Protogalaxies

The term protogalaxies or sometimes primeval galaxies (hereafter PGs) has been used in the literature with a range of different meanings, but a general definition may be something like this: progenitors of the present-day (normal) galaxies, in the early stages of formation; also, many, but not all, authors would also add to that the words 'at high redshifts'. The key phrase is 'the early stages of formation', which is fundamentally not well defined.

The ambiguities are mainly due to the fact that our understanding of what GALAXY FORMATION means has been evolving over the past few decades. Moreover, even at any given time, experts of different technical orientations (e.g. observers versus n-body simulators) who would otherwise agree about the general picture of galaxy formation would have different and equally legitimate definitions in mind. For example, an observer may mean 'the first major burst of star formation in a progenitor of a present-day elliptical galaxy', whereas a theorist may mean 'the peak merging epoch of dark halos of the fragments which assemble to produce an average galaxy today'; others would define a PG as a still gaseous body before any STAR FORMATION has taken place or as an overdense region of DARK MATTER in the very early universe, destined to become gravitationally bound and to collapse. In some sense all of them are right.

It may be useful here to offer a brief account of our present understanding of galaxy formation and then describe some of the relevant observations, thus defining PGs in an implicit way. While the subject is still evolving rapidly as of this writing (late 1999), most cosmologists would agree that the basic picture we now have is likely to be essentially correct. Instead of seeking a hypothetical magic epoch when PGs appear and then evolve into the familiar types of galaxies today, the focus is now on understanding of the complete history of galaxy and large-scale structure assembly, as well as the history of star formation in all galaxies and the chemical enrichment of gas in the universe at large. It is then almost a matter of taste which slice of formative history would be called the epoch of galaxy formation and which objects or which evolutionary stages of normal galaxies would be called PGs.

In what follows, we outline a broad-brush picture and give general arguments and order-of-magnitude estimates, which are likely to be at least roughly correct, however much our understanding of galaxy formation evolves in the forthcoming years. Ever more detailed models will be producing more precise predictions, which can then be tested by future observations.

From density fluctuations to protogalaxies

The basic paradigm of structure formation is that it takes place through the gravitational instability of bound, overdense regions in the early universe. The distribution of such regions in mass (or, equivalently, density) is quantified by the initial density perturbation spectrum, which is produced by some type of quantum fluctuation processes in the early universe. Sufficiently overdense regions separate from the universal expansion and collapse under their own self-gravity, roughly on the free-fall time scale:

$$t_{\rm ff} = \left(\frac{\pi^2}{8G}\right)^{1/2} R_{\rm init}^{3/2}\, M^{-1/2}$$

$$\approx 5\times 10^8\ {\rm yr} \times \left(\frac{R_{\rm init}}{100\ {\rm kpc}}\right)^{3/2} \left(\frac{M}{10^{12} M_\odot}\right)^{-1/2}$$

$$\approx\ 1.6\times 10^9 {\rm yr} \times \left(\frac{R_{\rm init}}{10\ {\rm Mpc}}\right)^{3/2} \left(\frac{M}{10^{15} M_\odot}\right)^{-1/2}$$

where $R_{\rm init}$ is the initial radius of the bound region and M is the enclosed mass. It can be seen that for a galactic or subgalactic mass fragment the corresponding cosmic epochs ($\sim 10^8$ yr) imply high REDSHIFTS of collapse, whereas CLUSTERS OF GALAXIES or larger structures may be still collapsing today. The free-fall time scale is a lower limit to the formation time scale, its duration and the cosmic epoch: in practice, it may take several free-fall times for a given mass overdensity to be assembled through a process of hierarchical merging and virialization. To a first order, this simple argument implies that the peak epoch of galaxy formation is likely to be found at the cosmic epochs of a few $\times 10^8$ to a few $\times 10^9$ yr, or roughly in the redshift range $z \sim$ 2–20 or so, whereas the epoch of cluster formation lasts many billions of years and is still going on now.

