Chapter 2
Fundamental Cosmological Observations and Data Interpretation

Contributions by Matthias Bartelmann, Charles L. Bennett, Carlo Burigana, Cesare Chiosi, Mauro D'Onofrio, Alan Dressler, Isabella Gioia, Günther Hasinger, Juan Francisco Macias-Perez, Piero Madau, Paola Marziani, John Mather, Francesca Matteucci, Keith Olive, John Peacock, Wolfgang Reich, Pierre-Marie Robitaille, Michael Rowan-Robinson, Gary Steigman, Matthias Steinmetz, Jack W. Sulentic, Massimo Turatto, and Simon D.M. White

2.1 Outline of the Chapter

The last decades have seen an exponential increase of experiments and observations aimed at establishing the structure, the evolution, and the constituents of the Universe, covering essentially the whole electromagnetic spectrum.

The interviews collected in this chapter describe the main evidence at the base of the currently accepted cosmological scenario and discuss the interpretation of experimental data that support it. Some alternative (heretical?) ideas on the interpretation of redshifts and Cosmic Microwave Background (CMB) radiation are also included.

We start presenting the most important change of cosmological paradigm that has occurred in the last two decades: the transition from the Cold Dark Matter (CDM) to the ΛCDM models, described in Sect. 2.2 by John Peacock and in Sect. 2.3 by Massimo Turatto from two different points of view. Far from being a simple re-parameterization of the various kinds of energy contents of the Universe, the new scenario has dramatically impacted on our vision of the fate of the Universe and of its "recent" dynamical evolution and structure formation. It raises the problem of the cosmological constant with the related question of the dark energy (DE) content, and the alternatives to these possibilities that are linked to fundamental physics, from early Universe processes to gravitational theories.

The final step towards the new cosmological paradigm largely relies on the study of type Ia Supernovae (SNe), the pros and cons of which, in cosmological and astrophysical context, are described by Massimo Turatto in Sect. 2.3 and by Francesca Matteucci in Sect. 2.4. The reliability of these objects as distance indicators is clearly the center of our interest in their interviews. How robust is the present indication of an accelerating Universe coming from SNe? How far can we trust in such indicators at redshifts so far from that of nearby type Ia SNe?

Such discussions inevitably bring us to seek new and more powerful distance indicators. Do we have any? In Sect. 2.5, we ask Paola Marziani why the luminous quasars cannot be useful distance indicators for tracing the structure of the Universe, despite the fact that they can be observed up to redshifts ($z \sim 6$) larger than that of

M. D'Onofrio and C. Burigana (eds.), *Questions of Modern Cosmology: Galileo's Legacy*, DOI 10.1007/978-3-642-00792-7_2,

type Ia SNe. Of course, speaking of quasars, it was natural for us to explore the well known question of the anomalous redshifts observed in the past by Halton Arp for some of these objects. We discuss such problems with Jack Sulentic in Sect. 2.6.

Discussion on the empirical cornerstones of the Concordance Model develops along the following lines.

We begin with the cosmological nucleosynthesis, presented in Sect. 2.7*, in which we ask the points of view of Keith Olive and Gary Steigman on the standard Big Bang Nucleosynthesis (BBN) theory and on the empirical tests proving this scenario. The interview with Keith is more focused on nuclear reactions and empirical tests of the BBN, while Gary will more closely follow the link between nucleosynthesis process and the expansion of the Universe, addressing the complementary information coming from BBN, CMB, and Large Scale Structure (LSS) and the possible alternatives to the standard theory.*

The second big cornerstone, the CMB, has been reviewed by the Nobel Laureate John Mather, by Charles Bennett, and by Juan Francisco Macias-Perez. John and Charles delineated the most important aspects of the Cosmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (WMAP) all-sky space missions. These projects have strongly impacted on our ideas on the evolution of the Universe. Their interviews can be found in Sects. 2.8 *and* 2.8.2*. John will review the cosmological context before and after COBE, the great success of the mission, and the experiments on-board COBE. Charles, after a brief discussion of the cosmological questions left open by COBE, will introduce us to the aims and strategy of WMAP, presenting its main scientific achievements. Juan reviews for us the fundamental results from balloon-borne experiments realized in the years between COBE and WMAP (see Sect.* 2.8.3*).*

One key aspect for the interpretation of microwave data is to understand in which way the foreground signals coming from the Milky Way and extragalactic sources are separated from the truly cosmological CMB signal. We start by describing the properties of the far infrared foreground and of dust emission, mainly observed at high resolution through balloon experiments. Juan will address these questions in Sect. 2.8.4*, where he gives a concise description of the algorithms of component separation. He also presents an overview of the properties of the Inter Stellar Medium (ISM) in Sect.* 2.8.5*. Later on, Wolfgang Reich will remind us the characteristics of the radio foreground in Sect.* 2.8.6*, presenting the most recent all-sky radio maps painstakingly obtained both in total intensity and polarization and discusses the role of magnetic fields in our Galaxy.*

Despite the big success of CMB experiments, the interpretation of the CMB data is not uniformly accepted yet. For this reason, in the spirit of this book, we decided to give space even to the more radical opposition. Pierre-Marie Robitaille will give his point of view on CMB in Sect. 2.8.7*. He will discuss, in particular, the interpretation of the Planck and Kirchhoff data and the origin of the CMB monopole signal.*

From the truly diffuse CMB background, we then move ourselves towards the complex problem of X-ray background, the astrophysical sources responsible for it, and the contribution of X-ray astronomy to the development of the current *cosmological scenario. These themes are addressed here by Günther Hasinger in Sect. 2.9, where he points also out the effect of AGN evolution and massive Black Hole (BH) feedback.*

Having brought the discussion on primordial BHs, we continue our interviews with the difficult problem of how such primeval objects and structures emerged from the dark era, when the Universe entered in its nonlinear phase. Piero Madau will review for us the story of the dark ages and the appearance of first stars in Sect. 2.10. He also addresses the question of the feedback from BHs in galaxy formation in Sect. 2.10.3.

The sections that follow are mainly dedicated to the cosmological information coming from the "nearby" Universe, dominated by the presence of self-gravitating structures of dark and visible matter. The largest of such structures are galaxy clusters. In Sect. 2.11, Alan Dressler will describe the general properties of clusters and their importance for the current cosmology. In particular, he revisits the information that can be extracted from the Morphology–Density relation, which he discovered some years ago, the problem of the bias in the measurements of mass, and the role of the scaling relations, such as the Fundamental Plane and the Tully–Fisher relations, in the cosmological context.

Galaxy clusters are indeed important tracers of the mass distribution in the Universe. In particular, the relationships between X-ray luminosity and temperature, and between temperature and mass, due to the hot gas in the Intra Cluster Medium (ICM), can give us a measure of the clustering of matter around such structures. Isabella Gioia, in Sect. 2.12, will provide a multifrequency view of the properties of galaxy clusters and their use for extracting information on the cosmological parameters.

Of course, speaking of clusters, it was inevitable to address one of the most crucial problems of present cosmology: the dark matter (DM). Does it exist and what is it? How is it distributed around galaxies and clusters? Simon White and Matthias Steinmetz have reviewed the problem in Sect. 2.13 and in related subsections, discussing also the contribution to these studies coming from cosmological simulations. While Simon presents the empirical evidence that call for the existence of DM, Matthias focuses his interview on the pros and cons of the standard CDM scenario.

The lensing phenomenon is closely linked to the problem of DM. Although predicted by Einstein and observed for the first time by Eddington during the solar eclipse of 1919, the cosmological exploitation of lensing started only during the last 10 years. What are its properties and what is its role in determining the distribution of dark mass around structures? Matthias Bartelmann in Sect. 2.14 will review the physical concepts at the basis of our expectations for cosmology coming from weak and strong lensing.

The interviews so far included mainly concern some fundamental categories of physical observables providing cosmological information. These are mass and type of matter, both visible and dark, in the form of astronomical objects or diffuse components, their geometrical (angular and redshift–distance) distributions, and their spectral energy properties. The category considered in the next section is instead directly related to the exploitation of cosmic time evolution of stars, that, *as well known, are the best known astrophysical clocks. Cesare Chiosi will illustrate the constraints coming from stellar evolution in determining the time scale of the Universe. In Sect. 2.15, he revisits the problem of the stellar ages from the perspective of star clusters and that of the integral color of galaxies.*

In the final interview of this chapter, the cosmological question of the value of the Hubble constant H_0 from various observables is addressed. Although the discovery of the expansion of the Universe by Hubble was one of the starting points of modern cosmology, we decided to move this discussion to the end of the chapter, because H_0 is now based on various kinds of observables, from astronomical objects to diffuse cosmic background, and the analysis of discrete objects assumes the formation of astronomical structures described in the previous sections. Michael Rowan-Robinson reviews the present situation of H_0 in Sect. 2.16.

Let us start with the interview of John Peacock, who will now clarify the epoch of transition from the CDM to the LCDM scenarios.

2.2 From CDM to ΛCDM Paradigm

Dear John *(Peacock)***, recent surveys of galaxies and clusters have significantly contributed to our current knowledge of the structure of the Universe and of its evolution. Would you like to summarize here the most important results coming from such studies in the context of observational cosmology?**

Structure in the visible Universe has driven cosmological enquiries from very early on. Once the basic nature of galaxies as stellar systems in motion had been established by the work of Slipher and Hubble, astronomers were inevitably led to ask how galaxies had originated. Even before the detection of the CMB, the theory of gravitational collapse from small initial fluctuations had been worked on extensively, and many of the key elements of modern understanding were in place by the early 1970s, including the idea of characteristic patterns in cosmic structure being imprinted by the transition between an early radiation dominated era and matter domination. These ideas were developed further during the 1970s, at the same time as the first comprehensive attempts at quantifying the inhomogeneities in the galaxy distribution: the correlation-function programme whose results were summarized in the hugely influential book by Peebles [392].

Peebles's book marked a true turning point in the subject, as it was almost immediately followed by two key theoretical advances. The most fundamental was the development of inflation, which proposed the heroic vision of a quantum origin for cosmic structure. Furthermore, the simplest prediction of inflation was seen to be fluctuations that were adiabatic (equal fractional perturbations to the matter density and the photon number density), and nearly scale-invariant in character (metric perturbations that were fractal-like, so that deviations from flatness were of equal magnitude in each logarithmic range of spatial wavelength). The other main step was the idea that the main matter component might be a collisionless relic particle. Initially, the main candidate of interest was the massive neutrino, but it was rapidly appreciated that this idea was a dead end: randomly directed streaming of neutrinos while they are relativistic would erase all galaxy-scale structures [55]. This led to the simpler idea of CDM , in which the particle is massive and any small-scale damping is unobservably small [393]. The consequences of this for the CMB were quickly worked out [54, 577]. Thus, in an astonishing burst of activity, all the elements were put in place of the theoretical picture that still applies today, a quarter of a Century later.

The CDM model made two specific predictions: (1) that there would be a break in the power spectrum of matter fluctuations at a length around $c \times t$ at the time of matter–radiation equality; (2) that the characteristic angular scale of CMB fluctuations would depend on space-time curvature, being around 1° for a flat Universe, but moving to smaller angles for an open Universe. As we know, it has been possible to use these predictions to measure the character of the Universe with astounding precision. The history of this is rather interesting, and shows that galaxy surveys really led the way in the 1990s, although by now the CMB data give us much the most important and accurate constraints. In 1990, for example, only upper limits on CMB fluctuations existed – but these were still important. The small-scale limits on 10-arcmin scales from the Caltech group were rather stringent, and it was already clear that CDM models that were heavily open could be excluded. If flat models were preferred, it came down to a choice between the $\Omega_m = 1$ Einstein–de Sitter Universe, or one that was vacuum dominated, satisfying $\Omega_m + \Omega_v = 1$. Throughout the 1980s, most cosmologists would have plumped for the former alternative – based largely on worries about the fine-tuning involved in a small cosmological constant that becomes important only around the present. But by 1990, strong evidence had accumulated in favor of this alternative. The CDM spectrum contains a break at the horizon scale of matter–radiation equality, and observations of this break scale allow the combination $\Omega_m h$ to be pinned down (where $h \equiv H_0/(100\,\mathrm{km\,s^{-1}\,Mpc^{-1}})$ is the usual dimensionless Hubble parameter at the present time, i.e., the reduced Hubble constant). A lower matter density implies relatively larger fluctuation on large scales, and clear evidence for these was seen in the projected clustering properties of the Automatic Plate Measuring (APM) galaxy survey, which was produced from scans of UK Schmidt Telescope photographic plates. The preferred value of $\Omega_m h \simeq 0.2$ argued for a low matter density, even though the Hubble constant was not so well known then as it is now. The only way of reconciling the small-scale CMB constraint with the requirement of a low matter density from galaxy clustering was to assume that the Universe was dominated by a cosmological constant – so that the sum of this vacuum energy and the matter content yielded a flat Universe. This argument was made with admirable clarity in a Nature paper by Efstathiou et al. [147].

Although the logic was impressive at the time, there was still considerable resistance to the conclusion. I certainly remember being deeply unhappy with the idea of a fine-tuned vacuum density, and spent the early 1990s looking for ways out: basically trying to see if nonlinear evolution and scale-dependent bias could help. It was abundantly clear that $\Omega_m = 1$ was dead, but I was philosophically more attracted to the idea that the Universe might be open (with, say, $\Omega_m \simeq 0.4$ and a low Hubble parameter) than to accept the reality of vacuum domination. These slightly open models looked a progressively less good match to the data as time went by, although their rejection was delayed by the first Perlmutter et al. paper [402] on the supernovae Hubble diagram, which conclusively rejected what we would now regard as the correct vacuum-dominated model. It was only in 1997, just at the time I was finishing my textbook on cosmology, that it became clear that the SN story was changing and also falling in line with the picture that LSS+CMB had been painting since 1990. At that stage, I abandoned any further resistance to the idea of nonzero Λ – but it is interesting to see how the suggestion has arisen that the "discovery of DE" came out of a clear sky with the 1998 SN papers. I certainly feel fortunate in the alignment of the timing of all this with finishing my book: if I'd been more efficient and got it done when the publishers first wanted, it would have been horribly out of date within a couple of years. As it was, I was able to in effect write a first-hand witness account of the birth pangs of the present standard model. Although we know many numbers much more accurately than we did 10 years ago, it is astonishing how little has changed over the past decade in terms of our basic set of ideas. I cannot decide if this is a cause for celebration or depression; certainly, the nature of the subject has altered out of all recognition from the glory days of the 1980s, and the kind of creative cosmological speculation that was common then is less easy to carry off today.

