strongly absorbs ultraviolet light, creating a spectral discontinuity that can be used to identify young, high-redshift galaxies [50].
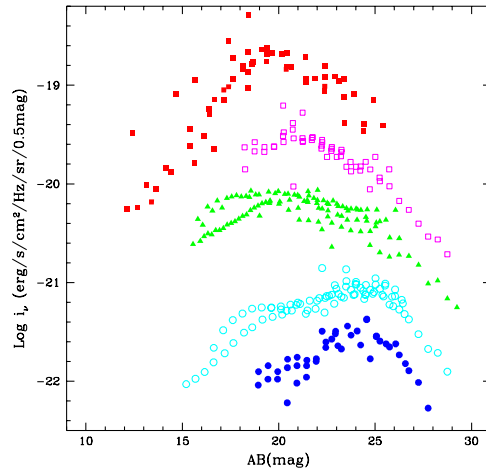
FIGURE 2. The contribution of known galaxies to the extragalactic background light per magnitude bin as a function of U (*filled circles*), B (*open circles*), V (*filled triangles*), I (*open squares*) and K (*filled squares*) magnitudes. For clarity, the B, V, I and K values have been multiplied by a factor of 2, 8, 20, and 40, respectively.

flattening of the number counts has the interesting consequence that the galaxies that produce $\sim 60\%$ of the blue EBL have $B < 24.5$. They are then bright enough to be identified in spectroscopic surveys, and are indeed known to have median redshift $\langle z \rangle = 0.6$ [32]. The quite general conclusion is that there is no evidence in the number-magnitude relation down to very faint flux levels for a large amount of star formation at high redshift. Note that these considerations do not constrain the *rate* of starbirth at early epochs, only the total (integrated over cosmic time) amount of stars – hence background light – being produced, and neglect the effect of dust reddening.

Figure 3 shows the total optical EBL from known galaxies together with the recent *COBE* results. The value derived by integrating the galaxy counts [44] down to very faint magnitude levels [because of the flattening at faint magnitudes of the number-magnitude relation most of the contribution to the optical EBL comes from relatively bright galaxies] implies a lower limit to the EBL intensity in the 0.3–2.2 μm interval of $I_{\rm opt} \approx 12 \, {\rm nW \, m^{-2} \, sr^{-1}}$.[4] When combined with the FIRAS and DIRBE measurements ($I_{\rm FIR} \approx 16 \, {\rm nW \, m^{-2} \, sr^{-1}}$ in the 125–5000 μm range), this gives an observed EBL intensity in excess of 28 nW m^{-2} sr^{-1}. The correction factor needed to account for the residual emission in the 2.2 to 125 μm region is probably $\lesssim 2$ [11]. (We shall see below how a population of dusty AGNs could make a significant contribution to the FIR background.) In the rest of this talk I will adopt a conservative reference value for the total EBL intensity associated with star formation activity over the entire history of the universe of $I_{\rm EBL} = 40 \, I_{40} \, {\rm nW \, m^{-2} \, sr^{-1}}$.

---

[4] The direct detection of the optical EBL at 3000, 5500, and 8000 Å derived from *HST* data [3] implies values that are about a factor of two higher than the integrated light from galaxy counts.
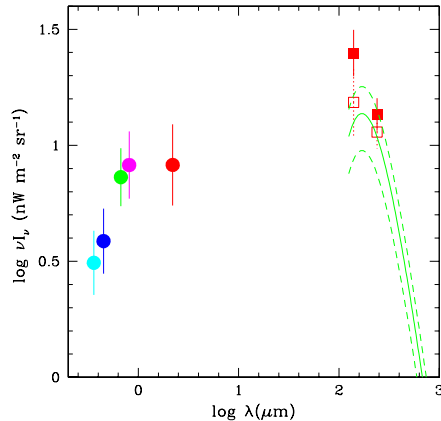
FIGURE 3. Spectrum of the extragalactic background light as derived from a compilation of ground-based and space-based galaxy counts in the U, B, V, I, and K-bands (*filled dots*), together with the FIRAS 125–5000 μm (*solid and dashed lines*) and DIRBE 140 and 240 μm (*filled squares*) detections [15], [25]. The *empty squares* show the DIRBE points after correction for WIM dust emission [29].