Collapse

Since the non-baryonic dark matter appears to dominate the total mass, the overdense regions, i.e. proto-dark-halos, can start collapsing even while the universe is still ionized, e.g. at $z \sim 10^4$, with the baryonic matter following. The smallest mass fluctuations collapse the fastest; however, they are also most readily erased by a variety of damping processes operating in the early universe, such as the streaming of matter and photons, sound waves, etc. It is now believed that the smallest structures which survive at the recombination epoch are similar in mass to GLOBULAR CLUSTERS or DWARF GALAXIES, i.e. $(10^5$–$10^8)M_\odot$. They may be the basic building blocks of galaxies.

Images of the COSMIC MICROWAVE BACKGROUND (CMB) photosphere show a snapshot of overdense regions at the recombination epoch at $z \sim 1100$. At the present time, the resolution of such observations corresponds to physical scales of large clusters and superclusters of galaxies rather than galaxies themselves. Nevertheless, these observations support the basic picture of structure formation via gravitational instability.

Any energy dissipation in the baryonic component of the mass (also called the cooling of PGs) would of course accelerate the collapse and lead to the formation of denser objects, since systems which dissipate energy at a fixed mass become more tightly bound. The distinction between galaxies and large-scale structures such as galaxy groups or clusters (or larger) is that the galaxies (or

their protogalactic fragments) can cool faster than the free-fall time. Physical mechanisms which enable this energy dissipation in collapsing PGs include the inverse-Compton cooling of hot gas on the CMB radiation (CMBR) photons and shocks in the infalling and colliding gas clouds.

Thus galaxies (or their building blocks) become distinct concentrations within the overall large-scale structure, whose evolution is dominated by the gravity of the dark matter. Simple arguments based on the mean density of galaxies, the spin-up of galaxy disks, etc all suggest that they collapsed by about a factor of 10 or 20 in radius, whereas pure dissipationless collapse can produce only a factor of 2 (this is a direct consequence of the virial theorem for systems bound by gravity). Given the ratio of the comoving rms separation of galaxies today and their typical radii, this must have happened at $z < 20$ or so. It is also hard to have an effective cooling of PGs at redshifts much higher than that.

There are two fundamental aspects of galaxy formation: assembly of the mass (mostly dark), and conversion of the primordial gas (see GAS IN GALAXIES) into stars, and the subsequent chemical enrichment through SUPERNOVA explosions. The former is easier to model numerically; the latter is what is actually observable. The two may be connected, as bursts of star formation are likely to result from mergers of gas-rich PG fragments.

Assembly of protogalaxies

At every spatial (or mass) scale, there is a gradual, hierarchical merging of fragments into ever larger structures. This process operates first at the smaller scales, and moves to ever larger ones as the age of the universe increases and as the smaller fragments are consumed in mergers. While mergers of galaxies are relatively rare (yet readily observed) in the present-day universe, such processes must have been much more common at higher redshifts, when galaxies and their initial building blocks were closer together, with less developed peculiar velocities, and when there were more fragments to merge with (see GALAXIES: INTERACTIONS AND MERGERS). However, this process of hierarchical assembly never really ceases, and today it is the principal mechanism of galaxy transformation.

In 1960s and early 1970s, a simple picture of galaxy formation as a monolithic collapse of a single large PG cloud was promoted by Eggen, Lynden-Bell and Sandage, with a majority of the POPULATION II (bulge and halo) STARS formed more or less during a single free-fall time. This picture has now been completely replaced by the hierarchical formation scenario. The remaining issue is when and where did most of the stars form? A merger of a large number of PG fragments on a time scale comparable with the free-fall time of the entire system, and accompanied by a rapid star formation activity, is in practice indistinguishable from the old monolithic collapse picture. It is quite possible that massive ELLIPTICAL GALAXIES in clusters formed in such a manner. In another extreme scenario, mergers of large fragments occur over a more protracted time scale, comparable with the Hubble time, with most of the stars already formed in the merging units and some formed in merger-induced starbursts. It is likely that there was a full spectrum of galaxy assembly scenarios or formation histories at work.

