Physics – Uspekhi **68** (2) 177 – 186 (2025)

PHYSICS OF OUR DAYS

PACS numbers: 98.62.Ai, 98.80.Bp, 98.80.Es

70th ANNIVERSARY OF THE BIRTH OF V.A. RUBAKOV

Galaxies in the first billion years of the Universe's expansion

O.K. Sil'chenko

DOI: https://doi.org/10.3367/UFNe.2024.11.039810

Contents

Abstract. **The Universe started its expansion 13.8 billion years ago. A hundred million years later, the gas component of the Universe's matter 'matured' to the point of forming stars. This epoch should be inspected to search for the beginning of the emergence of galaxy populations in the Universe — large gravitationally bound star systems. Modern astronomical observation tools — ground-based interferometers and space-borne telescopes — allow the properties of galaxies to be studied directly at the earliest stages of their evolution, i.e., in the first billion years of the Universe's expansion, including their shapes, sizes, masses, star formation rates, and nuclear activity. This brief review presents the latest results of such studies.**

Keywords: evolution of galaxies, early Universe, astronomical observations

1. Introduction

Galaxies — the largest gravitationally bound conglomerates of stars and diffuse baryonic matter, gas, and dust — are the main visible populations of the Universe. Modern galaxies demonstrate various morphologies (shapes), different kinematics (rotation), and diverse masses of the stellar component — from the most modest, hundreds of solar masses (ultra-faint dwarf satellites of the Milky Way), to hundreds of billions of solar masses (central spheroidal galaxies in clusters). Figure 1 shows a classification of the morphological types of galaxies, the idea of which was published in its final form by Edwin Hubble in 1936 [1]. At the roughest level, galaxies are divided into spheroidal, with a uniform structure and stellar population, and disk, with large stellar disks having a nonuniform structure. The radial extension of the disks is much greater than their thickness; to a first approximation, they can be considered thin and round stellar systems, the kinetic energy of which is concentrated in an ordered circular rotation. However, disk galaxies are usually structurally two-component: spheroidal subsystems called bulges are often observed in their centers, and the classification of galaxies as 'early-type' or 'late-type' as well as the refinement of their position on the Hubble sequence, is largely determined by the contribution of the bulge to the total luminosity (mass) of the galaxy.

We still use the Hubble scheme to classify galaxies; moreover, high spatial resolution Hubble Space Telescope (HST) images of deep fields have shown that most galaxies in the Universe have approximately the same morphology as those around us before the redshift $z \approx 1$, i.e., in the last eight-to-nine billion years [2]. But how exactly did such forms of galaxies develop? What was their origin? And how did galaxies begin? Astronomers are currently looking for answers to these questions, and the source of our knowledge in this area is both cosmological models that prescribe a strict time scale for the appearance of galaxies in the Universe and fix their initial forms and subsequent evolution, and direct observations of galaxies at different stages of their evolution.

The proposal to directly *observe* the evolution of galaxies over several (many) billion years seems exotic only at first glance. In fact, nature itself provides us with almost exhaustive opportunities for the consistent study of galaxies in the process of their evolution. Due to the finite speed of light, the further the galaxy being studied is from us, the longer it took the light emitted by it to reach us; accordingly, observing it now with our local tools, we actually see it as it was billions of years ago. Cosmological models that describe the uneven expansion of the Universe allow us to relate the apparent redshift of a galaxy to the distance to it. Thus, by measuring the redshift of a galaxy, we simultaneously estimate how long it took its light to reach us. At a redshift of $z = 0.5$, we observe galaxies as they were 5 billion years

O.K. Sil'chenko
Lomonosov Moscow State University,
Sternberg State Astronomical Institute,
Universitetskii prosp. 13, 119234 Moscow, Russian Federation
E-mail: sil@sai.msu.ru, olga@sai.msu.su

Received 12 September 2024, revised 31 October 2024
Uspekhi Fizicheskikh Nauk **195** (2) 188 – 198 (2025)
Translated by A.I. Ulitkin

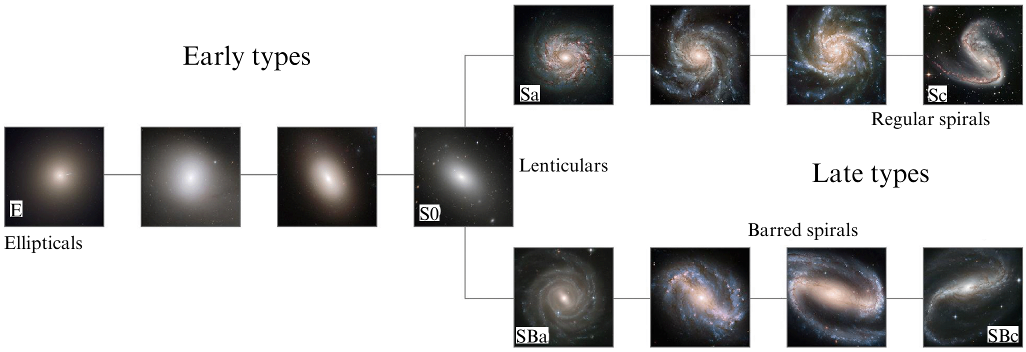

Figure 1. Hubble's galaxy morphological classification.

ago, and at a redshift of $z = 1$, we perceive them as they were 8 billion years ago. At a redshift of $z = 5$, we see galaxies 1 billion years after the Big Bang, which is not far from observing the event of the emergence of the first galaxies.

The James Webb Space Telescope (JWST) [3], launched on December 25, 2021, was designed to see further back into the history of the very first galaxies than ever before and released its first scientific results by the end of 2022. The JWST has two major advantages over the HST that allow hope for successful observations of the most distant and very first galaxies. First, the JWST's mirror is almost three times larger than Hubble's; observations from space, where Earth's atmosphere does not blur the images of the observed objects with its turbulence, make it possible not only to collect more light, but also to obtain images with a three times better spatial resolution in the same spectral range. This means that the JWST enables astronomers to see faint compact objects that are smeared over a larger number of receiver pixels in the HST data and therefore become poorly distinguishable above the sky background. The second advantage of the JWST is that, in contrast to Hubble, its sensitivity range is shifted to the infrared region of the spectrum. The most distant galaxies also exhibit the largest red shifts; this means that the maximum in the spectral energy distribution, assuming a normal stellar population, shifts from the optical to the near-infrared range. The long-wavelength limit of HST observations is the H filter (1.6 μm). The JWST has two main instruments: the NIRCam with its seven filters covers the range from 1 to 4.5 μm, while the MIRI can make observations in the range from 5 to 28 μm. Both instruments also have a spectral mode of observations. As a result, the first deep exposures of the NIRCam presented the scientific world with galaxies at red shifts of up to, presumably, 17.