MODELING GALAXY EVOLUTION

In the past few years two different approaches have been widely used to interpret faint galaxy data [12]. In the simplest version of what I will call the 'traditional' scheme, a one-to-one mapping between galaxies at the present epoch and their distant counterparts is assumed: one starts from the local measurements of the distribution of galaxies as a function of luminosity and Hubble type and models their photometric evolution assuming some redshift of formation and a set of parameterized star formation histories [52]. These, together with an initial mass function (IMF) and a cosmological model, are then adjusted to match the observed number counts, colors, and redshift distributions. Beyond the intrinsic simplicity of assuming a well defined collapse epoch and pure-luminosity evolution thereafter, the main advantage of this kind of approach is that it can easily be made consistent with the classical view that ellipticals and spiral galaxy bulges (both redder than spiral disks and containing less gas) formed early in a single burst of duration 1 Gyr or less [2]. Spiral galaxies, by contrast, are characterized by a slower metabolism, i.e. star formation histories that extend to the present epoch. In these models, typically, much of the action happens at high-redshifts.

A more physically motivated way to interpret the observations is to construct semianalytic hierarchical models of galaxy formation and evolution [56]. Here, one starts ab initio from a power spectrum of primordial density fluctuations, and follows the formation and hierarchical merging of the dark matter halos that provide the early seeds for later galaxy formation. Baryonic gas gets accreted onto the halos and is shock-heated. Various prescriptions for gas cooling, star formation, feedback, and dynamical friction are adopted, and tuned to match the statistical properties of both nearby and distant galaxies. In this scenario, there is no period when bulges and ellipticals form rapidly as single units and are very bright: rather, small objects form first and merge continually to make larger ones. Galaxy do not
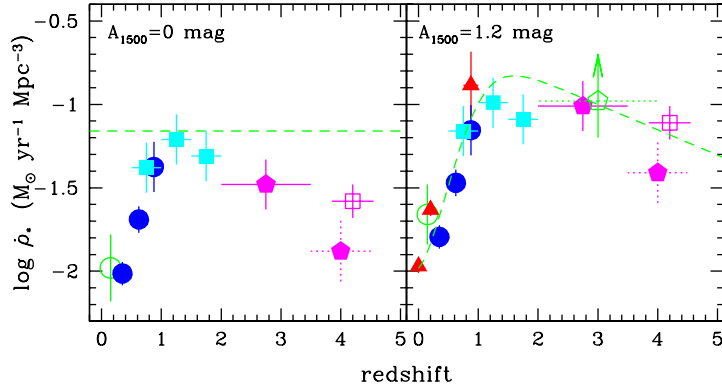
FIGURE 4. *Left*: Mean comoving density of star formation as a function of cosmic time. The data points with error bars have been inferred from the UV-continuum luminosity densities of [32] (*filled dots*), [8] (*filled squares*), [37] (*filled pentagons*), [54] (*empty dot*), and [51] (*empty square*). The *dotted line* shows the fiducial rate, $\langle \dot{\rho}_* \rangle = 0.054 \ \mathrm{M_\odot \, yr^{-1} \, Mpc^{-3}}$, required to generate the observed EBL. *Right*: dust corrected values ($A_{1500} = 1.2$ mag). The Hα determinations of [17], [53], and [19] (*filled triangles*), together with the *SCUBA* lower limit [27] (*empty pentagon*) have been added for comparison.

evolve as isolated objects, and the rate of interaction was higher in the past. The bulk of the galaxy population is predicted to have been assembled quite recently, and most galaxies never experience star formation rates in excess of a few solar masses per year.

STAR FORMATION HISTORY

Recently, it has become familiar to follow an alternative method, which focuses on the emission properties of the galaxy population as a whole. It traces the cosmic evolution with redshift of the galaxy luminosity density and offers the prospect of an empirical determination of the global star formation history of the universe and IMF of stars independently of the merging histories, complex evolutionary phases, and possibly short-lived star formation episodes of individual galaxies. The technique relies on two basic properties of stellar populations: a) the UV-continuum emission in all but the oldest galaxies is dominated by short-lived massive stars, and is therefore a direct measure, for a given IMF and dust content, of the instantaneous star formation rate; and b) the rest-frame near-IR light is dominated by near-solar mass evolved stars, the progenitors of which make up the bulk of a galaxy's stellar mass, and is more sensitive to the past star-formation history than the blue (and UV) light. By modeling the "emission history" of the universe at ultraviolet, optical, and near-infrared wavelengths from the present epoch to high redshifts, one should be able to shed light on some key questions in galaxy formation and evolution studies: Is there a characteristic epoch of star and metal formation in galaxies? What fraction of the luminous baryons observed today were already locked into galaxies at early epochs? Are high-z galaxies obscured by dust? Do spheroids form early and rapidly? Is there a universal IMF?