The interplay of mass assembly and star formation fundamentally determines the galaxy morphology and the origin of disks in SPIRAL GALAXIES. In general, random merging of fragments leads to the formation of spheroidal (or, more accurately, triaxial) systems such as the bulges, stellar halos and dark halos, where kinetic energy needed to support the system against its own gravity is in random motions. In contrast to that, if protogalactic gas settles into a dark halo potential well gradually, it radiates away its kinetic energy, but retains most of the angular momentum (which can be only mechanically transferred away). PGs acquire their angular momenta through tidal interactions in the early universe, as passing-by mass concentrations exert tidal torques on each other. The gas then assumes the lowest-energy configuration possible for a given amount of angular momentum, which is a centrifugally supported, rotating thin disk (similar arguments apply to the origin of any disks in astronomy, including PROTOPLANETARY DISKS, ACCRETION DISKS, etc, not just DISK GALAXIES). Any stars formed within the gaseous disk then inherit the same kinematics.

Disks are dynamically fragile: major mergers would disrupt their orderly rotation and convert them to random motions. Thus, any disk galaxy is unlikely to have had a major merger since its disk was formed, but accretion of smaller satellites is still possible. Conversely, one often hears assertions that (some) ellipticals are made by merging spirals, which is indeed observed in the nearby universe. Obviously this is not the only possible path: a young elliptical may be made by merging a thousand small fragments early on. While some ellipticals may be formed in late, major mergers, many studies of elliptical galaxy properties in clusters at low and moderate redshifts suggest that most of their star formation was completed at high redshifts, $z > 3$ or higher, although some fraction of their stars may have formed since then.

At large redshifts, the merging was much more rapid and frequent, and the formation of durable, massive galactic disks was not likely. As the merging subsided and galaxies were carried apart by the universal expansion, the necessary gradual assembly of disks increased (incidentally, the same argument is used to explain the morphology–density relation in clusters and groups of galaxies: spirals are found mainly in low-density regions, and vice versa for ellipticals). Thus, we expect that the epoch of halo and bulge formation in galaxies in general preceded the epoch of disk formation, which is consistent with their star formation histories: bulges and ellipticals show little or no star formation today, and have spectra and colors characteristic of old stellar populations, where the bulk of the stars formed ~6–12 billion years ago, whereas the disks are actively forming stars now,

and apparently had similar star formation rates over their entire history.

To summarize, the expected history of galaxy formation may be something like this. The first PG fragments (gravitationally bound clumps in the early universe) may have had masses comparable with those of dwarf galaxies, $M \sim (10^6\text{–}10^8) M_\odot$ at the epoch of recombination, at $z \sim 1100$. Over the subsequent few billion years they merged hierarchically into ever more massive structures, stimulating star formation in the process. The first stars may have formed within a few hundred million years, at redshifts $z \sim 20 \pm 10$. The process of hierarchical assembly of PGs reached a peak maybe 1–3 billion years later, at redshifts $z \sim 2.5 \pm 1$. At subsequent epochs, this activity subsided as galaxies were carried apart by the universal expansion, and the supply of fragments and gas to merge with was diminished. Galaxy formation became GALAXY EVOLUTION. Most of the early star formation was in proto-ellipticals and bulges, with the disks forming on longer time scales.

Assembly of the mass into the present-day galaxies is one fundamental aspect of galaxy formation, but since most of the mass is dark this process is not readily observable. The other fundamental aspect is the conversion of primordial gas into stars and the early evolution of stellar populations, which is what we can observe. Let us then examine the sources of luminosity in PGs.

Luminosity of protogalaxies

The first is the release of excess binding energy, since PGs have collapsed. This is true regardless of the details of mass assembly, be it hierarchical merging or a monolithic collapse. A variety of arguments suggest that the PGs collapsed by about a factor of 10. First, the mean densities of galaxies exceed the extrapolated densities of the large-scale structure down to the typical radial scales of galaxies (a few kpc) by about a factor of $\sim 10^3$, suggesting a collapse by about a factor of 10. Second, the spinup of galactic disks requires a collapse by a similar radial factor. We can estimate the amount of the binding energy released by assuming that the present-day galaxies are in a virial equilibrium and estimating their total intrinsic kinetic energy. According to the virial theorem, this is equal to the amount of the excess binding energy lost in the process of virialization: a PG starts as a marginally bound system, with the kinetic energy equal to the potential one, but ends up as a virialized system where the kinetic energy is one half of that, with the other half presumably radiated away or carried out by galactic winds. An estimate of the kinetic energy of a normal, average galaxy today can be written as