Thank you John. Indeed the last 10 years have seen enormous progress in both observations and theories, but not yet decisive to establish the cosmological scenario. Could we interpret this as a symptom of a crisis? We will address this question later in this book. For the moment we are interested in highlighting the transition from the CDM to the ΛCDM from the point of view of type Ia SNe at high redshifts.

2.3 Type Ia SNe as Probe of the Paradigm Shift

Dear Massimo *(Turatto)*, since the discovery of the cosmic acceleration of the Universe prompted by observations of high-z SNe-Ia, the ΛCDM scenario has been confirmed by CMB experiments, in particular by WMAP. Could you please briefly review the role of SNe as cosmological tracers?

Supernovae are celestial objects that, even if for short time, shine as bright as their entire host galaxy ($M_B \sim -19$), making them detectable up to cosmological distances with large telescopes and modern detectors. For this reason, SNe are unmatched probes of the different evolutionary conditions of the Universe. In particular, the subclass of SNe called of type Ia, which can be recognized for the characteristic light curve and spectral features (see [561, 562] for a modern SNe taxonomy), has the specific property of having a relatively small dispersion of luminosity at maximum light, making them unique distance indicators. As explained in detail by Francesca Matteucci in the Sect. 2.4, type Ia SNe are the outcome of thermonuclear explosion of White Dwarfs reaching the Chandrasekhar limit by accretion of material from a companion. The fact that in first approximation they are similar one to each other is therefore not surprising.

In the 1990s, with the improved photometric calibration it became evident that type Ia SNe show a significant diversity at optical wavelengths (hence are not strictly standard candles), but it was also found that simple relations exist between the shapes of the light curves and the absolute magnitudes at maximum [402, 405, 449]. Therefore, in analogy to what happens with Cepheids, type Ia SNe can be recovered as distance indicators (*standardizable* candles). This method has proved to be very effective, but has a major caveat, the lack of a satisfactory theoretical interpretation. The fact that type Ia SNe might indeed be much better standard candles in the near-IR [296], where also the reddening is much less than a problem, is of little help in the current context because of the drift in the luminosity peak with redshift to even longer wavelengths, with the known observational complications and of the lower luminosity of type Ia SNe in the IR. If this finding will be confirmed, the James Webb Space Telescope (JWST) might exploit the near-IR properties of type Ia SNe.

The Hubble diagram built with the shape-corrected luminosities of nearby type Ia SNe has a dispersion less than 0.2 mag, that is, 10% in distance (see [316] and references therein). A direct output of the Hubble diagram is the determination of the Hubble constant once the absolute magnitude of type Ia SNe is known independently. Making use of the calibration of nearby type Ia SNe by Cepheids, a recent determination provides $H_0 = 73 \pm 4(\text{stat}) \pm 5(\text{sys})\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$ [451] (see also Sect. 2.16 by Michael Rowan-Robinson). The Hubble diagram of type Ia SNe shows evidence for a *local bubble* with local ($v < 7{,}000\,\text{km}\,\text{s}^{-1}$) expansion velocities larger than the cosmic average [258, 600].

But type Ia SNe allow to go further and to explore the cosmic expansion rate up to look-back times of about two-third the age of the Universe ($z \sim 1.5$). In 1998, two competing SN teams [403, 450] announced the independent discovery that type Ia SNe at $z \sim 0.5$ are fainter than predicted in an empty Universe and, therefore, the expansion of the Universe is accelerated, possibly due to the presence of a new (dark) energy component, which opposes the gravitational pull or a modification of gravitation theory. Combining SN measurements with those obtained from the LSS and CMB, a so-called Concordance Model has emerged, in which the Universe is flat and filled with about 4% baryons, 20% DM, and 76% DE (see [183] for a recent review). These claims are now supported by better statistics provided by the Supernova Legacy Survey (SNLS) and the Equation of State Supernovae trace Cosmic Expansion (ESSENCE) collaborations [16, 591], which have measured light curves for several hundred type Ia SNe in the $0.3 < z < 0.9$ range. In addition, a SN search carried out with the Hubble Space Telescope (HST) confirmed that the higher-redshift ($z > 1$) Universe is decelerating (see Fig. 2.1), and was able to sample the transition from a deceleration in the past to the current acceleration [452]. The SN data together with those from Baryonic Acoustic Oscillations (BAO) and CMB have been used to constrain the equation of state of the DE ($w = P/(\rho c^2)$) (see Figs. 14 and 15 of [293]). Presently the available data are consistent with $w = -1$ with no time variation, though other models cannot be excluded.

The cosmological results just mentioned are getting more robust as larger data set are collected but the number of issues remain, both technical and astrophysical, which limit the accuracy of the results (see [316] for an extensive overview).

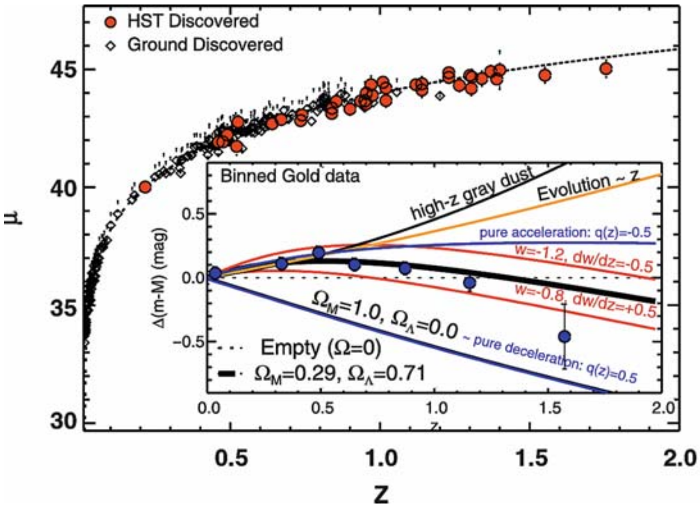

Fig. 2.1 SN Ia Hubble diagram. SNe Ia from ground-based discoveries in the gold sample are shown as *diamonds*, HST-discovered SNe Ia are shown as *filled symbols*. Overplotted is the best fit for a flat cosmology: $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$. *Inset*: Residual Hubble diagram and models after subtracting empty universe model. The gold sample is binned in equal spans of $n\Delta z = 6$, where n is the number of SNe in a bin and Δz is the redshift range of the bin. From [452]

Detailed treatments of the systematics have shown that more accurate photometric calibration especially when dealing with mosaic detectors or with data coming from different telescopes are needed. Moreover, it is important to monitor the stability and the uniformity of photometric systems as well as the atmospheric temporal and spatial variations. Even small uncertainties at the edges of the photometric band passes can affect significantly the K-corrections of the feature-rich SN spectra or the computation of light curve models based on spectral templates. Different light curve fitting methods used so far seem to give consistent results but still can be the source of systematic biases. The problem of removal of host galaxy contamination has not been solved, as well as the contamination of the SN samples by possible type Ia SNe impostors, like bright SNIb/c [32, 549]. Several are the astrophysical issues deserving further investigation, among which are the gravitational lensing and the existence of peculiar velocity fields. The presence of noncanonical ($R_V \sim 2$) reddening laws in the host galaxies [154, 297] might be a serious concern especially in the early Universe. Last but not least, there is the critical assumption that lies at the basis of the observational cosmology with SNe, that is, the observational properties of SNe do not change with redshift. The currently available data show that the spectroscopic and light curve behaviors of type Ia SNe at low and high redshifts are rather similar [16, 50, 452, 591] and the cosmological parameters derived in different host environments [541] or SN subsamples [293] are consistent, but still this is an issue that must be continuously investigated. Not only that, before drawing any firm conclusion on the nature of our Universe, we need a full theoretical comprehension of these fascinating, but largely unknown objects, and to answer basic but still unsolved questions. What is the true nature of the progenitors? What is(are) the explosion mechanism(s)? What drives the observed diversity? What is(are) the parent Population(s)?

Do there exist other types of SNe that can be of interest for cosmology?

Contrary to the thermonuclear type Ia SNe, Core-collapse SNe (CCSNe), descending from stars more massive than 8–10 $M_\odot$, are far from being standard candles and show huge variations in the peak luminosity and in the shape of the light curve, due to different configurations of the exploding stars at the moment of the explosion and to different energetics of the explosion itself. They range from the faint ($M_V \sim -14$) SNe [389] to the bright ($M_R \sim -22$) and hyper-energetic SNe like 2005ap and 2006gy [5,432,519]. Nevertheless, some CCSNe can be used as distance indicators by mean of the Expanding Photosphere Method (EPM [278]), which allow to derive their distance independent of the calibration of lower rungs of the distance ladder. The uncertainty in the determination of dilution factor, which accounts for the difference of the SN spectra from that of a Black Body (BB), seems to overcome by the modern incarnation of the method (Spectral-fitting Expanding Atmosphere Model (SEAM) [25]), which exploits a detailed spectral modeling for each object. An empirical method has been recently developed [226], which makes use of the luminosity of the extended plateau characterizing the light curve of the hydrogen-dominated type IIP.

CCSNe are attractive in cosmology for other reasons. Because of the short evolutionary life-times of their progenitors, (<30Myr) the determination of the rate of their explosion provides a direct measurement of the on-going star formation rate (SFR) for an assumed initial mass function (IMF). The study of the SN rate as a function of the redshift thus provides a trace of the star formation history (SFH). The most recent determinations confirm the steep increase of CCSN rate (SFR) by a factor of 3 for a look-back time of 3 Gyr [65].

Stripped-envelope CCSNe (SNIb/c), that is, those whose progenitors have lost the H (SNIb) and He (SNIc) layers by massive stellar winds or by interaction with a nearby companion star, have been recently associated to the Gamma-ray burst (GRB) of long duration. In particular, the association seems to hold for high-energy, broad-lined events like the type Ic SN 1998bw ($E > 10^{52}$ ergs) [190, 252]. A continuous distribution of properties seems to connect these *hypernovae* to less energetic objects like 1994I, passing through intermediate objects like SNe 2002ap and 2006aj.

Finally, overall SNe of all types are the major contributors to the chemical enrichment of the Universe by returning to the ISM the heavy elements synthesized during the hydrostatic and explosive burning. The impact of various SNe types on the chemical evolution of galaxies are extensively discussed by Francesca Matteucci in the next section.

Thank you Massimo. It seems that the uncertainties on the distance of SNe is still a matter of controversy today: going below a 10% uncertainty on distance is very difficult for several reasons. Furthermore, as you say correctly, can we exclude a luminosity evolution of type Ia SNe with redshift? As we do not really know the *details of the explosion mechanisms and the nature of the progenitors, can we trust in them as good standard candles? Let us ask Francesca Matteucci about that.*

2.4 SNe Physics and the ΛCDM Scenario

Dear Francesca *(Matteucci)*, **the current lack of a full theoretical explanation of the physics of the explosion of SNe may be potentially dangerous for the cosmological Concordance Model. Can you explain why?**

The most dangerous fact for the cosmological Concordance Model is the possibility that type Ia SNe are not standard candles. This could happen if the mechanism of explosion would be different among the progenitors of such SNe, including the possibility that high redshift type Ia SNe are different from the local ones. This would therefore challenge the assumption that type Ia SNe can be considered as standard candles. What about the explosion mechanism? What do we know? It is commonly believed that type Ia SNe originate from the explosion, by carbon-deflagration, of a carbon oxygen white dwarf, which has reached its limiting mass for stability, namely the Chandrasekhar mass ($M_{\rm Ch} \sim 1.44 \rm M_{\odot}$), as illustrated in Fig. 2.2, where a possible scenario for the progenitor of type Ia SNe is presented. So, each SN Ia should be the result of the explosion of a fixed mass. This would ensure that the maximum luminosity of such SNe is always the same as it was believed until a few years ago, when type Ia SNe with different maximum luminosity were discovered.

However, Phillips [405] pointed out that there is a significant intrinsic dispersion in the absolute magnitudes at maximum light of local type Ia SNe. This result was interpreted to arise from a possible range of masses for the progenitors or from variations in the explosion mechanism. Both interpretations could cast doubt on the use of type Ia SNe as very accurate standard candles, particularly at large redshifts where Malmquist bias[1] could be an important effect. It has then been proposed that some SNe type Ia could be the result of the explosion of a sub-Chandrasekhar white dwarf ($0.6 - 1.0\ M_{\odot}$, see, e.g., [592]). Variations in the explosion mechanism could also produce a dispersion in the absolute magnitudes [275]; besides deflagration, other possible explosion mechanisms are detonation, delayed detonation, pulsating delayed detonation, and tamped detonation, although carbon-deflagration is preferred as it produces the right amount of chemical elements observed in SN spectra. However, there has been shown to exist a correlation between the maximum absolute magnitude of SNe Ia and the rate of decline of their luminosity after the maximum, which therefore allows one to calculate the maximum luminosity in any case, and therefore to retain the SNe Ia as standard candles.

[1] The Malmquist bias is a selection effect in observational astronomy. In particular, if a sample of objects (e.g., galaxies, quasars, etc.) is flux-limited, then the observer will see an increase in average luminosity with distance. This is of course because the less luminous sources at large distances will not be detected. The solution is to use a sample that is not magnitude limited such as a volume limited sample.