The comoving volume-averaged history of star formation follows a relatively simple dependence on redshift. Its latest version, uncorrected for dust extinction, is plotted in Figure

4 (*left*). The measurements are based upon the rest-frame UV luminosity function (at 1500 and 2800 Å), assumed to be from young stellar populations [34]. The prescription for a 'correct' de-reddening of these values has been the subject of an ongoing debate. Dust may play a role in obscuring the UV continuum of Canada-France Reshift Survey (CFRS, $0.3 < z < 1$) and Lyman-break ($z \approx 3$) galaxies, as their colors are too red to be fitted with an evolving stellar population and a Salpeter IMF [37]. Figure 4 (*right*) depicts an extinction-corrected version of the same plot. The best-fit cosmic star formation history (shown by the dashed-line) with such a universal correction produces a total EBL of 37 nW m^{-2} sr^{-1}. About 65% of this is radiated in the UV+optical+near-IR between 0.1 and 5 μm; the total amount of starlight that is absorbed by dust and reprocessed in the far-IR is 13 nW m^{-2} sr^{-1}. Because of the uncertainties associated with the incompleteness of the data sets, photometric redshift technique, dust reddening, and UV-to-SFR conversion, these numbers are only meant to be indicative. On the other hand, this very simple model is not in obvious disagreement with any of the observations, and is able, in particular, to provide a reasonable estimate of the galaxy optical and near-IR luminosity density.

THE STELLAR BARYON BUDGET

With the help of some simple stellar population synthesis tools it is possible at this stage to make an estimate of the stellar mass density that produced the integrated light observed today. The total *bolometric* luminosity of a simple stellar population (a single generation of coeval, chemically homogeneous stars) having mass M can be well approximated by a power-law with time for all ages $t \gtrsim 100$ Myr,

$$L(t) = 1.3 \, L_\odot \frac{M}{M_\odot} \left(\frac{t}{1 \, \text{Gyr}} \right)^{-0.8} \tag{1}$$

(cf. [6]), where we have assumed solar metallicity and a Salpeter IMF truncated at 0.1 and 125 $M_\odot$. In a stellar system with arbitrary star formation rate per unit cosmological volume, $\dot\rho_*$, the comoving bolometric emissivity at time t is given by the convolution integral

$$\rho_{\text{bol}}(t) = \int_0^t L(\tau)\dot\rho_*(t - \tau)d\tau. \tag{2}$$

The total background light observed at Earth ($t = t_H$) is

$$I_{\text{EBL}} = \frac{c}{4\pi} \int_0^{t_H} \frac{\rho_{\text{bol}}(t)}{1 + z} dt, \tag{3}$$

where the factor $(1 + z)$ at the denominator is lost to cosmic expansion when converting from observed to radiated (comoving) luminosity density. From the above equations it is easy to derive in a EdS cosmology

$$I_{\text{EBL}} = 740 \text{ nW m}^{-2} \text{sr}^{-1} \langle \frac{\dot\rho_*}{\text{M}_\odot \, \text{yr}^{-1} \, \text{Mpc}^{-3}} \rangle \left(\frac{t_H}{13 \, \text{Gyr}} \right)^{1.87}. \tag{4}$$

The observations shown in Figure 3 therefore imply a "fiducial" mean star formation density of $\langle \dot\rho_* \rangle = 0.054 \, I_{40}$ M$_\odot$ yr^{-1} Mpc^{-3}. The total stellar mass density observed today is
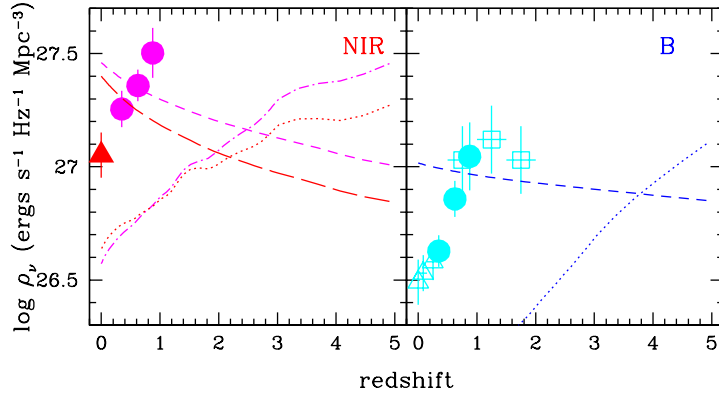
FIGURE 5. *Left*: Synthetic evolution of the near-IR luminosity density at rest-frame wavelengths of 1.0 (*long-dashed line*) and 2.2 μm (*short-dashed line*). The model assumes a constant star formation rate of $\dot{\rho}_* = 0.054$ M$_\odot$ yr^{-1} Mpc^{-3} (Salpeter IMF). The *dotted* (2.2 μm) and *dash-dotted* (1.0 μm) curves show the emissivity of a simple stellar population with formation redshift $z_{on} = 5$, and total mass equal to the mass observed in spheroids today [16]. The data points are taken from [32] (*filled dots*) and [18] (*filled triangle*). *Right*: Same but in the B-band. The data points are taken from [32] (*filled dots*), [13] (*empty triangles*), and [8] (*empty squares*).