And what did theorists expect to see at such red shifts? Figure 2 shows a picture drawn in 2003, at the dawn of the JWST development, to illustrate the expectations of the telescope for studying the early Universe. Generally speaking, no galaxies were expected at a redshift of $z = 20$; theorists thought that at this redshift only the very first single stars were born, just 200 million years after the Big Bang. Galaxies were expected to appear around $z = 10$. A contemporary, brilliant review by Bromm and Larson [4] on the formation of the first stars explains the complex interaction between the temperature of the cosmic microwave background, the virial temperature of the gas consisting of hydrogen and helium without an admixture of heavier elements (the gas of primordial chemical composition is a product of primordial nucleosynthesis in the first 20 minutes after the Big Bang) inside gravitationally bound, 'collapsed' dark matter halos,

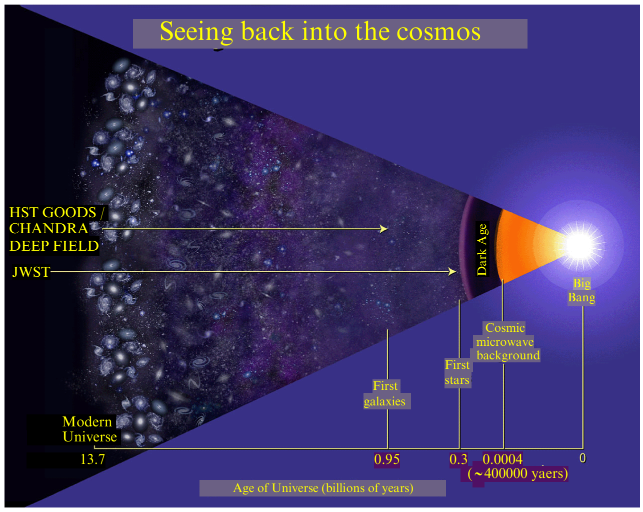

Figure 2. Concepts of time scales of Universe's evolution on eve of the launch of JWST project. Credit: NASA/ESA and Ann Feild (STScI).

and the thermal conditions necessary for the onset of star formation. These dark halos, each with a mass of a million solar masses, are capable of becoming the birthplace of the first, so-called Population III stars; Population III stars should initially form at a rate of one to five stars per halo. These are not yet galaxies — star formation inside such minihalos ceases very quickly due to the reverse injection of energy into the gas from young stars and supernovae, into which they rapidly evolve by the end of their short, 2 million year, life path. And more populated star systems are expected only when dark halos with a mass of 10^8 solar masses 'gather' in the Universe [5]. The entire evolution of matter distribution in the Universe within the framework of the modern cosmological model is subject to a strict hierarchy of masses and sizes. It is necessary to 'start' with an almost uniform distribution of matter in the Universe; first, small density fluctuations arise, which grow and then merge into ever larger ones, and thus the typical mass of the formed dark halo is strictly tied to the epoch when it was formed, i.e., to the red shift. Barkana and Loeb [6] demonstrate in their review how, by comparing the characteristic mass of the dark halo and the corresponding virial temperature with the temperature of the cosmic microwave background, decreasing with the Universe's expansion, one can determine the redshift at which the formation of stars is possible. With a halo mass of 10^6 solar masses and cooling on H_2 molecules (primordial gas), the value of z amounts to 16–24, while with a halo mass of 10^8 solar masses and cooling of gas already enriched in heavy elements by more traditional mechanisms, z is 8–12. These are the estimates on which expectations for the discovery of the first 'luminous' objects, stars, and galaxies by the JWST were based. What did the first data from the JWST actually show?

2. Ly-break galaxies at $z > 10$

But let us first dwell on the most effective, simple, and large-scale method for discovering galaxies at high redshifts and how it limits the properties of these galaxies *a priori*. That is, which distant galaxies we can discover and which we cannot, even if they exist.

In the 1990s, Charles C. Steidel et al. [7] proposed the Lyman break technique to search for distant galaxies. At first, the discussion was about the massive discovery of galaxies at a redshift of $z \sim 3$. The technique was based on a simple physical model. If we want to discover a very young galaxy that has only recently begun to form stars, such a galaxy should still have a lot of gas. Star formation most likely began in the center of the galaxy, where the gas density is higher, and neutral gas, not captured by star formation, should be concentrated on the periphery. Then, the light from the central cluster of young massive blue stars must necessarily pass through a dense 'coat' of neutral hydrogen, which will absorb all photons of the spectral continuum at a wavelength shorter than 912 Å (in the system of wavelengths of the galaxy) — this is the Lyman jump (Ly break). If we observe such a galaxy at a redshift of $z = 3$, the Lyman jump of 912 Å will move to the red side, and its wavelength will increase by a factor of $(1 + z)$, i.e., for a galaxy at $z = 3$, it will become equal to 3648 Å. This is the ultraviolet observed from Earth; for example, in the well-known Johnson photometric system, this wavelength falls into the U filter. Thus, Charles Steidel selected galaxies from ground-based deep photometric observations that are **not visible** in the U filter, but are bright in the neighboring blue B filter; and subsequent spectral observations confirmed that these galaxies are indeed located at $z = 3$ [7]. It then turned out to be very easy to extend this method to search for even more distant galaxies: those that are **not visible** in the B filter are at $z = 4$, those that are **not visible** in the V filter are at $z = 5$, and so on. Simple arithmetic shows that in order to detect a galaxy at $z = 10$, one needs to look for an object that is invisible at a wavelength of 1 μm and is bright, for example, in the H filter (1.6 μm). The Hubble Space Telescope could also do this: its record-breaking distant galaxy, GN-z11, is located at a redshift of $z \approx 11$ [8]. And the capabilities of the JWST allow it to search for — and successfully find — even more distant young galaxies. It is important to remember that the search method itself, the Ly-break technique, limits the possible characteristics of such galaxies: these are galaxies in which star formation has begun recently, and it most likely proceeds at a moderate rate, on the order of a few ten solar masses per year or several solar masses per year.

To take advantage of the greater penetrating power and better spatial resolution of JWST images, the first deep surveys with the NIRCam camera were aimed at areas of the sky that had already been studied in detail in a wide range of wavelengths, from X-rays to submillimeter, by the Hubble Space Telescope, large ground-based optical telescopes, and space telescopes. A total of 13 such areas were selected. The very first detailed report was received from the CEERS (Cosmic Evolution Early Release Science) area of 31.7 squared arc minutes, previously known as part of the CANDELS/EGS project. With particular excitement, these data, which included images in seven filters of the NIRCam instrument covering the range from 1 μm to 4.5 μm, were searched for Ly-break galaxies at redshifts $z > 10$; of these, 26 were initially announced in the range $9 < z < 16$ [9], and later the sample was increased to 88 galaxies in the range $8.5 < z < 14.5$ [10]. More precisely, the 'breaks' found in the JWST multi-wavelength photometric data for the $8.5 < z < 14.5$ candidates are breaks on the blue side of the Lyα emission, because at such redshifts the intergalactic medium is still neutral, and the hydrogen of the intergalactic medium, not directly associated with the galaxy, absorbs the light in the Lyα lines at all intermediate redshifts $z < z_{gal}$. In the first release [9], the champion in distance was the galaxy at $z = 16.6$, but later another redshift was chosen for it, $z = 4.9$ [11], because this break technique does not give a 100% certainty that we see exactly the Lyman jump — it could also be a Balmer jump or unusually wide emission lines that fall into several filters at once, as in this particular candidate [11]. Only spectral observations, when emission lines identified with specific atoms and ions are detected in the spectrum, give a 100% certainty of the redshift. Thus, for a number of galaxies from the CEERS sample, the previously estimated photometric redshifts $z \approx 9$ were spectrally confirmed when emission lines [OIII]λ4959,λ5007, shifted by the redshift to a wavelength of 4 μm, were detected and measured in their spectra [12] (Fig. 3). For another candidate at $z = 16.4$, CEERS-93316 [13], additional data helped in understanding its redshift, namely, photometric observations at a wavelength of 1 mm, carried out within the framework of the submillimeter survey SCUBA-2: the measured flux in the far-infrared range forces one to identify this galaxy as a very dusty galaxy with intense star formation at $z \approx 5$ [14] (see more about galaxies of this type below — so-called submillimeter galaxies, an order of magnitude more massive than Ly-break galaxies).