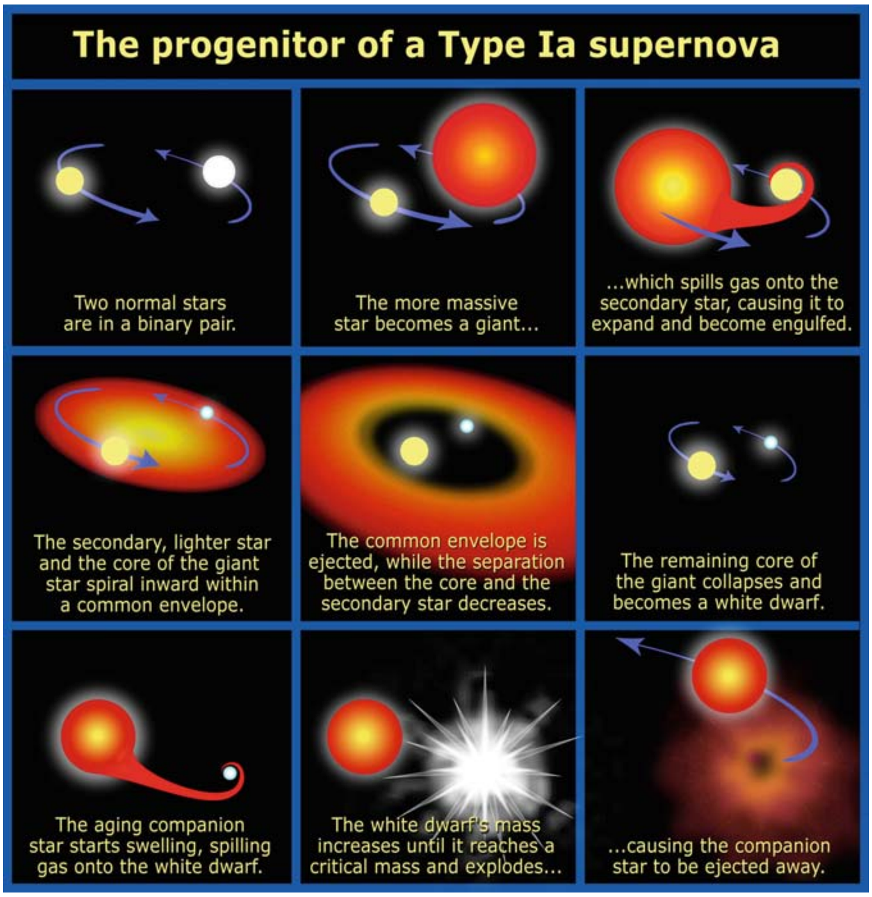

Fig. 2.2 The progenitor of a type Ia SN in the context of the single-degenerate model: here we have a binary system made of a C–O white dwarf plus a normal star. Illustration credit: NASA, ESA, and A. Feild (HST-ScI)

The other challenge to the Concordance Model of cosmology is the possibility that high redshift type Ia SNe are different from local ones, and therefore that the correlation between the maximum luminosity and the rate of decline is no more applicable. Howell et al. [246] pointed out that there are basically two groups of type Ia SNe, as suggested also by the studies of Mannucci et al. [336, 337]: one group is made of prompt SNe exploding on short timescales (less than 0.1 Gyr) and they are intrinsically more luminous and with broader light curves, the other group is made of SNe, which take several Gyrs to explode since the birth of the progenitors, have narrower light curves, and are less luminous. A possible interpretation is that the prompt SNe originate from white dwarfs with more massive progenitors than the delayed SNe. In particular, progenitors with main sequence masses between 5 and 8 $M_{\odot}$ should be related to the prompt SNe. They are brighter and produce more

^{56}Ni, whose decay is responsible for the observed light curve. As the cosmic star formation rate increases by a factor of 10 from redshift $z = 0$ to $z = 1.5$ [245], at high redshift, these SNe will be an order of magnitude more common, and even more so if it is assumed that ellipticals and spheroids formed preferentially at high redshift, as many observations seem to indicate. Therefore, it may be necessary to apply corrections for the evolution of SN Ia properties with redshift (beyond the correction for light curve shape established for local type Ia SNe), and future studies requiring increasing precision must take into account the effects of an evolving type Ia SN population.

While the elements (O, Ne, Mg, Si, S) are mainly produced in type II SNe in relatively short time-scales $(0.01 \div 0.03$ **Gyr), the Fe-peak elements are produced in SNe Ia in a longer time-scale, which has been estimated to be** $0.3 \div 1.0$ **Gyr and more through chemo-dynamical modeling of galaxies. Recently, high redshift (**$z > 6$ **or larger) quasars have been discovered with high iron abundance. Do you think that the time scale for metal enrichment may fall in contrast with standard cosmological scenario if more quasars with high metallicity will be discovered at similar or even higher redshifts? Could it be a serious challenge to the Standard Model?**

The most common interpretation of the abundance ratios in galaxies is the "time-delay" model, namely the delay with which some stars restore their nuclear products into the ISM relative to other stars. In particular, type II SNe, which originate from core-collapse of massive stars ($M \geq 10\,M_{\odot}$), restore their main nuclear products (O, Ne, Mg, Si, S, Ca) and a fraction of Fe on timescales of the order of $0.01 \div 0.03$ Gyr. On the other hand, type Ia SNe, which are believed to originate from exploding carbon–oxygen white dwarfs in binary systems, restore Fe, which is their main nuclear product, into the ISM on a large range of timescales going from 0.035 Gyr to a Hubble time. The typical timescale for the Fe enrichment from SNe of type Ia, which we can define as the time of the maximum for the type Ia SN rate, is a function of the assumed progenitor model for type Ia SNe and of the star formation history. Therefore, it varies from galaxy to galaxy, as different SFHs characterize the Hubble Sequence . This timescale can vary from $\sim$0.3 Gyr in ellipticals and bulges to $\sim$1 Gyr in the local disk and to $\sim$4–5 Gyr in irregular dwarf galaxies. These timescales have been evaluated [348] by assuming the "single-degenerate scenario" for the progenitor of type Ia SNe and that the star formation rate is decreasing in intensity and increasing in length going from early to late type galaxies, in agreement with observational evidence [273]. The single degenerate scenario for the progenitor of type Ia SNe suggests that a binary system made of a carbon–oxygen white dwarf plus a younger star can produce a SN Ia. In fact, when the younger companion evolves and becomes a red giant, it starts losing material onto the white dwarf, which can reach the Chandrasekhar mass limit and explode catastrophically leaving no remnant. This thermonuclear explosion produces $\sim 0.6 \div 0.7\,M_{\odot}$ of Fe plus minor quantities of the elements from C to Si. In this progenitor scenario, the most massive binary system that can contribute to a type Ia SN is made of an $8\,M_{\odot}$ plus a companion of roughly the same mass. This implies that the minimum time for the explosion of the first type Ia SNe is no longer than 0.035–0.040 Gyr. Another possible scenario for the progenitors of type Ia SNe is the double-degenerate one: in other words, two carbon–oxygen white dwarfs of roughly 0.7 $M_\odot$ merge after loosing angular momentum due to gravitational wave emission. In this case, the clock to the explosion is given by the lifetime of the originally smaller mass, as in the single degenerate scenario, plus the gravitational time delay, namely the time necessary to bring together the two white dwarfs that will give rise to a Chandrasekhar mass and explode as in the previous scenario. The nucleosynthesis products are the same in the two scenarios. The minimum time for the explosion of these systems will therefore be given by the lifetime of an 8 $M_\odot$ plus the minimum gravitational time delay (0.001 Gyr, see [219]). Having said that, we see that the minimum timescale for the explosion of SNe type Ia is practically the same for all galaxies and in any scenario (the difference is only 0.001 Gyr), whereas the timescale for the maximum in the SN rate, which is relevant for chemical enrichment, changes according to the SFH, as already discussed. Observational evidence for such prompt type Ia SNe has been recently provided by [336, 337]. These authors suggested, in fact, that roughly 50% of all type Ia SNe explode soon after stellar birth, in a time of the order of 10^8 years, whereas the remaining 50% of type Ia SNe explode on a much wider timescale distribution going from times larger than 10^8 years to a Hubble time (14 Gyr). They reached this conclusion by studying the dependence of the SN Ia rate on the colors of parent galaxies and the enhancement of the SN Ia rate in radio-loud early type galaxies. The fraction of prompt SNe Ia suggested by Mannucci and collaborators is higher than the fraction generally assumed in modeling the type Ia SN rate. In particular, the type Ia SN rate, in the framework of the single degenerate scenario (e.g., [220, 348]) predicts a fraction of prompt type Ia SNe not larger than 13% of the total.

To answer the question, we can say that there is no problem in explaining the high Fe abundance observed in high redshift ($z > 6$) quasars as long as we assume that galaxy formation, in particular the formation of large ellipticals, started at redshift $z \gg 6$. In the following we explain why Quasars are generally hosted by massive ellipticals and the quasar phenomenon is attributed to matter falling into a central BH (see Fig. 2.3). The chemical abundances measured from the broad emission lines in quasars indicate supersolar abundances of several chemical elements including Fe (e.g., [138, 225, 335]).

It has been shown by [347] that in massive ellipticals, the hosts of quasars, when a strong starburst is assumed together with a standard IMF [484] and the single degenerate scenario for the progenitors of type Ia SNe, the interstellar gas can reach solar Fe abundances on timescales of the order of 0.1 Gyr from the beginning of star formation (note that in such a galaxy the relevant timescale for Fe enrichment from type Ia SNe is $\sim$0.3 Gyr), and that at 1 Gyr the Fe abundance has already reached ten times the solar value, as is shown in Fig. 2.4, where we present the evolution of the abundances of several chemical species in the ISM of a large elliptical. In this model, it was assumed that the initial strong starburst ends when a galactic wind develops. During the starburst, type II SNe already produce a non-negligible fraction of Fe and most of the oxygen on timescales not longer than 0.03 Gyr. Actually, in a strong

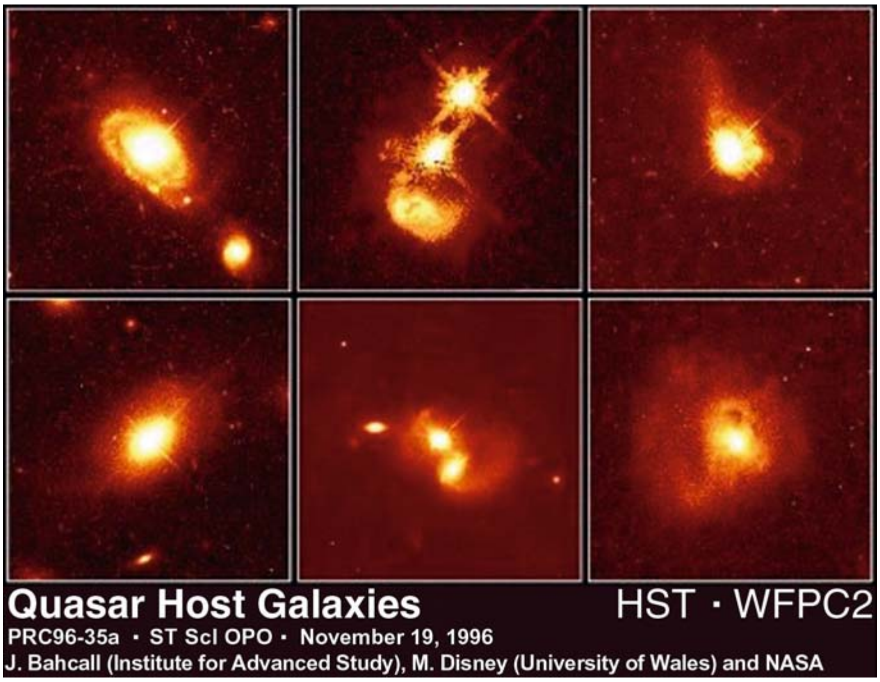

Fig. 2.3 Images of quasar host galaxies from the HST. The host galaxies are generally ellipticals

starburst like this one, type II SNe can produce by themselves enough Fe to enrich the gas up to the solar value. The wind occurs when the gas thermal energy, due to the energy deposited by SN (II and Ia) explosions, equates the potential well of the gas. The galactic wind carries away all the residual gas and then the elliptical galaxy evolves passively. At this point, the gas restored by the dying stars goes to feed the central BH and the quasar phase starts.

In Fig. 2.4, the occurrence of the wind is identified by the discontinuity in the curves (between 10^8 and 10^9 years). As one can see, after the occurrence of the galactic wind, the Fe abundance reaches values as high as 10 times solar, while the O abundance grows up to an abundance in excess of solar before the winds and remains constant and lower than the Fe abundance after the wind. This is due to the fact that after the wind, which removes most of the gas from the galaxy, the star formation (SF) process stops and therefore the O production, which is related to the short living massive stars, also stops. On the other hand, Fe continues to be produced by type Ia SNe, which continue to explode until the present time. This scenario for the evolution of the gas in ellipticals hosting quasars seems to reproduce not only the high Fe abundance at high redshifts but also the observed constancy of the Fe and other element abundances in quasars as functions of redshift.

The results of [347] have been confirmed in the following years by the models of [213, 414, 472]. Moreover, the rapid increase of the Fe abundance would further

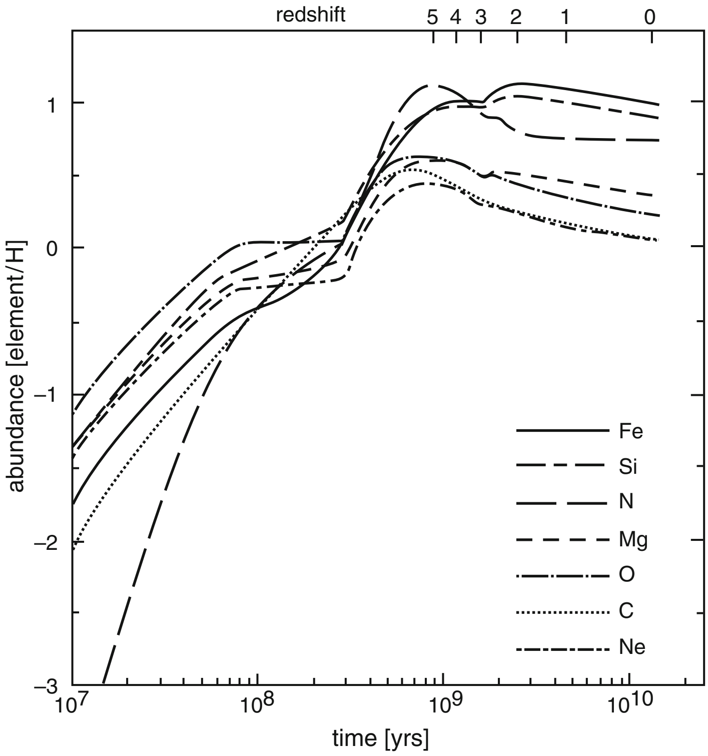

Fig. 2.4 Predicted time-dependence of several chemical elements, as indicated in the figure, relative to the solar abundances (0 means solar, 1 means 10 times solar, etc.) for an elliptical galaxy of $10^{12}\,M_{\odot}$ luminous matter. The adopted IMF is from [484] and the cosmology is the standard one with an assumed redshift of galaxy formation of $z = 10$. The time at which the abundances show a discontinuity refers to the time of occurrence of a galactic wind. From [347]

be strengthened if the type Ia SN rate suggested by [336] were adopted. In this case, in fact, the increase of the Fe abundance in time would be even faster than in Fig. 2.4, because of the fraction of prompt type Ia SNe higher than in the rate adopted by [347], as explained before.