$$\rho_*(t_H) = (1 - R) \int_0^{t_H} \dot{\rho}_*(t)dt \approx 5 \times 10^8 \, I_{40} \ \text{M}_\odot \, \text{Mpc}^{-3} \tag{5}$$

(corresponding to $\Omega_* = 0.007 \, I_{40}$), where R is the mass fraction of a generation of stars that is returned to the interstellar medium, $R \approx 0.3$ for a Salpeter IMF. The optical/FIR background therefore requires that about 10% of the nucleosynthetic baryons today [5] are in the forms of stars and their remnants. The predicted stellar mass-to-blue light ratio is $\langle M/L_B \rangle \approx 5$. These values are quite sensitive to the lower-mass cutoff of the IMF, as very-low mass stars can contribute significantly to the mass but not to the integrated light of the whole stellar population. A lower cutoff of $0.5 \, \text{M}_\odot$ instead of the $0.1 \, \text{M}_\odot$ adopted would decrease the mass-to-light ratio (and Ω_*) by a factor of 1.9 for a Salpeter function.

TWO EXTREME SCENARIOS

Based on the agreement between the $z \approx 3$ and $z \approx 4$ luminosity functions at the bright end, it has been recently argued [51] that the decline in the luminosity density of faint HDF Lyman-break galaxies observed in the same redshift interval [34] may not be real, but simply due to sample variance in the HDF. When extinction corrections are applied, the emissivity per unit comoving volume due to star formation may then remain essentially flat for all redshift $z \gtrsim 1$ (see Fig. 4). While this has obvious implications for hierarchical models of structure formation, the epoch of first light, and the reionization of the intergalactic medium (IGM), it is also interesting to speculate on the possibility of a constant star-formation density at *all* epochs $0 \leq z \leq 5$. Figure 5 shows the time evolution of the blue and near-IR rest-frame luminosity density of a stellar population characterized by a Salpeter IMF, solar metallicity, and a (constant) star formation rate of $\dot{\rho}_* = 0.054$ M$_\odot$ yr^{-1} Mpc^{-3} (needed

to produce the observed EBL). The predicted evolution appears to be a poor match to the observations: it overpredicts the local B and K-band luminosity densities, and underpredicts the $1\,\mu$m emissivity at $z \approx 1$.

At the other extreme, we know from stellar population studies that about half of the present-day stars are contained in spheroidal systems, i.e. elliptical galaxies and spiral galaxy bulges, and that these stars formed early and rapidly. The expected rest-frame blue and near-IR emissivity of a simple stellar population with formation redshift $z_{\rm on} = 5$ and total mass density equal to the mass in spheroids observed today is shown in Figure 5. In this model the near-IR and blue emissivities at $z = 4 - 5$ are comparable with the values observed at $z = 1$. *HST*-NICMOS deep observations may be able to test similar scenarios for the formation of elliptical galaxies at early times.

THE MASS DENSITY IN BLACK HOLES

Recent dynamical evidence indicates that supermassive black holes reside at the center of most nearby galaxies. The available data (about 30 objects) show a strong correlation (but with a large scatter) between bulge and black hole mass [38], with $M_{\rm bh} = 0.006\,M_{\rm bulge}$ as a best-fit. The total mass density in spheroids today is $\Omega_{\rm bulge} = 0.0036^{+0.0024}_{-0.0017}$ [16], implying a mean mass density of dead quasars

$$\rho_{\rm bh} = 1.5^{+1.0}_{-0.7} \times 10^6 \ {\rm M_\odot \, Mpc^{-3}}. \tag{6}$$

Since the observed energy density from all quasars is equal to the emitted energy divided by the average quasar redshift [49], the total contribution to the EBL from accretion onto black holes is

$$I_{\rm bh} = \frac{c^3}{4\pi} \frac{\eta \rho_{\rm bh}}{\langle 1 + z \rangle} \approx 18 \ {\rm nW \, m^{-2} \, sr^{-1}} \eta_{0.1} \langle 1 + z \rangle^{-1}, \tag{7}$$

where $\eta_{0.1}$ is the efficiency for transforming accreted rest-mass energy into radiation (in units of 10%). Quasars at $z \lesssim 2$ could then make a significant contribution to the brightness of the night sky if dust-obscured accretion onto supermassive black holes is an efficient process [21], [14]. [5]

THE END OF THE 'DARK AGES'

The epoch of reionization marked the end of the 'dark ages' during which the ever-fading primordial background radiation cooled below 3000 K and shifted first into the infrared and then into the radio. Darkness persisted until early structures collapsed and cooled, forming the first stars and quasars that lit the universe up again [45]. The application of the Gunn-Peterson constraint on the amount of smoothly distributed neutral material along the line of sight to distant objects requires the hydrogen component of the diffuse IGM to have been highly ionized by $z \approx 5$ [47], and the helium component by $z \approx 2.5$ [10]. From