$$|E_{\rm k}| \simeq \frac{1}{2} M_{\rm cool} \langle V_{\rm 3d}^2 \rangle$$

$$\simeq 6.2 \times 10^{58}\ {\rm erg} \times \frac{M_{\rm cool}}{10^{11} M_\odot} \left(\frac{V_{\rm 3d}}{250 {\rm km\ s^{-1}}} \right)^2$$

where $M_{\rm cool}$ is the total mass which can cool radiatively and $V_{\rm 3d}$ is the typical rms velocity. This estimate represents a minimum amount of energy released by a PG, even with no star formation. A comparable amount of binding energy is released in the collapse of protostars within a PG.

There are several possible mechanisms for release of this energy. The gas may be cooling via the inverse Compton effect on the photons of the CMBR. It may be radiated away in the shocks from colliding, still gaseous PG clouds and fragments. If the collapse happens on a free-fall time scale, $t_{\rm ff} \sim 10^8$ yr, the corresponding mean luminosity is $L_{\rm coll} \sim 2 \times 10^{43}$ erg s^{-1} $\sim 5 \times 10^9\ L_\odot$.

However, a much more energetically important process is nuclear burning in stars. It has been estimated that about a half of the STELLAR MASS in galaxies today is in the old populations, bulges and stellar halos, and thus about a half of all star formation in the universe's history can be associated with PGs (this is admittedly dependent on how one wishes to define PGs). The key to estimating this energy is through the resulting chemical enrichment: the same stars which produce the energy also produce the heavier elements. Most of the energy comes from hydrogen burning into helium, ~7 MeV per nucleon, i.e. ~0.7% or the initial rest mass of the processed hydrogen. The rest of the nucleosynthetic chain up to iron adds <0.1% to that.

Unfortunately, it is practically impossible to measure directly the accurate abundance of helium in the old STELLAR POPULATIONS today. Instead, we can use the observed abundances by mass of heavier elements and a conversion from chemical evolution models: $\Delta Y / \Delta Z = f$, where ΔY is the increase in the fractional mass abundance of helium, ΔZ is the increase in the fractional abundance of heavier elements (i.e. the mean metallicity), and f is typically found to be ~2–5. Since the initial abundance of heavier elements is essentially zero, $\Delta Z = Z$, which is observable. Estimates of the mean metallicity of old stellar populations are of the order of solar, i.e. $\langle Z \rangle \approx 0.01$–$0.02$. Thus the fraction of the rest mass of the primordial gas converted into energy during this initial chemical enrichment is

$$\Delta X \approx\ 0.07 f \langle Z \rangle \approx (4 \pm 2) \times\ 10^{-4}.$$

We can then estimate the total energy release from nuclear burning in stars responsible for this initial chemical enrichment:

$$E_{\rm nuc} \simeq M_\star\, c^2 \Delta X$$

$$\simeq 7.2 \times 10^{61} {\rm erg} \times \frac{M_\star}{10^{11} M_\odot} \frac{\Delta X}{4 \times\ 10^{-4}}$$

where $M_\star$ is the total mass burned in stars during the protogalactic starburst. This is much larger than the release of the binding energy and would thus be the dominant signal in detection of PGs. It may be a lower limit, since some of the produced metals must have been ejected in the resulting galactic winds and are not counted in our estimate of the mean metallicity of stellar populations today. Direct evidence for these ejected metals is in the intracluster gas seen in x-rays.

If this energy is released over a free-fall time scale, $t_{\rm ff} \sim 10^8$ yr, the corresponding bolometric luminosity would be $L_{\rm nuc} \sim 2 \times 10^{46}$ erg s^{-1} $\sim 5 \times 10^{12}\ L_\odot$, comparable with the luminosities of bright QUASARS and ultraluminous IRAS STARBURST GALAXIES today. Some authors have indeed proposed that such PG starbursts are the source of luminosity of quasars, but a majority belief is that quasars are powered by accretion onto massive BLACK HOLES. The mean star formation rate would be several hundred $M_\odot$ yr^{-1}. Even if this process proceeded on a more extended, hierarchical merging time scale $t_{hm} \sim 10^9$ yr, the resulting luminosities would still be an order of magnitude higher than the mean galaxy luminosity today (L_*) and comparable with that of the brightest cluster ellipticals today.