Before concluding, it is important to note that the cosmic age of 1 Gyr (the time at which the Fe abundance is maximum) corresponds, in the concordance cosmology ($\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 65$), to $z = 5$, if the redshift of galaxy formation is assumed to be $z_f = 10$. Therefore, there is no problem in explaining the high Fe abundances observed in quasars at redshift $z = 6$ and beyond.

In summary, the timescale of metal enrichment is not in conflict with the standard cosmological scenario as long as the redshift of galaxy formation for massive ellipticals is set at $z_f \geq 10$.

Thank you very much Francesca. It is remarkable how the studies of metallicities in the local Universe and in distant galaxies and high redshift Quasi Stellar Ob-

jects (QSOs) might represent a piece of the puzzle in the understanding of cosmic structures.

It is natural to ask now why extremely bright and high redshift objects like QSOs cannot be used as standard candles. Paola Marziani will explain us why not.

2.5 Cosmology with Quasars

Dear Paola *(Marziani)*, **quasars are among the most luminous sources in the Universe, and their optical luminosity is, in most cases, stable over periods of several years. Quasars have been discovered up to $z > 6$. Type Ia SNe are, in comparison, much dimmer sources, and even the most recent studies employ SNe only up to redshift $z \approx 1.9$. Why quasars have never been effectively used as standard candles?**

The question you are asking me is both challenging and embarrassing. It is challenging because a good standard candle needs to have a known, well-defined luminosity with a small intrinsic dispersion around an average value. Or, at least, a standard candle should be based on a calibration of a measurable property that tightly correlates with luminosity. Good standard candles, especially in a cosmological context, should then be easily recognizable and highly luminous.

2.5.1 The Challenge

There is no doubt that the last two properties are met by quasars. Quasars emit a fairly univocal spectrum, with prominent broad emission lines in the optical and in the UV range. And no doubt they can be very, very luminous: their absolute magnitude reaches $M_B \approx -30$, which corresponds to a luminosity 10^4 times that of Messier 31, the Andromeda galaxy. This is unfortunately only a part of the story. If quasars can be the most luminous sources in the Universe that can be stable over periods of several years (as opposed to GRBs), they can also be comparatively faint. We can immediately think of the other extreme, at low luminosity: the famed nucleus of NGC 4395 hosts the least luminous quasar known: its $M_B \approx -10$ is just 10 times the luminosity of a typical blue supergiant star [366]. And we know that quasars can have all luminosities in between the two extrema (which are a mind-wobbling 10^8 times apart!), with a luminosity function that is open-ended at low luminosity. Nor it has been possible to identify a flavor of quasars whose luminosity distribution is peaked or even tightly constrained.

Speaking of quasars, everyone naturally thinks of those star-like objects at high redshift. After all, the term quasar comes from *quasi stellar radio source*. Quasi-stellar because the spectrum did not look like that of a star when the first spectroscopic optical observations were carried out in the early 1960s. On the other hand, quasars looked unresolved on the photographic plates, exactly like a star. And the first quasar discovered was a powerful radio source, identified as 3C 273. Let us now make a jump of nearly 50 years. There has been a sort of *luminosity unification* of quasars. Decades of observations with ever-improving instruments found that the emission-line spectrum and the spectral energy distribution of quasars are very similar over a very wide range of luminosity. We see the same lines and almost the same line widths also in the bright nuclei of the relatively nearby galaxy, the one known as Seyfert galaxies (discovered 20 years earlier than quasars but not understood at the time) as well as in luminous quasars. We observe strong and ubiquitous hard X-ray emission.

Figure 2.5 conveys the meaning of these words in term of three spectra of quasars of widely different redshift and luminosity, even if the comparison is restricted to a narrow range around the HI Balmer line Hβ. The spectra show clearly that a very luminous quasar can look like a bright, nearby Seyfert galaxy, albeit it is important to stress that not all quasars look like the ones shown in Fig. 2.5. The luminous nuclei of galaxies showing a quasar-like spectrum have become to be known as AGN, but there is no discontinuity in the luminosity distribution between Active

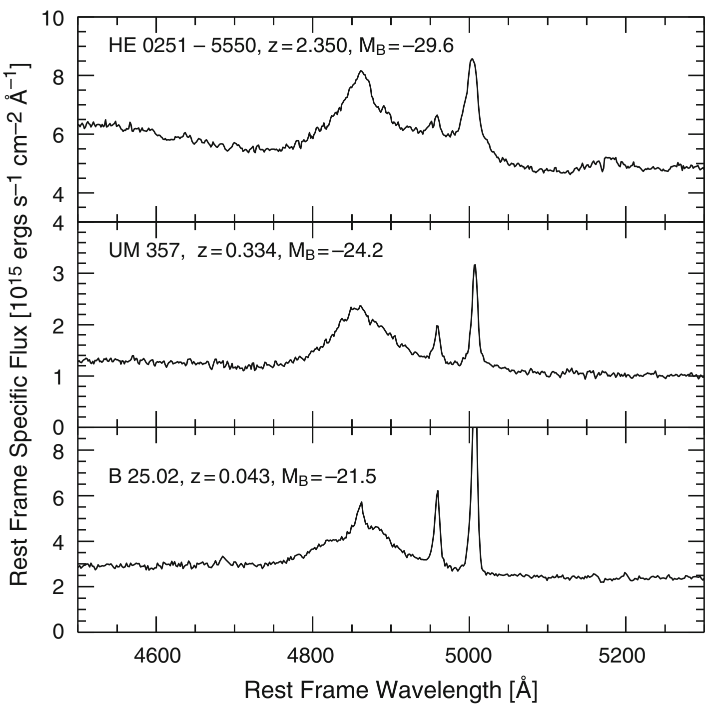

Fig. 2.5 The spectra of three quasars of widely different luminosity and redshift, covering the broad hydrogen Balmer Hβ line and the narrow, forbidden [OIII]$\lambda\lambda$ 4959,5007 lines. Note that the HE 0251-5550 is $\approx$2,000 times more luminous than *B* 25.02. While *B* 25.02 is a local Seyfert 1 galaxy, HE 0251-5550 is a distant quasars seen at a lookback time of 2/3 the age of the Universe. Yet, their Hβ spectra look very similar

Galactic Nuclei (AGN) and quasars. The distinction originated from the resolving power of the instruments: if a surrounding host galaxy is appreciable then it is an AGN; if not, then it is a quasar. We are presently able to resolve host galaxies up to redshift $z \approx 1$ [248], while the first quasar at $z \approx 0.15$ looked stellar. For much larger redshift, seeing or not seeing an underlying galaxy is still a matter of faith, but the results at $z < 1$ are still remarkable: $z \approx 1$ means a lookback time of roughly 7 Gyr, half the age of the Universe.

The previous digression is in part reassuring, as the ability to resolve the host galaxy is still a much needed confirmation of the assumption that distant quasars are luminous nuclei of galaxies as found locally. However, it also highlights why quasars are so cumbersome if one thinks of their potential use for measuring fundamental cosmological parameters like the Hubble constant H_0, the energy density associated to matter Ω_m, and to the cosmological constant Λ, Ω_Λ. We have a class of sources whose luminosity is spread over an enormous range in luminosity, and whose spectral properties is fairly similar over a large range of redshift, basically from local $z \approx 0$ Seyfert 1 nuclei to the most distant quasars at $z > 6$! And there are more sources of concern.

Quasars are anisotropic sources in most regions of the electromagnetic spectrum. Two main effects contribute to anisotropy: relativistic beaming in radio-loud sources, and obscuring material co-axial with the accretion disk in both radio-loud and radio-quiet AGN. This is the essence of the so-called *Unification Scheme* of quasars and AGN [248]. If a quasar is seen through a thick structure of absorbing gas and dust, its innermost emitting line regions will be obscured and the UV soft-X spectral energy distribution will be strongly affected. We will ignore these obscured AGN (conventionally called type-2) in the following. However, beaming and orientation effects are not yet fully understood as far as their influence on optical/UV spectroscopic properties of *unobscured* AGN (conventionally termed type-1) are concerned. Their occurrence in the main flavor of quasars, the ones that are radio-quiet (about 90% of all quasars), is even more enigmatic: for example, who can tell which radio-quiet AGN are seen pole-on in analogy to radio-loud BL Lacs and optically violently variable radio quasars?

But we know that orientation matters. Let me make just one example. Core-dominated and lobe-dominated quasars, which are thought to be sources whose radio jet is, respectively, almost aligned or grossly misaligned to our line-of-sights, show different Balmer line widths, by a factor $\approx$2. This has remarkable consequences on physical parameters estimation, as I will try to explain later.

For the moment it is important to keep in mind that quasars are such pranksters that they look different if they are seen along different line-of-sights. And that we do not understand well how. Is this enough not to plunge anyone into deep depression? And we are still not done. There is also the embarrassment.

The embarrassing side of your question is that quasars are plentiful: data for $\sim 10^5$ are presently available from the SLOAN Digital Sky Survey (SDSS) Data Release (DR) 6, more than 13,000 with $z \geq 2.3$ (for comparison, you can consider that only 200 quasars were catalogued in 1971!). Quasars are not only luminous sources, much more luminous than type Ia SNe ($M_\mathrm{B} > -30$ vs. $M_\mathrm{B} \approx -21.7$), they are also variable. The last resort is to look for parameters that we can easily measure and that tightly correlate with luminosity. Is it possible that, in 50 years of intensive quasar research, none has yet found even one suitable parameter?

2.5.2 *Exploiting Quasar Variability*

Let me recall that optical variability has been established as an identifying property of type-1 AGNs for more than three decades. AGN typically show continuum variations by 1–2 magnitudes with timescales ranging from days to years. Broad emission lines have also been found to vary. True, every attempt to find a periodicity in quasar variability patterns failed. A period–luminosity relationship is not even to be mentioned! But can we somehow exploit the variability of quasars?

A key idea is to consider that emission line variations lag the continuum variations with delays ranging from a few days to months in luminous Seyfert 1 nuclei. The cross-correlation function between the continuum and the emission line light curve then measures a time lag Δt_{obs} due to the travel time needed by continuum photons to reach Broad Line Region (BLR). This means that the distance of the BLR from a supposedly point-like, central continuum source can be simply written as

$$r_{\mathrm{BLR}} = \frac{c\Delta t_{\mathrm{obs}}}{1+z},$$

where the factor $(1 + z)$ reduces the observed time lapse to the time lag in the rest frame of the quasar. The evaluation of Δt_{obs} follows from several assumptions, mostly untested. Some of them may be even physically unreasonable. The coupled effects of a broad radial emissivity distribution, an unknown angular radiation pattern of line emission, and suboptimal temporal sampling of the light curve can cause errors that are difficult to quantify.

At any rate, the basic idea beyond exploiting time delays is to measure the linear size of a chosen structure from light travel times (i.e., in a way independent from H_0), and an angular size from a resolved, direct image of the same structure [156]. If we were able to measure the angular distance of the BLR from the central continuum source, then we could recover the cosmological angular distance d_{A} between us and the source, which could be written as

$$d_{\mathrm{A}} = \frac{r_{\mathrm{BLR}}}{\theta_{\mathrm{BLR}}} = f(z \mid H_0, \Omega_\Lambda, \Omega_m, \Omega_{\mathrm{k}}) \approx f(z \mid H_0), \text{if } z \ll 1.$$

But if the determination of r_{BLR} can be fraught by large uncertainty, the measurement of an angular size is even impossible. Trouble is that angular size measurements of the BLR are prohibitive with present-day technology: for the Seyfert 1 nucleus of NGC 5548, one finds $\Delta t \approx 21$ days; at $z = 0.017$, which means an $d_{\mathrm{BLR}} \approx 50\,\mu$arcsec. Even if the BLR linear size increases with luminosity as $\propto L^{0.7}$ [268], resolution better than $10\,\mu$arcsec is still required to resolve nearby quasars: for example, 3C 273 has $\Delta t \approx 387$ days at $z = 0.158$, and this delay corresponds right to $\approx$10 μarcsec. Clearly, the method is not yet applicable to any AGN near and far, although some measurements should become feasible with optical interferometers that are presently under development, like the Space Interferometry Mission Planet (SIM PQ) proposed to NASA [564].

In the meantime, methods based on time delays are probably bound to remain model-dependent until the required angular resolution will be achieved. A tickling attempt of the model-dependent kind has been based on the observed time-scale difference between continuum flux variations at different frequencies. No angular size measurement is required at the expense of a major assumption: the existence of an accretion disk around a massive BH. It is known that variation timescales are shorter at shorter wavelengths, so that the observed time delay $t(\lambda)$ in the optical-UV continuum is wavelength-dependent. The assumption that the optical-UV continuum is emitted by an illuminated, geometrically thin accretion disk allows to scale the disk radial temperature $T(r)$ with $r(\lambda)$, that is, with the delay multiplied by the speed of light [110]. The observed specific flux is then

$$f_\nu \propto t^2 d^{-2} \lambda^{-3}.$$

This relationship can give an H_0-independent distance d, suggesting $H_0 \approx 42 \pm 9$ from data for the Seyfert 1 nucleus of NGC 7469.

The story is different if there is an intervening galaxy between us and the quasar, and especially if the galaxy is not perfectly aligned with the quasar along our line of sight. In this case, a galaxy (or any other massive object like a cluster of galaxies) acting as a gravitational lens yields multiple, asymmetrically displaced images of the quasar. Following the intrinsic light variations of the quasars, one measures different time delays for the displaced images due to the path-length difference between the quasar and the earth, and also due to the gravitational effect on light rays traveling in slightly different potential wells. As a consequence, the computation of H_0 requires model-dependent assumptions on the gravitational potential of the intervening galaxy. The resulting H_0 value is usually below or in agreement [496] with the value obtained from the Cepheids, $H_0 = 72 \pm 8$. As multiple images often show an accessible angular separation, the method is promising and several campaigns (e.g., Supernovae and H_0 for the Dark Energy Equation of State (SHOES), COSmological MOnitoring of GRAvItational Lenses (COSMOGRAIL)) are underway to fully exploit its long-term potential. But I will not dwell on that, as I suppose that your question refers more to the intrinsic properties of quasars.