---

[5] It might be interesting to note in this context that a population of AGNs with strong intrinsic absorption (Type II quasars) is actually invoked in many current models for the X-ray background [33], [7].
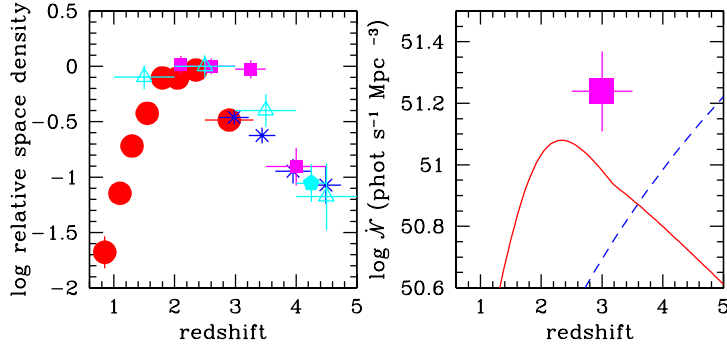
FIGURE 6. *Left*: comoving space density of bright QSOs as a function of redshift. The data points with error bars are taken from [24] *(filled dots)*, [55] *(filled squares)*, [46] *(crosses)*, and [28] *(filled pentagon)*. The *empty triangles* show the space density of radio-loud quasars [26]. *Right*: comoving emission rate of hydrogen Lyman-continuum photons (*solid line*) from QSOs, compared with the minimum rate (*dashed line*) which is needed to fully ionize a fast recombining (with gas clumping factor $C = 30$) EdS universe with $\Omega_b h^2 = 0.02$. Models based on photoionization by quasar sources appear to fall short at $z \approx 5$. The data point shows the estimated contribution from star-forming galaxies at $z \approx 3$, assuming that the fraction of Lyman continuum photons which escapes the galaxy H I layers into the intergalactic medium is 50% (see [35] for details).

QSO absorption studies we also know that neutral hydrogen at high-z accounts for only a small fraction, $\sim 10\%$, of the nucleosynthetic baryons [30].

A substantial population of dwarf galaxies having star formation rates $< 0.3\,\mathrm{M_\odot\,yr^{-1}}$, and a space density in excess of that predicted by extrapolating to faint magnitudes the best-fit Schechter function, may be expected to form at early times in hierarchical clustering models, and has been recently proposed [39], [35] as a possible candidate for photoionizing the IGM at these early epochs. Establishing the character of cosmological ionizing sources is an efficient way to constrain competing models for structure formation in the universe, and to study the collapse and cooling of small mass objects at early epochs.

The study of the candidate sources of ionization at $z = 5$ can be simplified by noting that the *breakthrough epoch* (when all radiation sources can see each other in the hydrogen Lyman-continuum) occurs much later in the universe than the *overlap epoch* (when individual ionized zones become simply connected and every point in space is exposed to ionizing radiation). This implies that at high redshifts the ionization equilibrium is actually determined by the *instantaneous* UV production rate [35]. The fact that the IGM is rather clumpy and still optically thick at overlapping, coupled to recent observations of a rapid decline in the space density of radio-loud quasars and of a large population of star-forming galaxies at $z \gtrsim 3$, has some interesting implications for rival ionization scenarios and for the star formation activity in the interval $< 3 < z < 5$.

The existence of a decline in the space density of bright quasars at redshifts beyond ~ 3 was first suggested by [40], and has been since then the subject of a long-standing debate. In recent years, several optical surveys have consistently provided new evidence for a turnover in the QSO counts [24], [55], [46], [28]. The interpretation of the drop-off observed in optically selected samples is equivocal, however, because of the possible bias introduced

by dust obscuration arising from intervening systems. Radio emission, on the other hand, is unaffected by dust, and it has recently been shown [48] that the space density of radio-loud quasars also decreases strongly for $z > 3$. This argues that the turnover is indeed real and that dust along the line of sight has a minimal effect on optically-selected QSOs. In this case the QSO emission rate of hydrogen ionizing photons per unit comoving volume drops by a factor of 3 from $z = 2.5$ to $z = 5$, as shown in Figure 6.