Thus, it is presently considered that the initially declared galaxies at $z = 16-17$ are not confirmed. However, in the sample [10], there are three objects with photometric redshifts of $z \approx 14$ and about 30 at $z = 10-12$, some of which have been confirmed spectrally. This is already enough to apply statistics. Using these data, the ultraviolet luminosity functions of galaxies at $z \approx 11$ have already been constructed and the integral volume cosmic densities of ultraviolet radiation of galaxies have been calculated, which are easily converted into cosmic densities of star formation rates. Finkelstein et al. [9, 10] formulated some results related to the evolution of galaxies up to $z = 12$, and these results decisively diverge from the predictions of the most sophisticated modern numerical simulations of the cosmological evolution of the Universe. In particular, the number of UV-bright galaxies at $z > 10$ is more than an order of magnitude greater than that predicted by the cosmological Illustris-TNG or SIMBA simulations. If we try to compare the absolute stellar magnitude from observations and the dark halo mass from the models (abundance matching method) [9], then the dark halo mass of 4×10^{10} solar masses corresponds in observations [9] to a galaxy twice as bright as expected at $z = 11$, according to the luminosity-to-dark halo mass relation predicted by the extrapolation method (from data at lower redshifts). Surprisingly, the evolution of the volume density of UV-bright galaxies at redshifts $z > 8$ reaches a plateau up to $z \approx 12$, while between $z = 4$ and $z = 8$ the volume density decreases, in agreement with the cosmic history of star

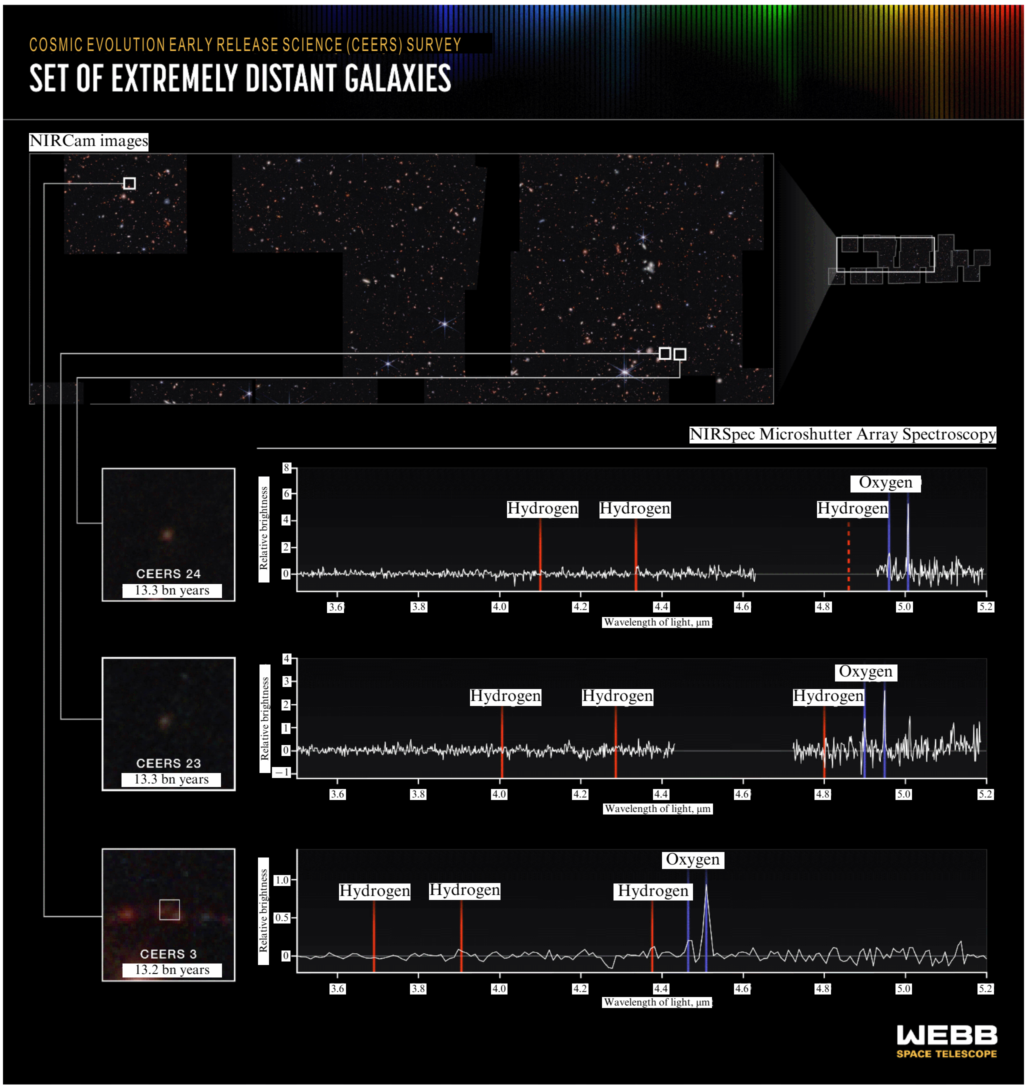

Figure 3. Examples of emission-line spectra of distant galaxies obtained by JWST NIRspec spectrograph. Credit: NASA, ESA, CSA, L. Hustak (STScI), S. Finkelstein (UT Austin), S. Fujimoto (UT Austin), and P. Arrabal Haro (NSF's NOIRLab).

formation, which shows a flat maximum between $z = 2$ and $z = 4$ and decreases at higher redshifts [15].

The second survey, a full photometry for objects of which appeared in literature already in 2024, is the JWST Advanced Deep Extragalactic Survey (JADES): on an area of 125 squared arc minutes, 717 galaxies with $z_{\rm phot} > 8$ have already been discovered, including 31 galaxies with $z_{\rm phot} > 12$ [16]. This survey covers two fields previously studied with both the HST and the large optical telescopes, GOODS-N and GOODS-S. For many objects in the JADES sample, there is high-precision photometry in the optical region of the spectrum based on measurements by the Hubble telescope, which makes it possible to refine the photometric redshifts and construct a more complete energy distribution in the spectrum. Since the statistics on objects here are already more abundant than that in the CEERS survey, it was possible to extend the evolutionary sequences to $z \sim 14$. Finally, it seems that it gives a hint to a drop in the cosmic volume density of ultraviolet luminosity between $z = 12$ and $z = 14$ by 2.5 times [17]. Can we hope that the JWST has reached the beginning of the era of galaxy formation at $z \sim 15$, including those dwarf galaxies inside dark halos with a mass of 10^8 solar masses that cosmologists promised? But for one of the most distant discovered galaxies, JADES-183348 at $z = 14.4$, it was possible to measure the stellar mass quite accurately, to a factor better than 2: it turned out to be equal to 10^9 solar masses [17]. This is a couple of orders of magnitude more than that expected for the very first galaxies predicted based on model considerations [5, 6].