2.5.3 Quasar Diversity and Quasar Evolution

At high z, we are observing quasars that can be very similar to the AGN we are observing at low z, in terms of line width, prominence of singly ionized iron emission, and equivalent widths of other emission lines [538]. Luminosity effects remain weak and prone to sample biases. As we will see better later, there are samples where the main luminosity correlations (the Baldwin effect) are not significant while several properties correlate with the luminosity-to-mass ratio L/M. This can be actually *measured* as the ratio between total luminosity (i.e., bolometric) and mass of the central compact object of a quasar. It is also important to stress that L/M is proportional to the Eddington ratio, that is, the ratio between bolometric and Eddington luminosity, which is considered, under some conditions, a limiting luminosity for the accretion process. We will exchange L/M and Eddington ratio as synonyms in the following.

The lack of a strong luminosity dependence may reflect a self-similarity in the accretion process, which is as yet not fully understood and exploited for quasar modeling, even if phenomenological analogies between accreting systems with stellar-mass BHs (the so-called "mini-quasars") and the supermassive BHs found in AGN are now recognized [360]. Two daring people [599] even suggested an analogy between quasars and two accreting white dwarves (not even BHs!), showing an optical emission line spectrum excitingly similar to the one of I Zw 1, the prototypical Narrow Line Seyfert 1 (NLS1) nucleus – save for different line widths, and save a factor of 10^8 in the mass of the accreting compact object.

True, it is also known that there is a strong *luminosity evolution* of quasars: we live in an Universe that, locally, does not possess luminous quasars. The most luminous quasars we detect now shone a long time ago, and very far from us, at $z \approx 2$ [448]. But now we see, among galaxies in the local Universe, the signature of that brilliant past: very massive compact objects (even a few billion solar masses) that were once accreting material at a pace large enough to sustain an enormous luminosity, and that are nowadays literally extinct, or accreting at very low Eddington ratio. On the other hand, not very far from us, we see sources accreting close to the Eddington limit, but whose masses are by far less than the ones of the quasar population that was once so luminous. As we shall see later, here is where the quasar spectral diversity comes in. For the moment, let me still consider in some more detail how luminosity seems to affect quasars.

2.5.4 The Baldwin Effect

Baldwin and co-workers noticed almost 30 years ago an inverse correlation between the equivalent width of the CIVλ1549 emission line and the apparent luminosity of bright quasars [21, 22]. Quoting the original 1978 paper [21],

> The data indicate that the luminosity of QSO emission lines increases as the 1/3 power of the continuum luminosity.

In other words, the lines, even if their luminosity increases, become less prominent over the underlying continuum with increasing luminosity. In the origin, the effect was believed to be fairly strong, with the equivalent width of CIVλ1549 decreasing proportionally to $L^{-\frac{2}{3}}$. Jumping to present times, if we focus the attention on lines emitted by ions of ionization potential $\gtrsim$50 eV (which we call high-ionization lines for brevity), then we observe a significant luminosity dependence. However, it is important to stress that the *Baldwin effect* has survived as a much weaker and very loose anticorrelation between specific luminosity and high-ionization lines equivalent width. Claims and counter-claims of a Baldwin effect on the basis of small samples (few tens of objects) are unreliable; the statistical weakness of the Baldwin correlation implies that the effect becomes significant only if a very large range in luminosity is considered, $4 \div 6$ decades, as also confirmed by Monte-Carlo simulations [536]. Results until mid-1999 have led to a standard scenario in which the slope of the Baldwin relationship between logarithm of equivalent width of CIVλ1549 (the most widely studied high-ionization line) and luminosity is ≈ -0.15, and not $-\frac{2}{3}$ as originally thought. The Baldwin effect seems to occur in all measurable high-ionization lines except NVλ1240, and the slope of the anticorrelation increases with the energy needed to create the ionic species from which a given line originates. These results have been basically confirmed by more recent studies based on large quasar samples [115, 218]. The anticorrelation of CIVλ1549 remains very weak, however, and cannot be exploited, as it is, for any cosmological purpose. The left panel of Fig. 2.6 shows the disarming spread of data points in a low-redshift sample.

Yet, neither do I share the pessimistic opinion that sees the Baldwin effect as a stalwart of a dogmatic view in quasar and cosmology research, nor do I share dogmatic views that have been widely accepted in the past and that considered the Baldwin effect as a "must be". What the Baldwin effect basically tells us is that

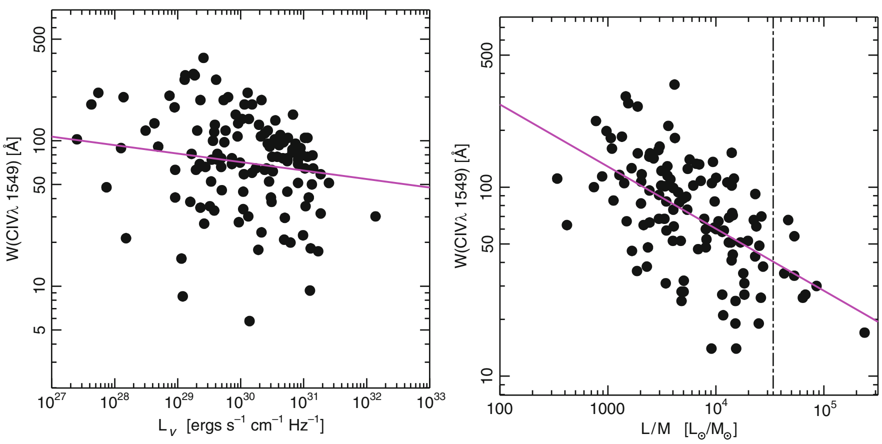

Fig. 2.6 *Left panel*: The weakness of the Baldwin effect in a sample of low-z quasars for which measurements of the CIVλ1549 high ionization line come from archived HST observations [540]. Abscissa is the specific luminosity; ordinate is the rest-frame equivalent width of CIVλ1549. *Right panel*: the Eddington ratio dependent "Baldwin effect", for quasars of the previous sample with a BH mass estimate from the Hβ line width. The abscissa is the luminosity to mass ratio expressed in solar units. log $L/M \approx 4.53$ corresponds to unity Eddington ratio (*dot-dashed vertical line*). We use the term "Baldwin effect" in an improper way here, as it customarily means an inverse correlation with luminosity. The inverse correlation with Eddington ratio is, however, much stronger than the one with luminosity

the quasar spectrum is not fully redshift/luminosity independent. What it says is that quasar spectra become systematically of lower-ionization with increasing redshift/luminosity, as low-ionization lines seem to be even less affected by luminosity, or not affected at all [538]. And that there is a large spread in equivalent width of high-ionization lines for a given specific continuum luminosity. Both aspects cannot be ignored. A few years ago, our group suggested that the Baldwin correlation may mainly reflect a combination of quasar luminosity evolution and of selection effects in the L/M ratio [536].

It is still debated whether the Baldwin effect is primarily evolutionary in its nature, but at this point a parenthesis must be opened to consider that quasar spectra are not all self-similar. Even if luminosity is not what affects spectra at most, there is still a considerable diversity in quasar spectral properties. From the optical/UV spectra, we go from sources of low overall ionization, prominent singly-ionized iron emission, and relatively narrow HI lines, low-equivalent width of CIVλ1549 to objects with weaker or almost absent FeII emission, broader lines, prominent high-ionization CIVλ1549. If we consider a CIVλ1549 equivalent width versus luminosity plot for low-z quasars, then we see that a considerable source of scatter is added by some low CIVλ1549 equivalent width sources at $\sim 10 \div 30$Å that tend to blur the Baldwin relationship. These sources are of the "low-ionization" kind, and include NLSy1s. Just a few years ago, two groups working on the CIVλ1549 emission feature from archived HST observations realized that CIVλ1549 equivalent width correlates much more strongly with L/M rather than with luminosity [19, 27]. Figure 2.6 shows the luminosity and the Eddington ratio correlation side-by-side for the same sample [540].

Actually, Eddington ratio seems to be relevant not only for physical conditions but also for the dynamics of the BLR gas. Low CIVλ1549 equivalent width sources can show prominent blueshifted CIVλ1549 with respect to the quasar rest frame, while sources with more prominent CIVλ1549 show no large shifts, at least at low-z [540]. Almost all low-redshift quasars belong to a sequence in the plane defined by the Full Width Half Maximum (FWHM) of Hβ and by the prominence of FeII emission in the optical spectrum. The main variable that seems to govern this sequence is again Eddington ratio [536]. We can safely conclude that sources with lower ionization spectra, including NLSy1s, are the ones radiating at higher Eddington ratio, although quantifying each quasar's Eddington ratio from spectral parameters is still an open issue.

2.5.5 Exploiting the Luminosity-to-Mass Ratio

It is now possible to glimpse a way out from the impasse. No strong dependence on luminosity, but several, easily measurable spectral parameter correlate pretty well with L/M. If we can find a very tight correlation with maybe a linear combination of observed spectral parameters and L/M, and we are so clever to measure the central BH mass, then it is obvious that we could retrieve a redshift-independent value of luminosity [340]. It is not yet possible to do that in a meaningful way. What we miss here is the equivalent of an Hertzsprung–Russell Diagram (HRD) for quasars. We have a view that is sketchy at best. Mass estimates have become an easy exercise, applied even to samples of tens of thousands of quasars, but they rely on two major assumptions: (1) that the gas motion giving rise to the line Doppler broadening is predominantly virial, and (2) that the size of the BLR correlates with optical or UV luminosity following a power law of index $0.5 \div 0.7$ [268]. This means that the M can be simply written as

$$M \propto r_{\mathrm{BLR}} \mathrm{FWHM}^2 \propto L_{\mathrm{opt}}^{\alpha} \mathrm{FWHM}^2,$$

where the FWHM of a suitable line is considered, preferentially a low-ionization line like the Balmer Hβ or MgIIλ2800 [352]. These mass estimates have a statistical value as the inferred uncertainty for individual sources (a factor of 3 at best) is still large. One can then apply a bolometric correction to retrieve the Eddington ratio, an approach seemingly rough but relatively stable as a matter of fact [341].

AGN have proved harder to understand than main sequence stars, whose physical characteristics are determined mainly by a single parameter, that is, mass, to a lesser extent by metallicity and age, and to virtually no extent by orientation. A 2D parameter space (the HRD) is sufficient to characterize both main sequence and nonmain sequence stars. It has become obvious that this is not possible for AGN, even for those with broad emission lines. The aspect-dependent phenomenology due to accretion of matter constrained in an accretion disk demands that at least an aspect parameter θ be taken into account. We think that, for typical quasars, the orientation parameter θ varies from a few degrees to $45° \div 60°$, beyond which the object appears as an obscured, type-2 source, and its BLR view is hidden from us [565]. Orientation matters in the width of the emission lines – this is known since almost 20 years and has been confirmed by several later studies – but we still do not know how to estimate the θ angle in individual sources, with the exception of a few very special objects [537]. We can be confident that orientation effects are a factor $\approx$2 in radio-loud sources, but we are afraid that the effects could be much larger if line emission is constrained in a strongly flattened system. This could be the case of at least some radio-quiet quasars but, at present, none knows for sure.

2.5.6 Guessing Further...

Where do we go from here? Is this all quasars can tell us? We do not have a satisfactory theory that connects accretion parameters of quasars to the structure of the line emitting regions and to measurable spectral properties. However, not even Galileo could count on theory of optics when he mounted his telescope, so please allow me a little explorative calculation. The question is too important to be dropped in this way. After all, we learned a few things since early attempts to exploit the most luminous quasars' Hubble diagram to confirm the expanding Universe scenario [20].

As already pointed out, we have been able to estimate masses and Eddington ratios for large samples of quasars. I know not everyone will agree with me but, as far as the present-day, "best" data are concerned, one can make two cautious statements. The first is that there is no convincing evidence of sources radiating above the Eddington limit. The second is that the extremely large BH masses are not observed. Rather, data are tantalizingly consistent with a maximum BH mass $\sim 5 \times 10^9\,M_\odot$. This mass limit is consistent with the largest spheroids observed in the present day Universe, for which there is a known spheroid–BH mass relationship [539]. Also, the mass function of BHs in quasars at $1.6 \lesssim z \lesssim 2.6$ seems to drop sharply above that value [573].

It is also important to stress that the BH mass follows from a measurement of time delay. Even if we then employ a correlation with luminosity [268], the mass value is independent from H_0, as changing H_0 will change the luminosity but will not affect the mass, that is, changing H_0 will just produce an horizontal shift in the plane r_{BLR} vs. L [550]. Therefore, one is tempted to assume that the most luminous quasars are the ones radiating at their Eddington limit and at their largest mass. In other words, we can estimate the maximum intrinsic luminosity of quasars. So we can derive the apparent magnitude of the brightest sources at each redshift. And this is a function of the cosmological parameters.

Stated this way, the argument is too simplistic. It is known that there is a strong luminosity evolution with cosmic age. But the argument could still be applicable to the most luminous quasars. We know that the luminous quasar population peaked at $z \approx 2$ [448]. Since then, quasars faded and some of them became even almost extinguished by our detection standards. So, sources at $z \ll 2$ should be considered with care: they may radiate well at Eddington ratio (we have very good examples of local AGN radiating close or at Eddington ratio), but the accreting masses could be lower than the assumed maximum mass due to the quasar strong luminosity evolution. If we consider the B photometric band, then one should take into account that $z \approx 1.6$ and $z \approx 3.0$ imply contamination by the strong UV lines of CIVλ1549 and hydrogen Lyα. In addition, shortward of Lyα, quasar spectra often show very complex patterns of narrow absorptions, due to the Lyα absorption by neutral gas clouds between us and the quasars. But why should not we give a look at the brightest quasars at least in the range $1.6 \lesssim z \lesssim 3.0$ once we keep in mind these problems?

It is really intriguing that observational data seem to constrain the cosmography. In Fig. 2.7, the three dashed lines describe the expected magnitudes for sources radiating at Eddington limit as a function of redshift for three different H_0 values (50, 75, 100). A simple k-correction, appropriate for sources at $z \approx 2$, has been assumed. Large H_0 values are not favored; rather, the observed brightest B magnitudes (corrected by Galactic absorption) apparently favor a small value of H_0, $\approx$50, not unlike methods based on gravitational lenses. Data collected from the Hamburg-ESO survey are the most homogeneous and are therefore preferred, but the previous "cosmological" conclusion is not strongly affected if we consider a query for quasars in the 12th edition of the catalogue by Véron-Cetty & Véron [572].