Galaxies with ongoing star-formation are another obvious source of Lyman-continuum photons. Since the rest-frame UV continuum at 1500 Å (redshifted into the visible band for a source at $z \approx 3$) is dominated by the same short-lived, massive stars which are responsible for the emission of photons shortward of the Lyman edge, the needed conversion factor, about one ionizing photon every 10 photons at 1500 Å, is fairly insensitive to the assumed IMF and independent of the galaxy history for $t \gg 10^7$ yr. Figure 6 (*right*) shows the estimated Lyman-continuum luminosity density of galaxies at $z \approx 3$. The data point assumes a value of $f_{\mathrm{esc}} = 0.5$ for the unknown fraction of ionizing photons which escapes the galaxy H I layers into the intergalactic medium. One should note that, while highly reddened galaxies at high redshifts would be missed by the Lyman-break color technique (which isolates sources that have blue colors in the optical and a sharp drop in the rest-frame UV), it seems unlikely that very dusty objects (with $f_{\mathrm{esc}} \ll 1$) would contribute in any significant manner to the ionizing metagalactic flux.

REIONIZATION

When an isolated point source of ionizing radiation turns on in a neutral medium, the ionized volume initially grows in size at a rate fixed by the emission of UV photons, and an ionization front separating the H II and H I regions propagates into the neutral gas. Most photons travel freely in the ionized bubble, and are absorbed in a transition layer. The evolution of an expanding H II region is governed by the equation

$$\frac{dV_I}{dt} - 3HV_I = \frac{\dot{N}_{\mathrm{ion}}}{\bar{n}_{\mathrm{H}}} - \frac{V_I}{\bar{t}_{\mathrm{rec}}}, \tag{8}$$

where V_I is the proper volume of the ionized zone, $\dot{N}_{\mathrm{ion}}$ is the number of ionizing photons emitted by the central source per unit time, $\bar{n}_{\mathrm{H}}$ is the mean hydrogen density of the expanding IGM, H is the Hubble constant, and $\bar{t}_{\mathrm{rec}}$ is the hydrogen mean recombination timescale,

$$\bar{t}_{\mathrm{rec}} = [(1 + 2\chi)\bar{n}_{\mathrm{H}}\alpha_B \, C]^{-1} = 0.3 \, \mathrm{Gyr} \left(\frac{\Omega_b h^2}{0.02}\right)^{-1} \left(\frac{1+z}{4}\right)^{-3} C_{30}^{-1}. \tag{9}$$

One should point out that the use of a volume-averaged clumping factor, $C \equiv \langle n_{\mathrm{HII}}^2 \rangle / \bar{n}_{\mathrm{HII}}^2$, in the recombination timescale is only justified when the size of the H II region is large compared to the scale of the clumping, so that the effect of many clumps (filaments) within the ionized volume can be averaged over. The validity of this approximation can be tested by numerical simulations (see Figure 7). Across the I-front the degree of ionization changes sharply on a distance of the order of the mean free path of an ionizing photon. When $\bar{t}_{\mathrm{rec}}$
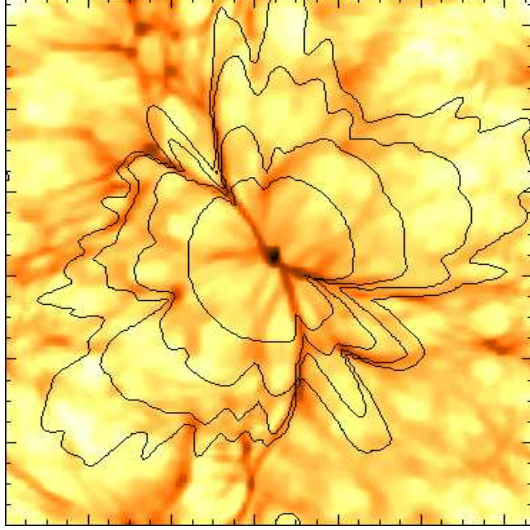
FIGURE 7. Simulating the reionization of the universe: propagation of an ionization front in a 128^3 cosmological density field. A 'mini-quasar' with $\dot{N} = 5 \times 10^{53}$ s^{-1} was turned on at the densest cell, in a virialized halo of total mass $1.3 \times 10^{11} \, M_\odot$. The box length is 2.4 comoving Mpc. The solid contours give the position of the front at 0.15, 0.25, 0.38, and 0.57 Myr after the quasar has switched on at $z = 7$. The underlying greyscale image indicates the initial H I density field. (From [1].)

is much smaller than the Hubble time, the growth of the H II region is slowed down by recombinations in the highly inhomogeneous medium, and its evolution can be decoupled from the expansion of the universe.