Finally, if formation of AGN accompanied some or all of these PG starbursts in young ellipticals and bulges, a comparable amount of energy may have been contributed by these early AGN. There is now growing evidence that massive black holes exist in the centers of most ellipticals and bulges, with the mean mass proportional to that of the stars, $M_{\rm bh} \sim 0.002 M_\star$. If about 10% of the rest mass of the black hole is radiated away in dissipative processes during its formation and growth (this is a typical estimate), then a comparable fraction of the bulge rest mass is radiated in the production of these ubiquitous black holes as in the nuclear burning in stars. It would be hard to separate observationally the effects of primordial starbursts in PGs housing quasars at large redshifts in terms of the energetics alone.

Thus, a 'generic' PG would have a typical bolometric luminosity $L_{\rm PG} \sim 10^{11\pm1} L_\odot$. To make an order-of-magnitude estimate of the observable fluxes, we note that in the relevant range of redshifts, $z \sim$ 3–10, and for a reasonable range of cosmological models, the luminosity distances are $D_{\rm L} \sim$ a few $\times 10^{29}$ cm, so that the expected bolometric fluxes are of the order of $\sim 10^{-15\pm1}$ erg cm^{-2} s^{-1}. Assuming for simplicity $L \sim \nu P_\nu$, the corresponding observed R band magnitudes would be $R \sim 25$ mag, or $F_\nu \sim 0.1$ mJy at $\lambda_{\rm obs} \approx 0.3$ mm, give or take an order of magnitude in flux. Much of this flux range is detectable with present (or soon forthcoming) instruments.

Detection of protogalaxies

The key question about detectability of PGs is that of obscuration. Most of the energy is released in the restframe UV, either from the photospheres of massive stars or from accretion disks in AGN, and would be now redshifted in the visible regime. On the other hand, if these PG starbursts were obscured by dust, the captured energy would be reradiated in the far-IR and would now be observable in the sub-mm regime (the AGN may also be detectable in x-rays). Approximately a third of all star formation in the present-day universe is obscured by dust, and perhaps a similar fraction applied to PGs (note, however, that some tens of millions of years of stellar evolution and chemical enrichment are needed before any dust can be produced, and the early PG phases must have been unobscured). This dichotomy of energy release channels leads to a dichotomy of detection strategies. We will discuss the two cases in turn.

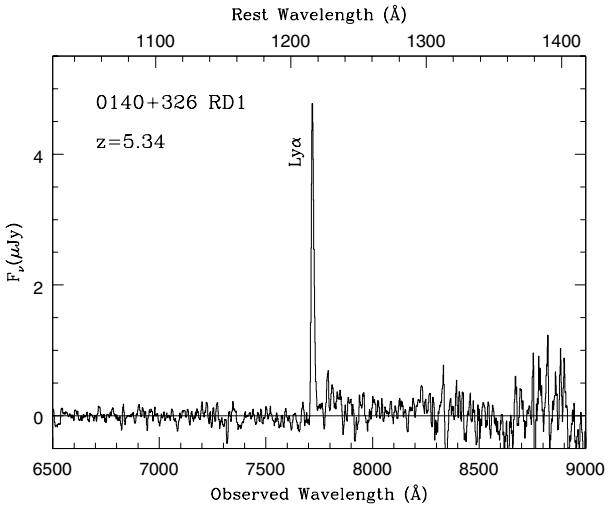

Figure 1. An example of a spectrum of a galaxy at $z = 5.34$, discovered serendipitously in a long-slit spectroscopic search. The Lyα emission line is very prominent. Objects like this can also be found in a narrow-band imaging search, provided that the line falls within the narrow filter bandpass. Reprinted with permission from Dey A *et al* 1998 *Astrophys. J. Lett.* **498** L93. © The American Astronomical Society.