Of course, a cosmological inference can be entirely washed away if quasars are subject to a physical limit in their mass accretion rate, so that very massive BHs

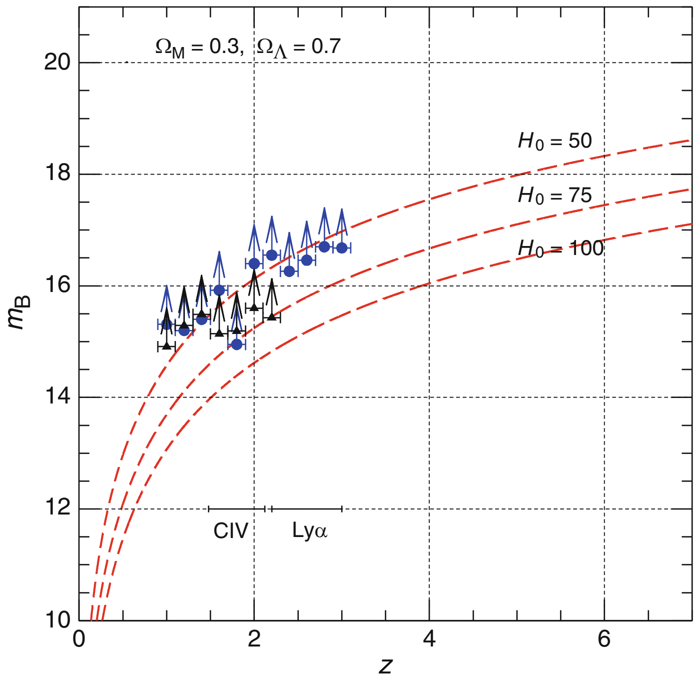

Fig. 2.7 The *dashed lines* show the expected magnitudes of the brightest quasars as a function of redshift for three different values of H_0, assuming a flat Universe with $\Omega_\Lambda = 0.7$ and $\Omega_m = 0.3$, and a "maximum" BH mass $M_{\rm BH} = 3 \times 10^9\,{\rm M}_\odot$. The data points represent the blue magnitudes (corrected by Galactic absorption) of the brightest quasars, in bins of $\delta z = 0.2$. *Blue filled circles*: Hamburg-ESO survey; *grey triangles*: data from the 2006 edition of the Véron-Cetty & Véron catalogue of quasars and AGN [572]. In this case, no data points were plotted for $z > 2.2$, as all brightest sources belonged to the Hamburg-ESO survey. The segments indicate the redshifts at which the strong emission lines of CIV1549 and Lyα are in the B passband

never radiate close to Eddington ratio, or, even simpler, if surveys missed the brightest quasars in the sky.

Resuming a cosmological interpretation, if we now compare models with $\Omega_\Lambda = 0$ and $H_0 \approx 50$ fixed with the same observational data of Fig. 2.7, we find that $\Omega_m \to 1$ is not favored: otherwise we would observe too many bright quasars. In the redshift range $1.6 \lesssim z \lesssim 2.5$, where our comparison seems safer, there is a discrepancy of $\sim$1 magnitude between the observed magnitudes and the prediction for $\Omega_m \approx 1$. But I am not going to show you this diagram.

I am anyway tempted to say that we live in a Universe where quasars become systematically fainter with redshift; the way they do and the way they shine when most luminous suggest a pretty "large" Universe by current standards, consistent with the currently accepted value of Ω_Λ or, if $\Omega_\Lambda = 0$ and $\Omega_{\rm k} = 0$, with an open Universe described by a small value of Ω_m.

Of course, the previous computations are only for illustrative purposes, even if they may not be very far from correct results. Quasar properties pose an enormous breath of challenges, and so it may not now sound surprising if quasars have not been yet used as widely successful cosmological probes. Nonetheless, I hope to have given you a glimpse of how a better understanding of quasar physics and evolution could make a difference for cosmology.

Thanks Paola for this overview of quasar properties. The nature of these objects has been a mystery since their discovery in the 1960s. Anyway, speaking of quasars, it is impossible to avoid the question of the anomalous redshifts observed for some of them. This is strictly connected to the cosmological interpretation of redshift on which we have founded our cosmology. We are now going to interview Jack Sulentic, who has always manifested his doubts about this interpretation.

2.6 The Heretical View on Cosmological Redshifts

Dear Jack *(Sulentic)***, are there viable alternatives to the cosmological explanation of redshift of distant galaxies and quasars?**

This is one of the questions posed in this book dealing with controversial ideas or observations. They were controversial 30 years ago and they are controversial now. Of course, many will say that the answers are well known. That the observational evidence has been refuted. That the subject is closed. I could respond to your question by restating all of the empirical evidence and reviewing all of the alternative ideas. But to what point? They (and technical references) can be found in many places (e.g., books such as "New Ideas in Astronomy" [43] and "The Red Limit" [171] or in the film "Universe: The Cosmology Quest" [357]). I will restate and update some of these results/ideas, but repeating them in detail would be a waste of time. The responses to the posed question will vary greatly from person to person. Some will say that these issues have been settled and others will say that they have never been seriously considered. I like to believe that I am somewhere in the middle and hope my responses reinforce that impression.

Moving on to the question.

There are actually two different questions that should be inferred from such a query:

1. Is there a need for an alternative explanation? Is there empirical evidence that places the standard explanation in doubt?

and

2. Whether needed or not, are there mechanisms capable of producing a pseudo-Doppler redshift?

Perhaps many careerists[2] would refuse to discuss the question beyond calling it nonsense. Others would refuse to decouple the two questions – stating that no empirical evidence would be convincing to them unless a suitable physical mechanism were already known. At the risk of being labeled a Baconian, I would argue

[2] *See the author definition of careerists and Baconians in Chap. 4.*

that empirical evidence should be judged independently of existing beliefs or ideologies. If a need for new mechanisms were felt, then I am sure that we would see many alternatives published in short order and by careerists who would see alternative models as a new avenue to advancement. Instead, should we believe that we live at a special time in the history of science where most of the basic laws have been discovered? Even if it were true one should resist the inclination to approach science this way because it would essentially preclude new discovery.

There have been vocal advocates for the view that empirical evidence does exist – even claims of overwhelming evidence for non-Doppler redshifts (e.g., [13,14,80–82,554]). Are iconoclasts always Baconian or can they also be careerists? Having spent a number of years exploring much of the evidence, I would not describe it as overwhelming but I would argue that it deserves to be taken seriously and tested. Some of those evidence will be discussed later. Much of it involves apparent associations between objects (e.g., galaxies and quasars) with significantly different redshifts.

Perhaps, the fairest thing to say is that the evidence for Doppler redshifts is somewhat more compelling than the evidence against. It is very difficult to look evenhandedly at both sets of evidence. The most compelling evidence for the paradigm is not, in my opinion, evidence for evolution in the Universe. Recall please the earlier mentioned low S/N observations of high z sources using large telescopes. The lack of evidence for evolution in the Universe is what surprises me. We find "old" galaxies at high redshift and we find quasars that are spectroscopically similar over the entire redshift range ($0.1 \lesssim z \lesssim 6$) that they are observed. This includes super-solar heavy element abundances in the highest redshift quasars. It also includes inferred BH masses as large as $\log M_{BH} \sim 9.5\,\mathrm{M}_{\odot}$ at all redshifts. One sees what one wants to see, and if one is a careerist, it is very difficult to see anything wrong with the paradigm that has advanced ones career. If a problem arises (e.g., super solar abundance in $z \sim 6$ objects), agility becomes a necessary careerist skill – add a small additional complexity (i.e., everything interesting happened before $z = 6$, where we could not see it happen because the quasars were enshrouded by dust or something—a new form of DM!) to the standard model and it fits! I think the most impressive support for Doppler redshifts is something much less subject to mixed interpretations. Something like gravitational lenses (and arcs) where we see multiple images of the same (?) high redshift quasar lensed by a lower redshift galaxy that is almost certainly in front of it.

Summarizing my response to the question: It is not clear that there are viable alternatives but few are looking for them. Feynman was not convinced by the evidence for non-Doppler redshifts, but said that, if convinced, he would look for a mechanism involving the correlation of light.

Given the success of theories in explaining so many high-energy phenomena in quasars and related sources, does it make still sense to reason like Faraday when we already have Maxwell's theory available?

What do you mean by success? Most astronomers know about gravity. Some (X-ray astronomers and jet theorists) understand aspects of plasma physics and Magneto

Hydro Dynamics (MHD). Few observations allow us to constrain well effects driven by electromagnetic forces. It is still nontrivial to measure magnetic fields in galaxies and quasars. Theorists can build models making sophisticated use of Maxwell's Equations. Again observations of jets are one area where the cross-talk between theory and observation has become quite sophisticated. Maybe there are other areas that I am unfamiliar with.

As BHs are mentioned in the questions and as they are regarded as the central engine driving quasars, perhaps they provide a better way to answer this question, or to illustrate why the answer to the question is "yes". We hypothesize a central supermassive object in quasars and that quasar activity manifests accretion onto this BH. We cannot see it and we cannot distinguish between different BH structures driven by spin, but we think that an accretion disk surrounds most BH and gives rise to much of the line emission. For a brief period, we thought to be able to distinguish Kerr and Schwarzschild configurations via the earlier mentioned 6.4 keV Fe Kα line arising at the inner edge of the accretion disk. But that was with low S/N (ASCA) data. Now we have higher S/N (XMM-*Newton* and *Chandra*) X-ray data and they rarely show the signature that we thought we saw before. Current attempts to prolong this game are sad and desperate. Recall the previously mentioned difficultly in getting time on large telescopes for high S/N observations in extragalactic astrophysics and cosmology? No careers have advanced with negative results or refutations involving better data.

The current desperate game involves estimating the masses of the central BHs, especially, estimating masses for high redshift quasars and comparing them to the local ones. The paradigm says that they should be smaller at earlier times, because they are thought to grow by accretion and mergers. But earlier predictions have a habit of being forgotten if the evidence points in another direction. We use emission line widths (or velocity dispersions) and assume that the lines arise in a virialized distribution of emitting clouds. This seems to work best for the Balmer lines of hydrogen. But the good lines are lost (redshifted out of the visible) at quite modest redshift $z \sim 0.7 \div 0.9$, then what do we do? Use other lines that from their measured properties do not likely arise from the clouds producing the Balmer lines and that show characteristics that throw the virial assumption into doubt? Virial assumption ($2T + P = 0$; here T and P are kinetic and potential energies, respectively)! We are talking here of potential energy and kinetic energy not of Maxwell's equations: Undergraduate physics. This embarrasses careerists who respect, and like to show off, complexity and technical prowess (Albert Einstein also mentioned this specific type of careerist in his tribute to Max Planck, see, e.g., [171]). It does not embarrass Baconians – we are where we are and we try to always move forward without jumping over the problem. We follow the Balmer lines out to $z \sim 3$ [539] and find the same large BH masses that are found locally. In fact, within about $z \sim 1$, the largest BH masses may be smaller than the ones observed at higher redshifts. There is some evidence that this BH obesity problem may extend to $z \sim 6.5$, as far as quasars are currently observed. There are a few areas, for example, modeling the details of jets, where Maxwell's equations can be applied and comparisons can be made between models and observations. But these detailed models do not tell us how jets are produced or why only a few percent of quasars manifest them.

2.6.1 On the Wolf Effect

Dear Jack *(Sulentic)*, **among plasma physicists, the Wolf effect has been considered a possible cause of noncosmological redshifts. Can you please explain why this effect should be relevant under the physical conditions expected for line emission from quasar? Can this effect account for the internal shifts observed between lines emitted by ions of widely different ionization potential?**

Emil Wolf is an impressive scientist who easily satisfies the definition of a truth seeker (Baconian). He was therefore rather naive when he suggested that a mechanism that came to be called the Wolf effect might have an application in astrophysics. This was a mechanism capable of producing non-Doppler shifts in spectral lines. Any application offering this possibility will be dismissed by careerists because any demonstration of a non-Doppler component, however small, could be said to open Pandora's Box. I accept some of the blame for encouraging him to explore such possibilities. I think he was genuinely surprised by the rancor and hostility that greeted his suggestion. The Wolf effect can be included in a general category of scattering mechanisms that can in principal (i.e., given the proper set of physical conditions) shift line emitting photons to longer (or shorter) wavelength. Others include Compton and Raman scattering mechanisms. Compton downshifting of photons is well known especially among X-ray astronomers. All of these mechanisms might play a roll in complex sources like quasars. Compton scattering was invoked [516] to explain a significant (but not cosmological) redshift observed in the 6.4 keV X-ray line discovered in many low redshift quasars in the 1980s. It was not warmly welcomed but now higher S/N spectra reveal that most of the redshifted lines were not real.

Scattering mechanisms do not seem promising as a way to produce all or most of the cosmological redshift. I am not qualified to discuss such models in detail, but the empiricism can provide first-order constraints for such models. Producing small shifts or asymmetries in emission lines is one thing, but shifting the bulk of the photons outside the envelope of the intrinsic (rest frame) line is quite another. A scattering process generally broadens a line and alters its shape. The broad and complex emission lines in quasars offer a tempting target for scattering applications. Electron scattering, for example, almost certainly has some small effect on emission line structure in quasars. Unfortunately, a large fraction of quasars also show one or more narrow emission lines with the same or very similar redshift as the broad ones. Scattering mechanisms also often produce wavelength-dependant shifts. In recent years, we have been able to compare the emission lines at UV, Optical, and IR wavelengths, and we find that all lines in a quasar show the same redshift within a scatter of at most $4 \div 5 \times 10^3\,\mathrm{km\,s^{-1}}$ (UV emission lines do show a systematic blueshift relative to optical lines in perhaps 60% of quasars). In summary, we do not know enough about the physical conditions within the central regions of quasars to rule out scattering mechanisms, but to produce a pseudo-cosmological redshift, they would have to scatter all of the lines from Lyα to Paschen α in a way to produce the same redshift and preserve their intrinsic widths. If I were looking for an astrophysical application of the Wolf effect in quasar astronomy, it would be to explain small scale shifts and asymmetric differences within a source.

In summary, the Wolf effect appears most promising for explaining smaller line shift and shape anomalies. It does not appear promising as a mechanism that might produce non-Doppler redshifts that could mimic observed cosmological redshifts.