In analogy with the individual H II region case, it can be shown that the hydrogen component in a highly inhomogeneous universe is completely reionized when the number of photons emitted above 1 ryd in one recombination time equals the mean number of hydrogen atoms. At any given epoch there is a critical value for the emission rate of ionizing photons per unit cosmological comoving volume,

$$\dot{\mathcal{N}}_{\rm ion}(z) = \frac{\bar{n}_{\rm H}(0)}{\bar{t}_{\rm rec}(z)} = (10^{51.2} \text{ s}^{-1} \text{ Mpc}^{-3}) \, C_{30} \left(\frac{1+z}{6}\right)^3 \left(\frac{\Omega_b h^2}{0.02}\right)^2, \tag{10}$$

which is independent of the (unknown) previous emission history of the universe: only rates above this value will provide enough UV photons to ionize the IGM by that epoch. One can then compare our estimate of $\dot{\mathcal{N}}_{\rm ion}$ to the inferred contribution from QSOs and star-forming galaxies. The uncertainty on this critical rate is difficult to estimate, as it depends on the clumpiness of the IGM (scaled in the expression above to the value inferred at $z = 5$ from numerical simulations [20]) and the nucleosynthesis constrained baryon density. The evolution of the critical rate as a function of redshift is plotted in Figure 6 (*right*). While $\dot{\mathcal{N}}_{\rm ion}$ is comparable to the quasar contribution at $z \gtrsim 3$, there is some indication of a deficit of Lyman-continuum photons at $z = 5$. For bright, massive galaxies to produce enough UV radiation at $z = 5$, their space density would have to be comparable to the one observed at $z \approx 3$, with most ionizing photons being able to escape freely from the regions of star formation into the IGM. This scenario may be in conflict with direct observations of local

starbursts below the Lyman limit showing that at most a few percent of the stellar ionizing radiation produced by these luminous sources actually escapes into the IGM [31].[6]

It is interesting to convert the derived value of $\dot{\mathcal{N}}_{\text{ion}}$ into a "minimum" star formation rate per unit (comoving) volume, $\dot{\rho}_*$:

$$\dot{\rho}_*(z) = \dot{\mathcal{N}}_{\text{ion}}(z) \times 10^{-53.1} f_{\text{esc}}^{-1} \approx 0.013 f_{\text{esc}}^{-1} \left(\frac{1+z}{6}\right)^3 \ \text{M}_\odot \, \text{yr}^{-1} \, \text{Mpc}^{-3}. \tag{11}$$

The star formation density given in the equation above is comparable with the value directly "observed" (i.e., uncorrected for dust reddening) at $z \approx 3$ [37]. The conversion factor assumes a Salpeter IMF with solar metallicity, and has been computed using a population synthesis code [4]. It can be understood by noting that, for each 1 $M_\odot$ of stars formed, 8% goes into massive stars with $M > 20 M_\odot$ that dominate the Lyman-continuum luminosity of a stellar population. At the end of the C-burning phase, roughly half of the initial mass is converted into helium and carbon, with a mass fraction released as radiation of 0.007. About 25% of the energy radiated away goes into ionizing photons of mean energy 20 eV. For each 1 $M_\odot$ of stars formed every year, we then expect

$$\frac{0.08 \times 0.5 \times 0.007 \times 0.25 \times M_\odot c^2}{20 \, \text{eV}} \frac{1}{1 \, \text{yr}} \sim 10^{53} \, \text{phot} \, \text{s}^{-1} \tag{12}$$

to be emitted shortward of 1 ryd.

CONCLUSIONS

Recent studies of the volume-averaged history of stellar birth are pointing to an era of intense star formation at $z \approx 1 - 1.5$. The optical datasets imply that a fraction close to 65% of the present-day stars was produced at $z > 1$, and only 25% at $z > 2$. About half of the stars observed today would be more than 9 Gyr old, and only 10% would be younger than 5 Gyr.[7] There is no 'single epoch of galaxy formation': rather, it appears that galaxy formation is a gradual process. Numerous uncertainties remain, however, particularly the role played by dust in obscuring star-forming objects. Our first glimpse of the history of galaxies to $z \sim 4$ leads to the exciting question of what happened before. Substantial sources of ultraviolet photons must have been present at $z \gtrsim 5$ to keep the universe ionized, perhaps low-luminosity quasars [23] or a first generation of stars in dark matter halos with virial temperature $T_{\text{vir}} \sim 10^4 - 10^5$ K [42], [22]. Early star formation provides a possible explanation for the widespread existence of heavy elements in the Lyα forest [9], while reionization by QSOs may produce a detectable signal in the radio extragalactic background at meter wavelengths [36]. A detailed exploration of such territories must await projected facilities like the *Next Generation Space Telescope* and the *Square Kilometer Radio Telescope*.

---

[6] At $z = 3$ Lyman-break galaxies radiate into the IGM more ionizing photons than QSOs if $f_{\text{esc}} \gtrsim 30\%$.

[7] Unlike the measured number densities of objects and rates of star formation, the integrated stellar mass density does not depend on the assumed cosmological model.