Most of the early, and some of the modern searches for PGs were based on the detection of recombination emission lines, and most notably the Lyα λ 1215.7 Å line of the ionized hydrogen. For an unobscured, young stellar population with a normal IMF, several per cent of the total bolometric luminosity should emerge in this line. A useful rule of thumb is that about two Lyα photons are produced for every three ionizing (i.e. $\lambda < 912$ Å) photons. The line is readily detected (and is very luminous) in quasars, but its detection from galaxies powered by star formation was only achieved in 1990s. An approximate conversion is that unobscured star formation rate of $1 M_\odot$ yr^{-1} would produce the Lyα line luminosity of about 10^{41} erg s^{-1} in the restframe, but this number is highly sensitive to the form of the stellar initial mass function, since most of the ionizing photons are produced at the high-mass end but most of the mass is in the low-mass stars.

There are two observational approaches to detection of high-redshift, emission line objects, including PGs: narrow-band imaging and slit spectroscopy. Most of the emission line PG searches in the visible light have targeted the Lyα line, but exactly the same methodology was also used in the near-infrared, where searches targeted the Balmer lines of ionized hydrogen (mainly Hα λ 6563 and Hβ λ 4861) and the nebular lines of ionized oxygen, [O II] λ 3727 or [O III] λ 5007.

In the former, deep images are obtained in a narrow-band filter (usually an interference filter, or a Fabry–Perot tunable filter) and compared with images in a

broad bandpass with a comparable central wavelength. Objects with a strong emission line falling in the narrow bandpass would appear much brighter than in the broad band, where this line flux would be averaged out. This method can cover a relatively large area, but only a very narrow slice of redshift, corresponding to the redshifted line traversing the width of the filter, i.e. $\Delta z = \Delta\lambda/\lambda_c$, where $\Delta\lambda$ is the filter width and λ_c its central wavelength. Typically, $\Delta z \approx 0.01$–0.02 is used in PG searches.

The second approach is to obtain deep long-slit spectra of 'blank' patches of the sky, hoping that an emission line object would be covered by the slit. This method covers only a small area but a large range in redshift. One benefit of this approach is that almost any deep spectra taken for some other purpose can be used for a serendipitous search for PGs (figure 1).

The first detections of Lyα emission from GALAXIES AT HIGH REDSHIFT were from powerful RADIO GALAXIES, which are ostensibly ionized mainly by an active nucleus, and from quasar companion galaxies. However, no clear population of 'ordinary' PGs, i.e. powered by star formation, was found despite considerable efforts until the late 1990s. The early limits and the more recent detections correspond to modest star formation rates, up to a few $M_\odot$ yr^{-1}.

This conspicuous absence of a large population of Lyα luminous PGs was readily explainable by invoking extinction: both Lyα photons and the ionizing UV photons which produce them are absorbed by interstellar dust, which is likely to be present at some level in all but the most primordial star-forming regions. To complicate things further, Lyα photons are resonantly scattered by the interstellar hydrogen, which greatly increases their probability of being absorbed before escaping the PG.

The real breakthrough in the detection of high-redshift, star-forming galaxies came with the use of a different technique, the Lyman-break method (figure 2). Actively star-forming PGs would have relatively featureless, flat continua, except for one very significant feature: a strong drop in flux shortward of the Lyman limit, at 912 Å in the restframe, as the more energetic photons are readily absorbed by interstellar hydrogen. Thus, by imaging in at least three filters, one of which would cover the part of the spectrum shortward of the redshifted Lyman break, and two of which would sample the relatively flat continuum longward of the break, one can select candidate high-redshift galaxies through the resulting color signature. Actual spectra are then obtain to verify the candidates and measure their redshifts. This method is very efficient, and to date many hundreds of high-redshift galaxies (i.e. at $z > 2$ or so) have been identified in this manner.