3. How to explain observations of galaxies at $z > 10$ in order to remove contradiction with LCDM model

Thus, it was a complete surprise to find galaxies at $z > 10$ with stellar masses on the order of $10^9 - 5 \times 10^9$ solar masses [18], and even galaxies an order of magnitude less massive at $z > 10$ were not expected within the framework of standard cosmology [6]. To remain within the classical hierarchy and in agreement with calculations of the rate of assembly of dark halos during the evolution of the Universe, we should assume that, at least for the brightest and most massive galaxies, the relationship between the stellar mass of a galaxy and the mass of the dark halo within which it formed is not the same as in the later Universe. Generally speaking, the identical spatial clustering of galaxies with stellar masses of $10^9 M_\odot$ and

$10^{10} M_\odot$ at $z > 10$ [18] seems to indicate that galaxies with such different stellar masses 'live' inside dark halos of *the same mass*. This means that the star formation rate (SFR) in these galaxies is not related to the rate of dark matter accretion onto their halos, which is markedly different from previously formulated ideas and theoretical approaches to constructing models of galaxy evolution based on observations of objects at $z < 8$ (see, for example, [19]).

What do observations tell us about the star formation regime in galaxies observed after the first billion years of the Universe's expansion have passed? Up to $z \sim 5$, the overwhelming majority of galaxies belong to the so-called 'main sequence.' This scaling relation, which characterizes a close correlation between the current star formation rate and the current stellar mass of the galaxy, was discovered for the nearby Universe in the SDSS survey [20], and then confirmed for earlier epochs [21] in other surveys. If we divide the stellar mass of a galaxy by its current star formation rate, it turns out that, at any redshift, the time needed to build up its stellar mass, assuming comparable rates of star formation, is on the order of the age of the Universe at a given redshift. Outside the main sequence and the spread of points determined by the accuracy of measuring the star formation rate, only 2–3% of all massive galaxies are found [21, 22]. Consequently, the vast majority of galaxies form stars in a smooth regime throughout their entire life, with an exponential star formation history and a characteristic decay time of about 7 billion years [21]. At the same time, violent bursts of star formation, inevitable, for example, for a large merger, when the star formation rate on a scale of $< 10^8$ years can increase by five or more times, are extremely rare. After this observational discovery, researchers began to relate the star formation rate to the external gas accretion rate through physical models and then to the accretion rate of dark matter onto the parent halo of the galaxy through cosmological models. It seems that at $z > 9-10$ such harmonious physics become inapplicable to the description of the real construction of the stellar component of galaxies.

What are the alternatives? Here is a short list of ideas proposed to explain the observed excess of bright (and possibly overly massive) galaxies at $z = 9-12$, which, after summing up observations from several JWST deep fields (CEERS, JADES, PRIMER, UNCOVER, and COSMOS-Web [18, 23, 24]), began to be statistically convincing. Ferrara et al. [25] proposed to completely remove dust from such galaxies in order to ensure a qualitative difference in the absorption of light from young stars between very young galaxies and those observed half a billion years after the Big Bang. But this should correspond to a very rapid chemical evolution, especially in the redshift range $z = 8-12$; meanwhile, there are no signs of the presence of Population III in the discovered young galaxies, and low-metallicity gas and stars are not observed in their measured spectra. It was also proposed to change the initial stellar mass function — qualitatively in the direction of theoretical expectations for Population III, but not so radically — to decrease only the slope in order to add massive stars. Then, the transition from the UV luminosity produced by massive stars to the total stellar mass will also be 'flatter,' and estimates of the stellar mass of galaxies will decrease. But these so-called 'top-heavy' initial stellar mass functions will lead to a very powerful 'feedback' from star formation — heating of gas by the stellar wind of massive stars and stopping of subsequent star formation, which we would, on the contrary, need to strengthen compared to model expectations. Hence, this is also a bad idea [26]. However, if it is necessary to reduce the feedback for young galaxies, one can also focus on the external UV background, which heats gas in galaxies in the Universe at lower redshifts, thereby reducing the star formation efficiency. If we remove this external ultraviolet, then the star formation efficiency (the fraction of the cold gas mass that will turn into stars in a single star formation event) at $z = 10-12$ will increase significantly, and the level of the cosmic star formation history (SFH) at such redshifts will exceed expectations from the SFH extrapolation calculated for a gas-to-star conversion efficiency of only a few percent [27]. There is a physical justification for disregarding the external UV background at the given redshifts: if reionization of the intergalactic medium began precisely at this epoch, at $z \sim 10$, then the intergalactic medium was transparent and opaque to UV light escaping from star-forming galaxies at $z < 10$ and $z > 10$, respectively, since at $z > 10$ it was still completely neutral. Consequently, one can expect that intergalactic neutral hydrogen at $z > 10$ should absorb all UV light with a wavelength shorter than the Lyman limit. Thus, this is a good, useful idea, as argued in [27].

But perhaps the most attractive idea (although with an unclear physical justification) is that of a change in the star formation regime from smooth to bursty in the range of redshifts from 5 to 9, i.e., between half a billion and one billion years after the Big Bang. When statistics on stellar mass (M_*) and SFR determinations for large samples of galaxies at redshifts up to 4–5 were collected, a difference began to emerge in the distribution of such early galaxies in the SFR vs M_* diagram from that which is observed at $z = 0$ or even at $z = 2$. Caputi et al. [28] studied a large sample of galaxies — about 2000 objects — from the Spitzer Matching Survey of the Ultra VISTA Ultra-deep Stripes (SMUVS) with Hα emission lines in the range $z = 3.9-4.9$. Their distribution in the SFR vs M_* diagram turned out to be clearly bimodal: in addition to the main sequence, where massive galaxies, $2 \times 10^9 - 4 \times 10^{10} M_\odot$, have a stellar mass buildup time of about 1 billion years, there is also a band of object concentration parallel to the main sequence, which is an order of magnitude higher in the star formation rate. This parallel sequence contains galaxies at the moment of a star formation burst. Caputi et al. [28] noted that the fraction of galaxies at $z = 3.9-4.9$ caught at the moment of a star formation burst depends on the stellar mass of the galaxy, and for the average mass of Ly-break galaxies, $10^{10} M_\odot$, it is about 10%. In a subsequent study, Rinaldi et al. [29] expanded the sample by involving an additional site and increasing the range of redshifts and stellar masses in question, with the threshold value of the latter decreasing to $10^6 M_\odot$; now, the studied sample of galaxies at $z = 4-5$ has increased to more than 4000 objects. The bimodality of the distribution of galaxies in the SFR vs M_* diagram has been fully confirmed. For the previously considered range of stellar masses, $3 \times 10^9 - 3 \times 10^{10} M_\odot$, the fraction of galaxies in the burst star formation regime has now been determined to be 30%; however, at large redshifts, $z = 5-6.5$, the fraction of galaxies in this range of stellar masses drops to 12% in the burst regime (according to the presented data). However, the strong bimodality of the distribution of observed galaxies in the SFR vs M_* diagram, with the presence of a parallel sequence of starburst galaxies, is a surprising and impressive result that was not at all predicted by cosmological models, such as Illustris TNG [29].