2.6.2 Anomalies with Quasars?

Dear Jack *(Sulentic)***, apparent connections like luminous bridges and tails between sources of widely different redshift (typically, a nearby galaxy and a distant quasar) have been known since long. A seminal case, the connection between NGC 4319 and Markarian 205, provoked a vigorous debate until a post-COSTAR HST image led to claims that the two sources are widely separated by time and space. In the last few years, several new odd sources have been discovered or studied with more advanced instrumentation, notably around NGC 7603, and close to the nucleus of NGC 7319. What is the astrophysical significance of Mark 205 and of the other alignments/superpositions? Can they be dismissed as chances or are they extremely unlikely occurrences that straightforwardly point toward a problem with our current understanding of redshifts?**

NGC4319 + Markarian 205 and NGC7603ab are two of the most famous examples of apparently connected objects with very different redshifts. Perhaps the strongest argument against their physical reality involves their rarity rather than their redshift differences. If all or most of the high redshift quasars are ejected from the nuclei of low redshift galaxies, as Arp has hypothesized, then one might expect to see many more with luminous connections. Of course, one can argue that they are ejected at high velocity but their rarity argues in favor of the chance projection hypothesis. Another problem: Markarian 205 is embedded in a host galaxy of like redshift. If quasars represent young matter ejected from galaxy nuclei as quasars with high non-Doppler redshifts, which subsequently develop a host galaxy, then why do we still see a connection in this case – where host galaxy of the quasar is already developed?

Sources like NGC4319+Markarian 205 are feared by careerists but are viewed with amusement and curiosity by Baconian types. These puzzles make science more fun – it must be a lot less fun when one is constantly worried about keeping an unblemished reputation and career advancement. Fear of these puzzles has been manifested in the way that the counter evidence has been advanced and accepted. Any logically fallacious and/or statistically weak argument against their reality is almost immediately accepted and circulated; the recent culmination of this trend –

surprising since there has been no technical discussion of NGC4319-Markarian 205 for almost 20 years – in an above mentioned HST press release declaring the case closed. The data was never published. Five minutes of manipulation (not modification) of the HST data reveals that the luminous "connection" is still there (as presented in [535]). The evidence supporting the reality of the feature has not been refuted or otherwise explained – it is simply disregarded. If the skeptics are so certain of the correctness of their view, then why are they afraid of the data and why would they warn young people away from studying such data? Even I realize that these strange configurations *must* be accidents – and that is all the more reason why I should study them until I understand them past this superficial reason for rejecting them. I am not upset about any of this because I do not work in this area anymore. However, I could not allow the press release to pass – but the cost? No HST time for any project for 13+ years – with 30+ proposals submitted during that time period.

Lets consider these two famous cases in more detail.

NGC4319+Markarian205 involves a low redshift ($z \sim 0.071$) quasar projected inside the arms of a low redshift ($z \sim 0.004$) spiral galaxy [12, 586]. A low surface brightness luminous filament appears to connect the two objects. Figure 2.8 shows an old image of the configuration that was obtained by Chip Arp many decades ago. It does not show the luminous connection but hundreds of pictures that do can be found on the web. It does shows the unusual brightness of Markarian 205 and the spiral arms of Markarian 205. This source is so threatening that its discoverer was motivated to apologize for discovering it [587]. A number of papers have claimed that the connection does not exist and/or is not what it seems to be. All could and have been easily refuted – at least the existence part. The earlier claims are still accepted and the refutations ignored. It matters little what was said in those papers. In other words, the evidence is irrelevant unless it gives the right answer, because we

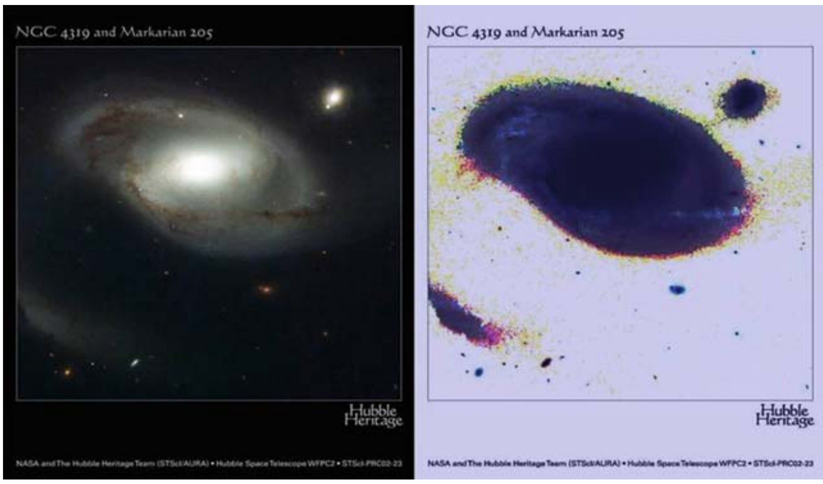

Fig. 2.8 Image of the discordant redshift configuration involving the spiral galaxy NGC4319 and higher redshift quasar Markarian 205. From NASA archive

now know so much that evidence alone has no meaning unless it fits into the existing paradigm. Lets assume that in this case this approach is OK. But what about other less blatant cases? What does it say about the way science is conducted? It serves no useful purpose to discuss in detail the different technical papers that have addressed the nature of the filament (see, e.g., [13, 266, 535, 587] or the film "Universe: The Cosmology Quest") as nothing has changed in the past 15 years.

A physical connection between two objects with significantly different redshift would produce a revolution in extragalactic astronomy. Baconians would say that this possibility, however remote, justifies careful study of discordant redshift pairs like the ones mentioned here. Careerists would argue that such configuration must be chance projections meriting no further study. It would be a "waste of telescope time" to study them. Our attempts to study NGC4319+Markarian 205 in more detail with HST were rejected. Perhaps it was thought that people like Arp and Sulentic would be too biased? Some HST time was subsequently given to an amateur astronomer (a high school teacher) who showed us the images he obtained with HST. They confirmed the connection. I warned him that he had got the wrong answer and would find the HST people reluctant to support publication. In fact, the (pre-COSTAR) results were never published. As mentioned earlier, the more recent post-COSTAR images confirm the luminous "connection".

So what is NGC4319+Markarian 205? Assuming the filament is real, and NOT a connection between discordant redshift objects, it is logically either related to "foreground" NGC4319 or "background" Markarian 205. NGC4319 is not alone. A large accordant redshift elliptical galaxy (NGC4291) lies only $\sim$6 arcmin away. This corresponds to a projected physical separation of only a few tens of kiloparsecs. Unless the two galaxies are much more widely separated along the line of sight than their redshifts suggest (they could show very similar redshift and still be separated by a megaparsec), they represent a close pair. NGC4319 does not show a typical Hubble or deVaucouleurs morphology (Fig. 2.8) and the unusual structure could be due to tidal interaction between NGC4291 and NGC4319. This does not, however, explain a high spatial frequency structure like the apparent connection. We observe no other similar features in NGC4319 pointing in random directions. The narrowness favors a tidal feature but at a larger distance and thus associated with Markarian 205. There is a faint object (compact galaxy?) close to Markarian 205 and it apparently shows the same redshift as Markarian 205 [530]. Thus Markarian 205 may not be alone and we know that gravitational interactions can produce tidal bridges and tails. The luminous connection might therefore have nothing to do with NGC 4319 and a lot to do with Markarian 205 and its like redshift neighbor. This interpretation might be testable if suitable telescope time were available. But would it not be a waste of telescope time to confirm what we already know to be the answer? Maybe observing time can be given to the unbiased people responsible for the HST press release, thus protecting astronomy from more biased interpretations.

The other configuration mentioned in the question involves NGC7603ab. This connection between two galaxies is, on many levels, a completely different type of association. The active quasar-like nucleus [289] lies in the assumed "parent" spiral galaxy ($z \sim 0.029$) and the companion involves a smaller early-type (S0 = lenticular) galaxy ($z \sim 0.056$) that shows only a stellar absorption line (i.e., stellar) spectrum. They appear to be connected by a filament that looks like a spiral arm of the parent galaxy. Figure 2.9 shows an image processed version of NGC7603ab using a 5 m Palomar plate obtained by H. Arp in the 1970s. I digitized the photographic image and displayed it in an unconventional way to assess the S/N properties of the lowest light levels, including the "bridge". The plate was scratched but the arm/connection is well seen. Two blobs in that arm/connection can also be seen as small dark spots. Twenty five years ago, we could not observe such faint objects spectroscopically – even with the Palomar 5 m – but I assumed they were HII regions in the spiral arm of NGC7603 and would likely show the same redshift as NGC7603. Arp was not so sure and he was correct! Recent new observations show that they have much higher redshifts ($z \sim 0.245$ and 0.394 [323]). Always a surprise when one looks more deeply with a larger telescope and new technologies.

In Fig. 2.9, the spiral arm/connection appears to terminate at the higher redshift galaxy, but even deeper images show that a much fainter arm extends beyond NGC7603b [511], lessening the impression that it is a connection. This result can

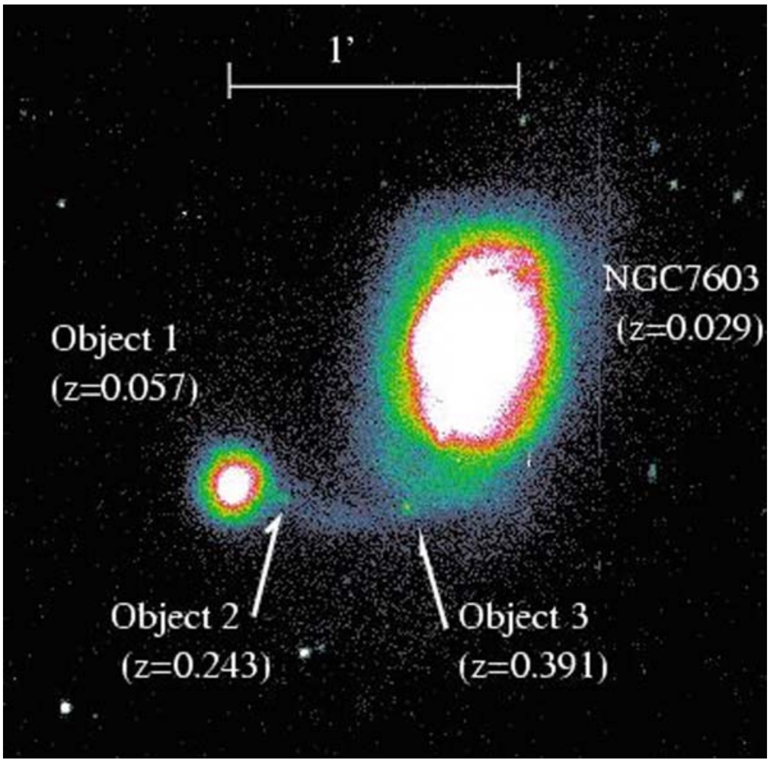

Fig. 2.9 Image of the discordant redshift pair NGC7603ab. The larger galaxy NGC7603a shows a quasar-like nucleus with strong broad emission lines, while the companion shows a higher-redshift stellar absorption line spectrum. Two even higher redshift emission line objects appear as *black dots* on the "bridge" connecting NGC7603ab. Credit [322]

be used to argue that the two galaxies are unconnected and unrelated. Perhaps a more powerful argument in favor of the connection involves the one-sidedness of the brighter feature interpreted as a spiral arm of NGC7603a. A spiral arm usually has a counterpart on the other side of the galaxy nucleus. Among the thousand brightest spiral galaxies on the sky, this striking asymmetry is almost unique. The next step for a Baconian (we lost the careerists two paragraphs ago) is to try and find another nearby galaxy that can be blamed for this configuration. It should show a redshift similar to NGC7603, allowing us to invoke a tidal interaction as the cause of the asymmetric spiral arm – conventionally NGC7603b and the two higher redshift blobs must be accidental projections along the same direction as that arm. My careerist side fully expected that another galaxy projected near NGC7603 would show a similar redshift. The field was recently included in the SDSS, providing redshift measures for all reasonable candidate galaxies. None of these galaxies show a redshift similar to NGC7603. The single asymmetric arm cannot be explained as a tidal feature because there is no suitable neighbor with whom NGC7603 could interact. A careerist will unblushingly say that the like redshift companion was eaten by NGC7603 just after it produced the asymmetry. Invoking things that we cannot see has become a cottage industry among the careerists (e.g., dark massive objects, dark matter, dark energy, dark galaxies). Naturally this invocation is distasteful to Baconians (after all why do we need observers if most of the Universe is invisible?) – but this explanation may be the correct one.

Moving away from two well-know examples, one can point out that there are now numerous statistical studies that show a clear excess of high redshift quasars in the vicinity of low redshift galaxies (e.g., [470, 534, 603]). The statistical results carry much more weight than any single connected pair of discordant redshift objects. These results are not much discussed unless the cause is attributed to “DM lensing”. Always something we cannot see. There is not much room for empiricism in modern astronomy but a wide berth for ideology. No more (or less?) far out would be the hypothesis that discordant redshift quasars are being ejected from the nuclei of lower redshift galaxies [13]. This interpretation involves two new ideas: ejection and discordant redshift compact objects of unknown nature. The ejected objects (quasars?) at the redshift distances of their parent galaxies would not be very luminous, similar to HII regions in the galaxies.

Actually the ejection idea is not so new – a “slingshot” ejection model having been proposed long ago [494]. An amusing recent discovery involves the best example of a (naked) quasar HE0450-2958 projected on the disk of a – fortunately like redshift – galaxy [332]. This discovery recently gave rise to a flurry of papers again, reviving a quasar ejection mechanism and most without crediting the pioneers of this idea (for an exception see [222]). Presumably, because the authors of the original paper showed their disloyalty to the standard paradigm by advocating the mechanism in connection with discordant redshift associations. Such treacherous acts are never forgiven. Have the respectable people who recently revised the quasar ejection hypothesis unwittingly opened Pandora’s box halfway? They accept that ejection can occur. Now all that is needed is a model to explain the physical nature of compact quasar ejecta with non-Doppler redshifts. No small order.

In summary, the famous cases involving apparent connected discordant objects deserve detailed study, no matter how certain we are that they are spurious. Our very certainty requires it. At the same time we have many studies showing an excess of higher redshift quasars near lower redshift galaxies. Unless magic (always unseen) effects can explain them, they represent a fundamental challenge to the standard paradigm. If they were not so controversial perhaps, they could have been better utilized to map the DM distribution. Alternatively, an ejection mechanism exists to explain how they got where they are. But no explanation for their discordant nature exists although scattering mechanisms mentioned earlier would be the place to look. Most of these quasars will show both broad and narrow emission lines with the same redshift, so it is clear that any of these potential non-Doppler redshift producing mechanisms must overcome a major challenge.