At somewhat higher redshifts, i.e. $z > 4$ or so, a variant of this technique is used to detect both galaxies and quasars. There the net absorption by the intergalactic Lyα forest clouds along the line of sight decreases the flux shortward of the Lyα line in the object itself and sufficiently so as to produce a similar color signature. For example, quasars and PGs at $z \sim 4$–4.5 or so are very red in the $B - R$ or a similar color owing to this absorption effect, yet blue in the $R - I$ color, which is intrinsic and is not affected by the Lyα forest. This is actually the principal discovery technique for $z > 4$ quasars, but it works equally well for galaxies at comparable redshifts. The effectiveness of this technique increases with redshift, as the optical depth of the Lyα forest increases.

What if the star-forming regions are effectively completely hidden by dust? The absorbed energy is then reradiated in the restframe far-infrared and would now be observable in the sub-mm and mm region of the spectrum. Adequate detection technology in these wavelength regimes is just maturing, and first detections of high-redshift sub-mm sources (intrinsically very luminous ones) are just starting as of this writing (late 1999). This is clearly a very promising area of research, and it is likely that future observations in this regime will uncover a more substantial, optically obscured component of PG population.

We note that the integrated emission of unobscured and obscured galaxies, both forming and evolving, is observable in the form of diffuse optical and far-infrared backgrounds. Discussion of this fascinating topic is beyond the scope of the present article, but suffice it to say that the data so far indicate that the two components (obscured and unobscured) are roughly comparable in terms of the energy content.

A complementary approach to a direct detection of PGs in emission is to detect them in absorption, in the spectra of unrelated background quasars. Metallic line absorbers and high column density hydrogen absorbers are now believed to originate in galaxies or their halos, which at high redshifts may still be in the process of assembly. This approach eliminates all selection biases inherent in direct photometric detection, but it is limited in the information it can provide, primarily the gas column density, metallicity (an indicator of the chemical evolution) and perhaps rudimentary kinematics. Direct detections of a small number of high-redshift galaxies responsible for absorption systems have been made, and the data suggest that they are very similar to the general field population detected with the Lyman-break technique at comparable redshifts.

Formation of galaxies and that of the large-scale structure are closely related at early epochs. It is plausible that the first galaxies form in the densest spots of the primordial density field. These 'density peaks' are expected to be intrinsically highly clustered, as they would be mainly located within the larger overdensity regions. The first PGs may thus be found in the cores of future rich clusters, and then galaxy formation would gradually spread to ever lower density environments. This concept of so-called biased galaxy formation is generic to almost all models of structure formation. There is now some evidence that galaxies and quasars at $z > 3$ or 4 may indeed be strongly clustered. It should be noted that this primordial large-scale structure derives from initial conditions, with the first PGs and quasars lighting up