With the first releases of deep JWST surveys, it became possible to extend such studies to large redshifts. Hodge et al. [30] analyzed the photometry of objects from the JADES survey: 3197 galaxies in the interval $6 < z < 7$, 1883 galaxies in the interval $7 < z < 9$, and 521 galaxies in the interval $9 < z < 12$. The availability of measurements in 23 photometric bands makes it possible to reconstruct recent (over the past 10^8 years) star formation history, and by its tilt, positive or negative, to classify a galaxy as being in a starburst or belonging to the main sequence [30]. Bimodality of star formation regimes was again found. If we consider the most massive part of the sample, $\log M_* > 8.8$, then about 87% of galaxies in the interval $6 < z < 7$ are in the main sequence. At large redshifts, the fraction of galaxies in the main sequence and in the burst regime are already comparable, and at $z > 9$, 90% of massive galaxies demonstrate the burst star formation regime. The evolution of the star formation regime is evident at a qualitative level: in the first half a billion years of the Universe's expansion, star formation in galaxies, even in the most massive ones, proceeded in the regime of violent bursts, with a time scale of several ten million years, and the physical nature of such a star formation regime has not yet been 'captured' in the standard cosmological model.

But, if this was really the case, then this would explain the anomalously bright early galaxies in the ultraviolet and the unexpectedly high cosmic density of star formation rates at $z = 10-12$: we only see galaxies at the moment of a star formation burst and do not see the same galaxies between star formation bursts.

4. Massive starburst galaxies at $z = 4-5$

However, Ly-break galaxies are not the only population of galaxies at large redshifts. The most massive ones at redshifts from 2 to 5 are submillimeter galaxies (SMGs) [31, 32]. These galaxies are 'searched for' in a completely different way than Ly-break objects, and, accordingly, have other characteristic properties. They are massive galaxies, with a typical stellar mass of about $10^{11} M_{\odot}$ and star formation rates from a hundred to a thousand or even up to three thousand solar masses per year [33]. They are no longer main sequence galaxies, but starburst galaxies, shifted upward on the SFR vs M_* diagram by at least four times in star formation rates. Such extreme star formations are always (both at large redshifts and in the local Universe) immersed in a dense dust cocoon; only the blackbody spectrum of dust heated by massive stars of young star clusters inside the cocoon emerges to the outside. At $z = 0$, the maximum of the energy distribution in the spectrum of such galaxies lies at a wavelength of 100 μm; accordingly, at $z > 3$, this maximum shifts to 400 μm and longer wavelengths. Submillimeter galaxies were discovered in the 1990s by the James Clerk Maxwell Telescope (JCMT) in Hawaii as a result of the SCUBA project, the first sky survey in two transparency windows of Earth's atmosphere, at 450 and 850 μm [34]. A whole series of objects bright at submillimeter wavelengths was discovered, and, indeed, when their redshifts were sorted out, these objects turned out to be massive galaxies with an intense burst of star formation at $z = 2-5$ [35]. Thus, already a billion years after the Big Bang, there is a whole population of massive galaxies in the Universe, with $M_* > 10^{10} M_{\odot}$, which also continue to form stars at a rate of a hundred solar masses per year. Accordingly, over the second billion years of the Universe's expansion, they should at least double their stellar mass.

Currently, SMGs are detected and studied 'en masse.' In addition to the JCMT SCUBA-2 survey [36], single-dish telescope in the Atacama Desert in Chile (LABOCA, the ALESS project) [37] and a specialized submillimeter telescope in Antarctica, the South Pole Telescope (SPT) [38], have been used to compile a list of SMG candidates, because a very dry atmosphere there transmits submillimeter radiation well. These single dishes make up the lists of candidates for SMGs, since the spatial resolution of single submillimeter dishes is worse than $10''$. Then, the galaxies from the obtained lists are identified and studied in detail using the Atacama Large Millimeter Array (ALMA) interferometer — a multi-wavelength instrument located in the same dry high-altitude Atacama Desert; therefore, it can help build maps both in the continuum at submillimeter wavelengths (hot dust) and in the CO, [OIII]λ88-μm, and [CII]λ158-μm lines (molecular and cold and warm atomic gas), with a resolution of better than a second. The ground-based interferometer in the Atacama Desert competes with HST in the spatial resolution of the obtained images of distant galaxies and sometimes with the JWST [39]!

Recently, completely unexpected results regarding the dynamic status of galactic disks have appeared based on the results of ALMA observations of the most distant SMGs and dusty star-forming galaxies (DSFGs) at $z = 4-5$ [40–43]. When comparing images in the continuum (heated dust) and in the emission line of atomic ([CII]λ158 μm) gas, it turns out that the distribution of cold neutral gas in galaxies is much more extended than the size of the dusty central region with intense star formation. That is, we are dealing with a violent burst of star formation in the center of a 'quiet' gas disk with a radius of several kiloparsecs — this is the birth of a bulge in a disk galaxy rather than the birth of an elliptical galaxy. Moreover, the gas disks of such galaxies turn out to be cold not only in temperature: they are dynamically cold — thin, up to 300 pc, with low velocity dispersion! They are less turbulent than the gas disks of galaxies at $z = 2$. The ALMA observations of distant faint galaxies with a resolution of $0.1''-0.2''$ (corresponding to a linear scale of 0.7–1.4 kpc at redshifts of 4–5) are very labor-consuming and require long exposures. Nevertheless, fourteen objects with such dynamic characteristics have already been identified. Measuring the line-of-sight velocities of cold neutral gas in the [CII] line over the entire disk of the galaxy made it possible to reconstruct the pattern of its rotation; it turned out to be regular and circular (without disturbances or traces of merging). Thus, the typical ratios of the rotation velocity to the velocity dispersion of gas clouds in massive galaxies, $\geqslant 10^{10} M_{\odot}$, at $z = 4-5$ turned out to be from 7 to 15, as in giant spiral galaxies of the local Universe.

Why did no one expect this? Previous studies of disk rotation in massive galaxies at $z = 2-3$ demonstrated the presence of thick (more than a kiloparsec thick) gas disks with a gas cloud velocity dispersion of up to 70–80 km s^{-1} [44]. When this was discovered, it was decided that it was a demonstration of the nature of turbulent gas cloud velocities associated with local gravitational instability. But then, the velocity dispersion simply depends on the amount of gas and is proportional to the product of the galaxy's rotation velocity and the fraction of gas in the galaxy's baryonic mass [44]. At $z = 2-3$, the gas velocity dispersion is higher than that at $z = 0$, because the fraction of gas in the baryonic mass of spiral galaxies at $z = 0$ does not exceed 10%, while it is already 40% at $z = 2-3$. Then, the typical velocity dispersion in gas disks with a redshift was expected to increase monotonically, because, in very young galaxies, gas should already dominate the baryonic mass of the galaxy. Non-monotonic evolution of the characteristic velocity dispersion of gas clouds in large-scale disks of massive galaxies is a completely unexpected and so far unexplained observational phenomenon. At least cosmological simulations, which include gravitational instability as a trigger for star formation and a source of energy release associated with this process, have so far promised a monotonic evolution of both the velocity dispersion of gas clouds and the ratio of the rotation velocity to the velocity dispersion. For example, TNG50 simulations from the Illustris project [45] predicted a monotonic drop in v/σ from 10 at $z = 0$ to a maximum of 3 at $z = 4$. However, one can notice a qualitative similarity in the evolution of the cosmic density of star formation rates and the velocity dispersion of gas clouds. This may be a hint that the nature of the turbulence of gas disks is not gravitational instability but the feedback effect, i.e., the supply of kinetic energy to the interstellar medium by young massive stars through their stellar wind and supernova explosions. Indeed, approximate model simulations of this effect, relating the velocity dispersion of gas clouds with the feedback effect of star formation, show that starburst galaxies at $z = 4-5$ fit precisely into the framework of this model [41, 46].