REFERENCES

1. Abel, T., Norman, M. L., & Madau, P. 1998, preprint (astro-ph/9812151)
2. Bernardi, M., et al. 1998, ApJ, 508, L143
3. Bernstein, R. A. 1998, preprint
4. Bruzual, A. C., & Charlot, S. 1999, in preparation
5. Burles, S., & Tytler, D. 1998, preprint (astro-ph/9803071)
6. Buzzoni, A. 1995, ApJS, 98, 69
7. Comastri, A., et al. 1995, A&A, 296, 1
8. Connolly, A. J., et al. 1997, ApJ, 486, L11
9. Cowie, L. L., et al. 1995, AJ, 109, 1522
10. Davidsen, A. F., Kriss, G. A., & Zheng, W. 1996, Nature, 380, 47
11. Dwek, E., et al. 1998, ApJ, 508, 106
12. Ellis, R. S. 1997, ARA&A, 35, 389
13. Ellis, R. S., et al. 1996, MNRAS, 280, 235
14. Fabian, A. C., & Iwasawa, K. 1999, MNRAS, in press
15. Fixsen, D. J., et al. 1998, ApJ, 508, 123
16. Fukugita, M., Hogan, C. J., & Peebles, P. J. E. 1998, ApJ, 503, 518
17. Gallego, J., et al. 1995, ApJ, 455, L1
18. Gardner, J. P., et al. B. E. 1997, ApJ, 480, L99
19. Glazebrook, K., et al. 1998, preprint (astro-ph/9808276)
20. Gnedin, N. Y., & Ostriker, J. P. 1997, ApJ, 486, 581
21. Haehnelt, M. G., Natarajan, P., & Rees, M. J. 1998, MNRAS, 300, 817
22. Haiman, Z., & Loeb, A. 1997, ApJ, 483, 21
23. Haiman, Z., & Loeb, A. 1998, ApJ, 503, 505
24. Hartwick, F. D. A., & Schade, D. 1990, ARA&A, 28, 437
25. Hauser, M. G., et al. 1998, ApJ, 508, 25
26. Hook, I. M., Shaver, P. A., & McMahon, R. G. 1998, in The Young Universe: Galaxy Forma-
    tion and Evolution at Intermediate and High Redshift, ed. S. D'Odorico, A. Fontana, & E.
    Giallongo (San Francisco: ASP), p. 17
27. Hughes, D., et al. 1998, Nature, 398, 241
28. Kennefick, J. D., Djorgovski, S. G., & de Carvalho, R. R. 1995, AJ, 110, 2553
29. Lagache, G., et al. 1999, preprint (astro-ph/9901059)
30. Lanzetta, K. M., Wolfe, A. M., & Turnshek, D. A. 1995, ApJ, 440, 435
31. Leitherer, C., et al. 1995, ApJ, 454, L19
32. Lilly, S. J., et al. 1996, ApJ, 460, L1
33. Madau, P., Ghisellini, G., & Fabian, A. C. 1994, MNRAS, 270, L17
34. Madau, P., et al. 1996, MNRAS, 283, 1388
35. Madau, P., Haardt, F., & Rees, M. J. 1998, ApJ, in press
36. Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
37. Madau, P., Pozzetti, L., & Dickinson, M. E. 1998, ApJ, 498, 106
38. Magorrian, G., et al. 1998, AJ, 115, 2285
39. Miralda-Escudé, J., & Rees, M. J. 1998, ApJ, 497, 21
40. Osmer, P. S. 1982, ApJ, 253, 280
41. Ostriker, J. P. 1999, this volume
42. Ostriker, J. P., & Gnedin, N. Y. 1996, ApJ, 472, L63
43. Pettini, M., et al. 1998, ApJ, 508, 539

44. Pozzetti, L., et al. 1998, MNRAS, 298, 1133

45. Rees, M. 1998, in The Next Generation Space Telescope: Science Drivers and Technological Challenges (Noordwijk: ESA Pub.), p. 5

46. Schmidt, M., Schneider, D. P., & Gunn, J. E. 1995, AJ, 110, 68

47. Schneider, D. P., Schmidt, M., & Gunn, J. E. 1991, AJ, 101, 2004

48. Shaver, P. A., et al. 1996, Nature, 384, 439

49. Soltan, A. 1982, MNRAS, 200, 115

50. Steidel, C. C., et al. 1996, ApJ, 462, L17

51. Steidel, C. C., et al. 1998, preprint (astro-ph/9811399)

52. Tinsley, B. M. 1980, Fundam. Cosmic Phys., 5, 287

53. Tresse, L., & Maddox, S. J. 1998, ApJ, 495, 691

54. Treyer, M. A., et al. 1998, MNRAS, 300, 303

55. Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, ApJ, 421, 412

56. White, S. D. M.,& Frenk, C. S. 1991, ApJ, 379, 25

57. Williams, R. E., et al. 1996, AJ, 112, 1335