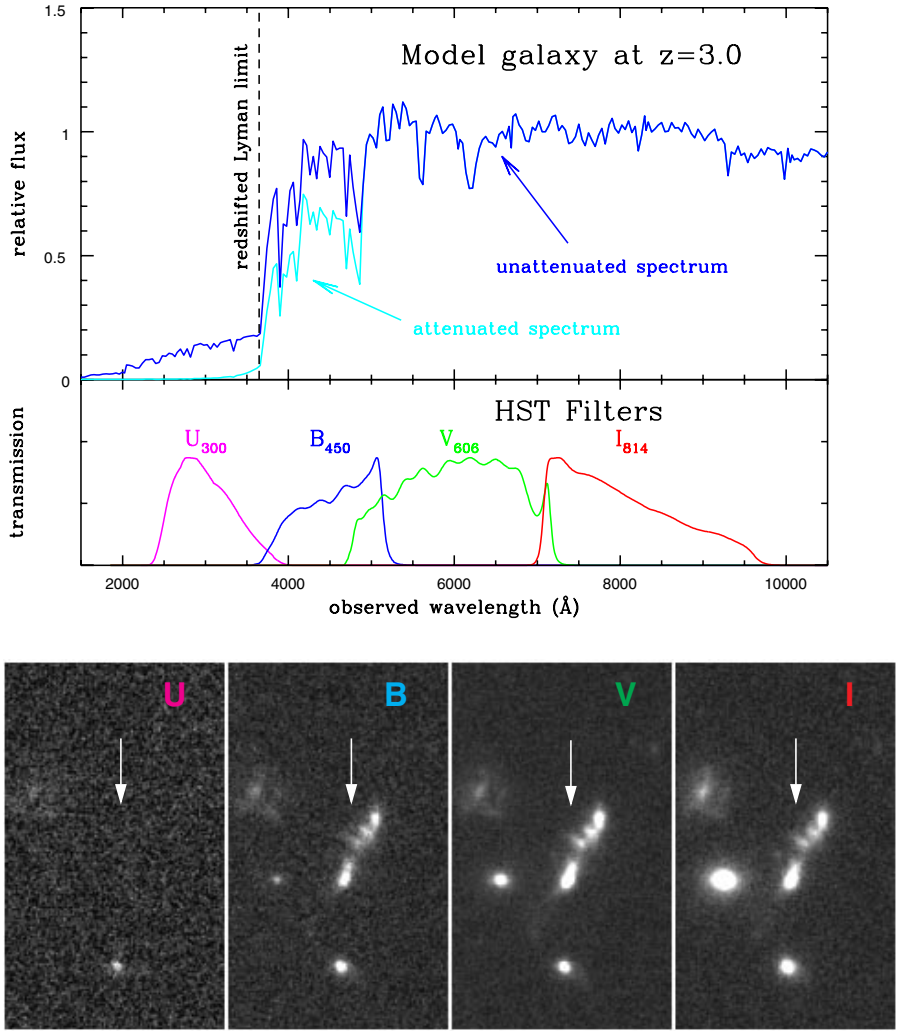

Figure 2. An illustration of the Lyman-break technique at work. The top panel shows a population synthesis model spectrum of an actively star-forming galaxy redshifted to $z = 3$, with and without attenuation due to the Lyα forest. Note the sharp drop in the continuum shortward of the Lyman break, at the observed wavelength of 912 Å $\times(1 + z)$. A set of transmission curves of filter used in the imaging of the Hubble Deep Field is shown below the spectrum. One would expect that such a galaxy would appear about equally bright in the B, V and R filters, but be practically undetectable in the U filter. This is indeed observed in the actual images of a galaxy at $z = 2.8$ found in the Hubble Deep Field images, shown in the bottom row. Reprinted with permission from Dickinson M 1998 *The Hubble Deep Field* ed M Livio *et al* (Cambridge: Cambridge University Press) p 219. © (The Cambridge University. Press).

the highest peaks of the primordial density field. This is distinct from the large-scale structure seen in the local universe, which grows gradually over the Hubble time through the gravitational instability.

The net result of the studies to date is at least a broad-brush understanding of the history of galaxy formation and evolution since $z \sim 5$ or so. The average star-forming activity has a broad maximum around $z \sim 2$ or 3, with a rapid decline since $z \sim 1$ or 2 towards the present day, $z \sim 0$. However, it is not clear whether we have yet seen the onset of galaxy formation. From absorption line studies we know that the universe was already fully ionized at $z \sim 5$, presumably by some combination of PGs and early AGN. While most of the stars may have formed since $z \sim 5$, clearly the first stars and thus PGs must have formed earlier.

We know that the universe entered 'dark ages' after the recombination at $z \sim 1100$, a few hundred thousand years after the big bang. Current theory suggests that the first protogalactic fragments started forming at $z \sim 10$–20 or so, a few hundred million years after the big bang. This 'cosmic renaissance' ended the dark ages and marked the onset of galaxy formation. The challenge for the future is to map this critical era, from $z \sim 20$ to $z \sim 5$, where the first PGs were forming.

Bibliography

This is a rapidly developing field, and perhaps the best way to learn about the up-to-date results is by browsing recent conference proceedings on cosmology, high-redshift universe, etc, of which there are many. However, the following reviews give more details on the subject of this article.

A good review of PG searches, problems, and techniques up to the mid-1990s is

Pritchet C 1994 *Publ. Astron. Soc. Pac.* **106** 1052–67

An updated review of searches for distant galaxies up to 1999 is

Stern D and Spinrad H 1999 *Publ. Astron. Soc. Pac.* **111** 1475–502

A good overview of the Lyman-break technique and some of the results is

Steidel C 2000 *Proc. Pritzker Symp. on the Status of Inflationary Cosmology* (Chicago, IL: University of Chicago Press)

S G Djorgovski