After TNG50 simulations from the Illustris project [45] failed to explain the observed cold disks of galaxies at $z = 4-5$, cosmologists put forward several more approaches to this problem. An authoritative team led by Avishai Dekel [47] attempted to search for an analogue of such objects in physical zoom-in cosmological models, using high-resolution dark matter simulation (55 pc) and the RAMSES gas-dynamic code. Their efforts resulted in the discovery of only one such model galaxy. It managed to maintain a thin gas disk at $z \approx 4$ for five rotation cycles due to a long (several hundred million years) episode of laminar coplanar accretion of cold gas from the cosmological filament, after which the disk of this model galaxy was destroyed by merging and a change in the feeding filament. The conclusion from this approach of the theorists to explaining the new observed class of galaxies was the following: cold gas disks of early galaxies are a transient phase, and such galaxies should be very few in number. In their recent calculation paper, Kohandel et al. [48] have presented SERRA cosmological numerical models with an adaptive resolution reaching 30 pc in the densest places; statistics on model galaxies in the mass range $10.0 < \log M_* < 10.3$ already number 142 objects in the redshift range $4 < z < 9$. Truly, the star formation rate in this model sample does not exceed 128 $M_\odot$ per year — this is at best the main sequence at such redshifts, not a star formation burst. But the necessary dynamic coldness precisely for the indicator of random motions of neutral gas clouds in the form of the [CII]λ158-μm line is observed in this group of model galaxies *on average*: the model range is $V/\sigma_{[CII]} = 8.5 \pm 2.2$ and does not depend on z in the redshift range $4 < z < 9$. This time, the authors claim that the dynamic coldness of their model galaxies is not transient, because it is traced over 10 revolutions; however, in absolute time units, this is still only 200 million years.

The situation became completely confusing when Parlanti et al. [49] considered observations of a sample of 22 galaxies in the redshift interval $4.2 < z < 7.6$, based on ALMA measurements of the [CII]λ158-μm line. Truly, this collection of archival observational data is very heterogeneous in quality: the spatial resolution varies from $0.2''$ to $1.5''$. In this sample, only main-sequence galaxies were deliberately left out, and SMGs were deliberately ruled out. If we limit ourselves to only the interval $5 < z < 7$, where the spatial resolution in linear measure is no worse than 1 kpc, then the observed interval is $V/\sigma_{[CII]} = 1-8$, where both hot and cold disks are present. Parlanti et al. [49] believe that their results diverged from the results of [40–43] due to the difference in typical star formation rates — the main sequence vs the burst regime. However, generally speaking, such an explanation is counter-intuitive: it turns out that, although the feedback effect from young stars accelerates turbulence, galaxies with violent bursts of star formation still show a lower dispersion of gas clouds in disks than that in galaxies with moderate star formation.

5. Supermassive black holes and their young host galaxies: which came first?

In the nearby Universe, almost every galaxy has a supermassive black hole at its center, with a mass of about 0.2% to 0.8% of the mass of the galaxy's spheroidal component (bulge) [50], if the spheroidal component dominates in the galaxy. In late-type galaxies, where bulges are difficult to distinguish, central black holes, however, are also found; they are especially noticeable when they manifest themselves as active nuclei emitting in the X-ray range and in broad optical emission lines. However, for these galaxies, where stellar disks dominate the structure, when trying to construct a correlation between the stellar component and the mass of the central black hole, it turns out that the mass of the black hole is much smaller than that in early-type galaxies of the same stellar mass: 0.01% to 0.025% [51, 52]. The existence of a close correlation between the stellar mass of the galaxy and the mass of the central black hole seems to indicate that both the galaxy and the central black hole grow synchronously during their evolution and probably as a result of the same events in the life of the galaxy. Cosmologists initially (in the 2000s) approved the concept of such synchronicity. Various hierarchical scenarios of evolution were discussed. For example, the merging of galaxies of comparable mass leads to a simultaneous doubling of both the mass of the stellar spheroid and the mass of the central black hole. And if we consider the evolution of a disk galaxy as completely determined by constant external accretion — of both gas and dark matter — then we can explain an increase in the stellar mass of the galaxy by star formation in the outer regions, fueled by gas accretion, with an increase in the mass of the central black hole through gas accretion onto the center of the galaxy. Such scenarios fit perfectly into the LCDM model of the evolution of the Universe in the course of hierarchical clustering of matter. But, as usual, the improvement in and achievements of observations significantly spoil the beauty of simple and harmonious models.

In recent years — literally in the last 5–7 years — astronomer-observers have managed to discover and study very distant quasars. The current champion in distance is a quasar at a redshift of $z = 8.68$ [53]. And, in total, over 200 quasars have been discovered at redshifts greater than $z = 6$ [54]. This is already enough to probe statistics! It turns out that even such early quasars, living in a Universe whose age was less than 1 billion years, had a central black hole mass of about 1 billion solar masses. Truly, the most distant quasar at $z = 8.68$ has a modest central black hole mass of $10^7 M_\odot$

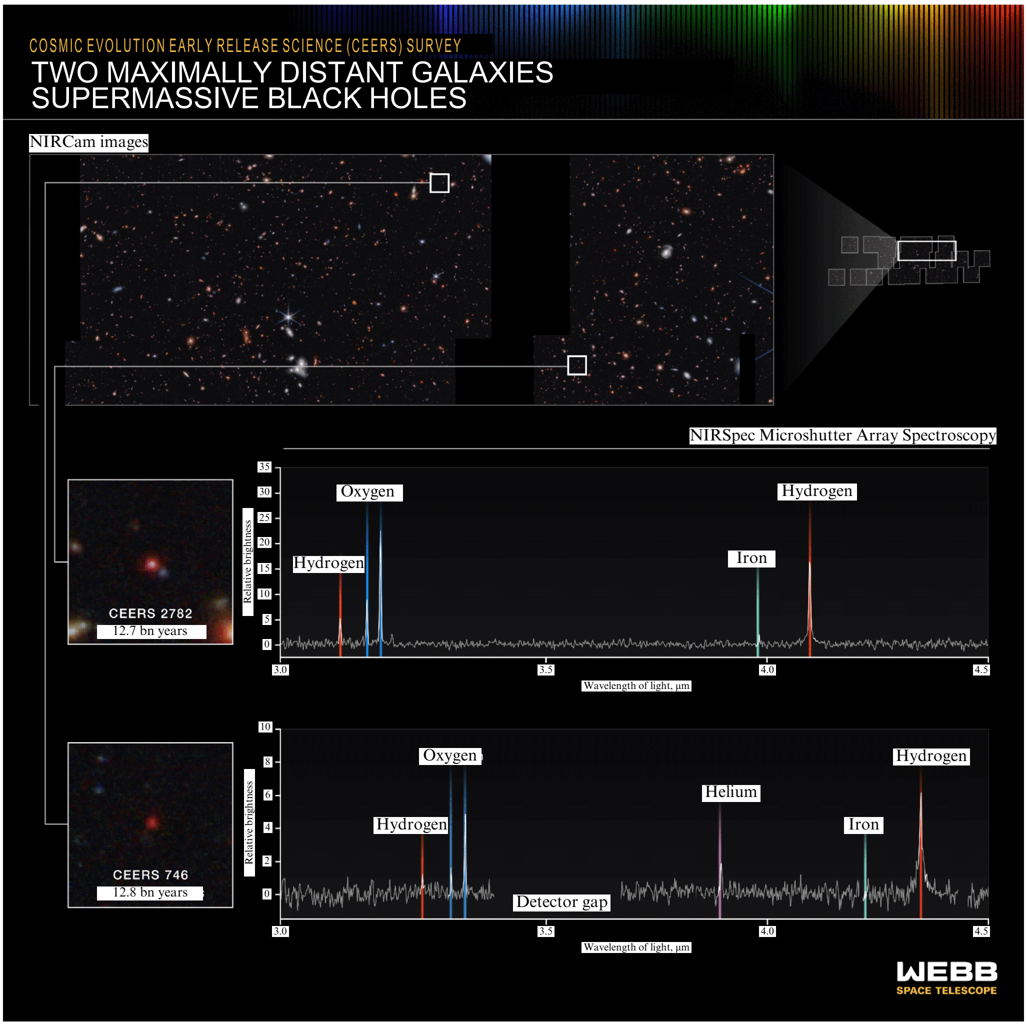

Figure 4. Examples of active galactic nuclei spectra obtained by JWST NIRspec spectrograph. Credit: NASA, ESA, CSA, Leah Hustak (STScI), S. Finkelstein (UT Austin), R. Larson (UT Austin), and P. Arrabal Haro (NSF's NOIRLab).

[53]. But already at $z = 7.642$ there is a quasar with a central black hole mass of 1.6 billion solar masses [55]. Of course, when we observe such distant objects, we first of all select the brightest of them ('selection effect'). Nevertheless, we always see the upper limit of masses! In the nearby Universe, it is several billion solar masses (for example, the central black hole in M87 has a mass of 6.5 billion solar masses [56]), but, in the distant Universe, the upper mass limit is also almost the same [57]!

The observational achievements of the JWST include not only distant quasars — active nuclei are found in even more distant *galaxies*, although with a lower luminosity than that of quasars, but also with very massive central black holes. They are detected and the masses of the central black holes are measured by the broad emission lines of the Balmer series, shifted by the cosmological redshift to the JWST/NIRspec range (Fig. 4). The maximum stellar mass of a galaxy at large redshifts seems to be less than that in the local Universe, in agreement with the hierarchical principle. Accordingly, the ratio of the mass of the central black hole to the stellar mass of the galaxy in the first billion years of the Universe's expansion *is observed* to be significantly greater than it is now: in the active nucleus of the galaxy from the JWST/UNCOVER survey, a black hole with a mass of $\log M_{\rm BH} = 8.17$ was discovered with a dwarf stellar mass of the galaxy $\log M_* < 8.7$ [58], and the champion in the ratio of the mass of the central black hole to the stellar mass of the galaxy — almost unity! — is the JADES GN 1146115 object, with $z = 6.67$ and $\log M_{\rm BH} = 8.61$ [59]. On average, for a dozen galaxies with active nuclei in the range $z = 4-7$, the ratio of the mass of the central black hole to the stellar mass of the galaxy fluctuates from 1% to 10% [60], i.e., it is tens of times higher than the corresponding ratio, also measured by broad Balmer lines, for galaxies with active nuclei at $z = 0$ [52]. On the other hand, the ratio of the *dynamic* mass of the entire galaxy to the mass of the central black hole (measured with the ALMA interferometer by the rotation speed of the galaxy's gas disk) for 27 distant quasars is also several times lower than that in the local Universe — from 20 to 200 instead of 500 [61]! Thus, observations indicate that the central black hole grows first, and only then does a galaxy appear around it. And the previously proposed mechanisms for the growth of the central black hole do not work for such a rapid accumulation of mass — it needs to accumulate a billion solar masses in half a billion years.

As soon as this circumstance was discovered, theorists actively began to study it, and almost immediately several scenarios were proposed for how to quickly — in half a billion years — grow a single black hole with a mass of a billion solar masses [62]. First of all, for this to happen, at a redshift of at least 20–40, there must already be 'seeds,' i.e., black holes of smaller masses. But how small? Scenarios are currently suggested for three classes of 'seeds': light, medium, and massive [63–65].

Light seeds, each hundreds to a thousand solar masses, may simply be post-evolutionary remnants of massive Population III stars — hypothetical first-generation stars formed from pristine zero-metallicity gas. Medium seeds can form as a result of the dynamic evolution of dense star clusters from these massive first-generation stars. With super-dense 'packing' of stars in the centers of dwarf protogalaxies, they will often collide and eventually simply merge into one black hole, almost without having time to exit the protostellar phase (loose protostars lose energy more easily during collisions, unlike already formed stars, and therefore are able to merge effectively). But theorists have scenarios for how seeds of black holes with a mass of up to a million solar masses each can be formed. Direct collapse of a gas cloud into a central black hole, bypassing the star formation phase, is considered. If gas flows rush to the center of a dark halo and if the gas is warm and cannot cool down, it will not fragment on the way into clumps of stellar masses and form a galaxy, but will gather directly in the center into a black hole with a mass of $10^5 - 10^6$ solar masses. The main point is not to let the gas cool down so that it cannot fragment into objects of smaller mass on the way to the center, while still in the disk. To this end, it is proposed to heat the gas with UV radiation from the outside. For example, such a heater can be a neighboring dwarf galaxy, which was still able to form stars (i.e., it is recognized that not every galaxy, especially a dwarf, should have a black hole in the center; dark halos with a central black hole and dark halos with a stellar population are separated). Interestingly, in order to successfully form a massive seed for a future supermassive black hole, star formation in the disk of the protogalaxy must be suppressed, i.e., the formation of the galaxy itself as a stellar system must be delayed. There are also more unusual scenarios that explain the existence of 'seeds;' such scenarios include the existence of primordial black holes inherited by the Universe from the Big Bang. Their upper mass limit, 10^6 solar masses, is suitable for seeds of future supermassive central black holes of galaxies [66].

If there is a massive gas accretion disk around this 'seed' black hole and if it is fed by gas from outside the galaxy, then efficient disk accretion with the Eddington limit or sometimes even with a super-Eddington rate makes it possible to grow a supermassive black hole of a billion solar masses in 700 million years. Episodes of super-Eddington accretion in the early evolution of supermassive black holes are now becoming increasingly popular with theorists [67, 68]. First, a high rate of super-Eddington accretion will help grow a supermassive black hole of a billion solar masses in the proposed hundreds of millions of years even from a 'light seed.' Secondly, if the seeds are heavy, the assumption of episodes of super-Eddington accretion provides an opportunity to 'catch' such a growing supermassive black hole *between* accretion episodes, and therefore to find it in a 'sleeping' state. This state, for example, perfectly explains the observation result [59], where the mass of the black hole is only 2.5 times smaller than that of the stellar component of the galaxy (i.e., the ratio of the masses of the black hole and the galaxy is 1000 times greater than that in the local Universe). At the same time, it demonstrates an extremely low current accretion rate, only 2% of the Eddington limit. Finally, the first observed examples of super-Eddington accretion onto a central black hole in an early galaxy have already appeared: this is the GN-z11 galaxy, which has long been the champion in distance, at a redshift of $z = 10.6$, for which the JWST measured a high-quality spectrum. Having analyzed the broad emission lines in this spectrum, Maiolino et al. [69] announced in *Nature* that they see accretion onto a central black hole with a mass of $\log M_{\rm BH} = 6.2$ at a rate equal to five Eddington rates.

Thus, from the point of view of theory, there is now no problem obtaining lonely supermassive black holes in the center of dwarf dark halos in the first billion years of the Universe's expansion. And then stars will begin to form in the gas disk surrounding the supermassive black hole and a young galaxy will grow around it. In fact, in order for us to see the manifestation of this supermassive black hole as an active galactic core, we need a different accretion regime, a different gas supply, which, perhaps, makes sense to associate with a star formation burst in a large gas disk of a galaxy.

The work was performed under the State assignment of Lomonosov Moscow State University.

References

1. Hubble E *The Realm of the Nebulae* (New Haven, CT: Yale Univ. Press, 1936)
2. Papovich C et al. *Astrophys. J.* **631** 101 (2005)
3. Gardner J P et al. *Space Sci. Rev.* **123** 485 (2006)
4. Bromm V, Larson R B *Annu. Rev. Astron. Astrophys.* **42** 79 (2004)
5. Bromm V, Yoshida N *Annu. Rev. Astron. Astrophys.* **49** 373 (2011)
6. Barkana R, Loeb A *Phys. Rep.* **349** 125 (2001)
7. Steidel C C et al. *Astrophys. J.* **462** L17 (1996)
8. Oesch P A et al. *Astrophys. J.* **819** 129 (2016)
9. Finkelstein S L et al. *Astrophys. J. Lett.* **946** L13 (2023)
10. Finkelstein S L et al. *Astrophys. J. Lett.* **969** L2 (2024)
11. Arrabal Haro P et al. *Nature* **622** 707 (2023)
12. Arrabal Haro P et al. *Astrophys. J. Lett.* **951** L22 (2023)
13. Donnan C T et al. *Mon. Not. R. Astron. Soc.* **518** 6011 (2023)
14. Zavala J A et al. *Astrophys. J. Lett.* **943** L9 (2023)
15. Bouwens R et al. *Astrophys. J.* **902** 112 (2020)
16. Hainline K N et al. *Astrophys. J.* **964** 71 (2024)
17. Robertson B et al. *Astrophys. J.* **970** 31 (2024)
18. Casey C M et al. *Astrophys. J.* **965** 98 (2024)
19. Lilly S et al. *Astrophys. J.* **772** 119 (2013)
20. Brinchmann J et al. *Mon. Not. R. Astron. Soc.* **351** 1151 (2004)
21. Noeske K G et al. *Astrophys. J.* **660** L43 (2007)
22. Rodighiero G et al. *Astrophys. J. Lett.* **739** L40 (2011)
23. Donnan C T et al. *Mon. Not. R. Astron. Soc.* **533** 3222 (2024)
24. McLeod D J et al. *Mon. Not. R. Astron. Soc.* **527** 5004 (2024)
25. Ferrara A, Pallottini A, Dayal P *Mon. Not. R. Astron. Soc.* **522** 3986 (2023)
26. Cueto E R et al. *Astron. Astrophys.* **686** A138 (2024)
27. Harikane Y et al. *Astrophys. J. Suppl.* **265** 5 (2023)
28. Caputi K I et al. *Astrophys. J.* **849** 45 (2017)
29. Rinaldi P et al. *Astrophys. J.* **930** 128 (2022)
30. Ciesla L et al. *Astron. Astrophys.* **686** A128 (2024)
31. Blain A M et al. *Phys. Rep.* **369** 111 (2002)
32. Casey C M, Narayanan D, Cooray A *Phys. Rep.* **541** 45 (2014)
33. Michalowski M J et al. *Mon. Not. R. Astron. Soc.* **469** 492 (2017)
34. Smail I, Ivison R J, Blain A W *Astrophys. J. Lett.* **490** L5 (1997)
35. Chapman S C et al. *Astrophys. J.* **622** 772 (2005)
36. Geach J E et al. *Mon. Not. R. Astron. Soc.* **465** 1789 (2017)
37. Wardlow J L et al. *Mon. Not. R. Astron. Soc.* **415** 1479 (2011)
38. Vieira J D et al. *Astrophys. J.* **719** 763 (2010)
39. Hodge J A et al. *Astrophys. J.* **876** 130 (2019)
40. Rizzo F et al. *Nature* **584** 201 (2020)
41. Rizzo F et al. *Mon. Not. R. Astron. Soc.* **507** 3952 (2021)
42. Lelli F et al. *Science* **371** 713 (2021)
43. Roman-Oliveira F, Fraternali F, Rizzo F *Mon. Not. R. Astron. Soc.* **521** 1045 (2023)
44. Wisnioski E et al. *Astrophys. J.* **799** 209 (2015)
45. Pillepich A et al. *Mon. Not. R. Astron. Soc.* **490** 3196 (2019)

46. Roman-Oliveira F, Rizzo F, Fraternali F *Astron. Astrophys.* **687** A35 (2024)
47. Kretschmer M, Dekel A, Teyssier R *Mon. Not. R. Astron. Soc.* **510** 3266 (2022)
48. Kohandel M et al. *Astron. Astrophys.* **685** A72 (2024)
49. Parlanti E et al. *Astron. Astrophys.* **673** A153 (2023)
50. Kormendy J, Ho L C *Annu. Rev. Astron. Astrophys.* **51** 511 (2013)
51. Graham A W, Sahu N *Mon. Not. R. Astron. Soc.* **518** 2177 (2023)
52. Reines A E, Volonteri M *Astrophys. J.* **813** 82 (2015)
53. Larson R L et al. *Astrophys. J. Lett.* **953** L29 (2023)
54. Fan X, Bañados E, Simcoe R A *Annu. Rev. Astron. Astrophys.* **61** 373 (2023)
55. Wang F et al. *Astrophys. J. Lett.* **907** L1 (2021)
56. The Event Horizon Telescope Collab., Akiyama K et al. *Astrophys. J. Lett.* **875** L6 (2019)
57. Stone M A et al. *Astrophys. J.* **964** 90 (2024)
58. Kokorev V et al. *Astrophys. J. Lett.* **957** L7 (2023)
59. Juodžbalis I et al. *Nature* **636** 594 (2024); arXiv:2403.03872
60. Pacucci F et al. *Astrophys. J. Lett.* **957** L3 (2023)
61. Decarli R et al. *Astrophys. J.* **854** 97 (2018)
62. Haemmerlé L et al. *Space Sci. Rev.* **216** 48 (2020)
63. Inayoshi K, Visbal E, Haiman Z *Annu. Rev. Astron. Astrophys.* **58** 27 (2020)
64. Valiante R et al. *Mon. Not. R. Astron. Soc.* **457** 3356 (2016)
65. Sassano F et al. *Mon. Not. R. Astron. Soc.* **506** 613 (2021)
66. Dolgov A D *Moscow Univ. Phys. Bull.* **79** (Suppl 1) 324 (2024) https://doi.org/10.3103/S0027134924701017; arXiv:2401.06882
67. Volonteri M, Silk J, Dubus G *Astrophys. J.* **804** 148 (2015)
68. Lupi A et al. *Astron. Astrophys.* **686** A256 (2024)
69. Maiolino R et al. *Nature* **627** 59 (2024)