Thank you Jack.

After the treatment of the main change of paradigm from CDM to ΛCDM and of the use of astronomical candles for mapping the expansion of the Universe, we now move to sections dedicated to the main empirical cornerstones of the standard cosmological model. Having mentioned in previous sections the problem of element abundances as a result of stellar evolution, it is important to understand where the simplest elements come from in the early Universe. We then start with the interviews of Keith Olive and Gary Steigman on BBN, its observational tests, and its relation to the results coming from the analysis of CMB and LSS.

2.7 Cosmological Nucleosynthesis

2.7.1 Theory of Cosmological Nucleosynthesis

Dear Keith *(Olive)*, **cosmological nucleosynthesis is one of the main probe of the Standard Cosmological Model. Could you sketch the fundamental concepts and nuclear reactions involved in BBN theory?**

Element abundances offer several unique probes into physical processes throughout the history of the Universe. Indeed, one of the most fundamental questions in science relates to the chemical origins of the elements and their nuclear isotopes. By far, most of the natural elements are synthesized in stars, but a handful trace their origins back to the first few minutes of the Universe. In fact, BBN offers the deepest reliable probe of the early Universe, being based on well-understood Standard Model physics. Predictions of the abundances of the light elements, D, ^{3}He, ^{4}He, and ^{7}Li, synthesized shortly after the Big Bang are in good overall agreement with the primordial abundances inferred from observational data, thus validating the standard hot Big Bang cosmology (see [173, 382, 582]). This is particularly impressive, given that these abundances span nine orders of magnitude – from ^{4}He/H $\sim$ 0.08 down to ^{7}Li/H $\sim 10^{-10}$ (ratios by number).

All the heavier elements have been synthesized in stars. Abundance patterns and ratios also offer a unique glimpse into the chemical evolution of the Galaxy and Inter

Galactic Medium (IGM). Indeed abundance ratios, observed in systems of varying degrees of metallicity, allow one to trace the star formation history of the Universe, and can in principle determine the very nature of the first stars.

2.7.1.1 BBN Theory in Short

The Universe as described by the Big Bang theory began in an extremely hot, dense, and largely homogeneous state. Today, the Universe is nearly 14 billion years old, but at the time of the formation of the light elements, the Universe had existed only for minutes. Because the Universe cools as it expands, the early Universe was very hot and at the time of BBN, the temperature exceeded 10^{10} K. The density of neutrons and protons was about $10^{17}\,\mathrm{cm}^{-3}$ when nucleosynthesis began. Though small when compared to terrestrial densities, it was far larger than the average density of normal matter in the Universe today, $10^{-7}\,\mathrm{cm}^{-3}$. As will be described later, equilibrium processes governed the production of the light nuclei as the Universe cooled.

The nucleosynthesis chain begins with the formation of deuterium in the process $p + n \rightarrow \mathrm{D} + \gamma$. However, because of the large number of photons relative to nucleons (or baryons), $\eta_{\mathrm{B}}^{-1} = n_\gamma/n_{\mathrm{B}} \sim 10^{10}$, deuterium production is delayed past the point where the temperature has fallen below the deuterium binding energy, $E_{\mathrm{B}} = 2.2\,\mathrm{MeV}$ (the average photon energy in a BB is $\bar{E}_\gamma \simeq 2.7T$). This is because there are many photons in the exponential tail of the photon energy distribution with energies $E > E_{\mathrm{B}}$ despite the fact that the temperature or $\bar{E}_\gamma$ is less than E_{B}. The degree to which deuterium production is delayed can be found by comparing the qualitative expressions for the deuterium production and destruction rates,

$$\begin{aligned} \Gamma_p &\approx n_{\mathrm{B}}\sigma v \\ \Gamma_d &\approx n_\gamma \sigma v\, \mathrm{e}^{-E_{\mathrm{B}}/T} \end{aligned} \qquad (2.1)$$

When the quantity $\eta_{\mathrm{B}}^{-1}\exp(-E_{\mathrm{B}}/T) \sim 1$, the rate for deuterium destruction $(\mathrm{D} + \gamma \rightarrow p + n)$ finally falls below the deuterium production rate and the nuclear chain begins at a temperature $T \sim 0.1\,\mathrm{MeV}$.

In addition to the $p + n \rightarrow \mathrm{D} + \gamma$ reaction, the other major reactions leading to the production of the light elements tritium (T) and ^{3}He are

- $\mathrm{D} + \mathrm{D} \rightarrow p + \mathrm{T}$ $\quad$ $^3\mathrm{He} + n \rightarrow p + \mathrm{T}$,
- $\mathrm{D} + \mathrm{D} \rightarrow n + {}^3\mathrm{He}$ $\quad$ $\mathrm{D} + p \rightarrow \gamma + {}^3\mathrm{He}$.

Followed by the reactions producing ^{4}He

- $^3\mathrm{He} + \mathrm{D} \rightarrow p + {}^4\mathrm{He}$ $\quad$ $\mathrm{T} + \mathrm{D} \rightarrow n + {}^4\mathrm{He}$.

The gap at $A = 5$ is overcome and the production and destruction of mass $A = 7$ are regulated by

- $^3\mathrm{He} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Be}$, followed by the decay of ^{7}Be

- $^7\mathrm{Be} + e^- \rightarrow {}^7\mathrm{Li} + \nu_e$

as well as the reactions involving ^{7}Li directly

- $\mathrm{T} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Li}$ $\quad {}^7\mathrm{Be} + n \rightarrow p + {}^7\mathrm{Li}$ $\quad {}^7\mathrm{Li} + p \rightarrow {}^4\mathrm{He} + {}^4\mathrm{He}$.

The gap at $A = 8$ prevents the production of other isotopes in any significant quantity.

When nucleosynthesis begins, nearly all the surviving neutrons end up bound in the most stable light element ^{4}He. Heavier nuclei do not form in any significant quantity both because of the absence of stable nuclei with mass number 5 or 8 (which impedes nucleosynthesis via $^4\mathrm{He} + n$, $^4\mathrm{He} + p$, or $^4\mathrm{He} + {}^4\mathrm{He}$ reactions) and the large Coulomb barriers for reactions such as the $\mathrm{T} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Li}$ and $^3\mathrm{He} + {}^4\mathrm{He} \rightarrow \gamma + {}^7\mathrm{Be}$ reactions listed earlier. Hence the primordial mass fraction of ^{4}He, conventionally referred to as Y_p, can be estimated by the simple counting argument

$$Y_\mathrm{p} = \frac{2(n/p)}{1 + n/p} \simeq 0.25\,. \tag{2.2}$$

There is little sensitivity here to the actual nuclear reaction rates, which are important in determining the other "left-over" abundances: D and ^{3}He at the level of a few times 10^{-5} by number relative to H, and ^{7}Li/H at the level of about 10^{-10} (when $\eta_{10} \equiv 10^{10}\eta_B$ is in the range $1 \div 10$).

Historically, BBN as a theory explaining the observed element abundances was nearly abandoned due to its inability to explain *all* element abundances. Subsequently, stellar nucleosynthesis became the leading theory for element production [83]. However, two key questions persisted. (1) The abundance of ^{4}He as a function of metallicity is nearly flat and no abundances are observed to be below about 23%. In particular, even in systems in which an element such as oxygen, which traces stellar activity, is observed at extremely low values (compared with the solar value of O/H), the ^{4}He abundance is nearly constant. This is very different from all other element abundances, with the exception of Li. (2) Stellar sources cannot produce the observed abundance of D/H. Indeed, stars destroy deuterium and no astrophysical site is known for the production of significant amounts of deuterium. Thus we are led back to BBN for the origins of D, ^{3}He, ^{4}He, and ^{7}Li.

The resulting elemental abundances predicted by standard BBN are shown in Fig. 2.10 as a function of η_B [118]. The plot shows the abundance of ^{4}He by mass, Y, and the abundances of the other three isotopes by number. The bands indicate the central predictions from BBN, while their thickness corresponds to the uncertainty in the predicted abundances. The uncertainty range in ^{4}He reflects primarily the 1σ uncertainty in the neutron lifetime.

In the standard model with three neutrino flavors, the only free parameter is the density of baryons that sets the rates of the strong reactions. Thus, any abundance measurement determines η_B, while additional measurements overconstrain the theory and thereby provide a consistency check. BBN has thus historically been the premier means of determining the cosmic baryon density. With the increased

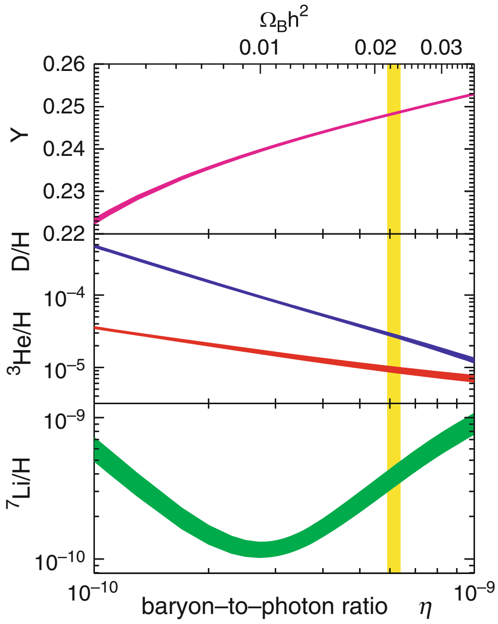

Fig. 2.10 The predictions of standard BBN. Primordial abundances as a function of the baryon-to-photon ratio η_B. Abundances are quantified as ratios to hydrogen, except for ^{4}He, which is given as a mass fraction. The bands give the 1σ uncertainties about the central values of the abundances as a function of η_B, reflecting the uncertainties in the nuclear and weak interaction rates. The vertical band shows the WMAP derived value of η_B. From [118]

precision of microwave background anisotropy measurements, it is now possible to use the CMB to independently determine the baryon density. The third year WMAP data implies [524]

$$\eta_{10} = 6.11 \pm 0.22. \tag{2.3}$$

Equivalently, this can be stated as the allowed range for the baryon mass density expressed today as a fraction of the critical density: $\Omega_B = \rho_B/\rho_{crit} \simeq \eta_{10} h^{-2}/274 = (0.0223 \pm 0.0008)h^{-2}$. This range in η_B is shown as a vertical strip in Fig. 2.10.

The promise of CMB precision measurements of the baryon density suggests a new approach in which the CMB baryon density becomes an input to BBN. Thus, within the context of the Standard Model (i.e., with $N_\nu = 3$), BBN becomes a zero-parameter theory, and the light element predictions are completely determined to be within the uncertainties in η_B and the BBN theoretical errors. Comparison with light element observations then can be used to restate the test of BBN–CMB consistency, or to turn the problem around and test the astrophysics of post-BBN light element evolution [119]. Alternatively, one can consider possible physics beyond the Standard Model (e.g., with $N_\nu \neq 3$) and then use all of the abundances to test such models.

Thank you Keith.

Dear Gary *(Steigman)*, **what are in your opinion the key aspects of the BBN theory and how are the primordial abundances of the light element isotopes**

(D, ^{3}He, ^{4}He, and ^{7}Li) connected with the ratio between the densities of baryons and photons and the expansion rate of the Universe?

The present Universe is observed to be expanding and to be filled with radiation, the CMB radiation whose spectrum is very precisely that of a BB at a temperature of 2.725 K. As the Universe expands, the average density of all its constituents decreases and the temperature of the CMB decreases as well. Conversely, in the past, the density and temperature were higher; the earlier the epoch in the evolution of the Universe, the hotter and denser were its constituents. The very early Universe was a hot, dense primordial soup of all the particles we know (from accelerator experiments and from the "standard model of particle physics"). The physical properties of the Universe, such as its rate of expansion and the temperature and density of its constituents can be tracked quantitatively using the standard Friedmann–Lamaître–Robertson–Walker (FLRW) cosmology, based on Einstein's theory of General Relativity (GR). According to this "standard cosmological model", when the Universe was only a fraction of a second old, the temperature corresponded to a thermal energy in excess of a few mega electron volt and the density was very high, so that collisions among the particles present at that time (neutrons, protons, electron-positron ($e^{\pm}$) pairs, neutrinos and photons – the blue-shifted CMB photons) were very rapid compared to the rate at which the Universe was expanding. It is during this epoch, from a fraction of a second to several minutes, that collisions among the neutrons and protons build the light elements, deuterium, helium-3, helium-4, and lithium-7 in primordial nucleosynthesis.

The calculation of the BBN-predicted primordial abundances involves following the nuclear and weak interactions as they interconvert neutrons and protons and, as they build neutrons and protons into more complex nuclei. For a more detailed description of this physics than is presented here, see my recent review [528]. The predicted abundances depend mainly on the density of nucleons, often called baryons (B). As the Universe is expanding, all densities evolve with time. A useful measure is in terms of the baryon density parameter $\eta_{\rm B}$, formed by the ratio of the number densities of baryon and CMB photons.

Aside from the extra photons produced when the $e^{\pm}$ pairs annihilate in the early Universe, this ratio remains constant as the Universe expands (and cools). The results of the BBN calculation in the standard model are shown as a function of $\eta_{10} = 10^{10}\eta_{\rm B}$ in Fig. 2.10, where the mass fraction of ^{4}He, Y, and the ratios of D, ^{3}He, and ^{7}Li to hydrogen (by number) are shown as a function of η_{10}.

The primordial abundances, especially that of ^{4}He, also depend on the early Universe expansion rate, as measured by the Hubble parameter, H, which, for the standard cosmology, depends on the square root of the energy density, ρ. As, at the time of BBN, the energy density is dominated by the contributions from relativistic particles (photons, $e^{\pm}$ pairs, light neutrinos), any deviation from the standard model, with three flavors of neutrinos (ν_e, ν_μ, ν_τ), may be parametrized by the expansion rate factor, $S \equiv H'/H = (\rho'/\rho)^{1/2}$ or, by the effective number of neutrinos, $\rm N_\nu \equiv 3 + \Delta N_\nu$,

$$S \equiv H'/H = (\rho'/\rho)^{1/2} = (1 + 7\Delta \mathrm{N}_\nu/43)^{1/2}. \qquad (2.4)$$

As the neutron-to-proton ratio is determined by the weak interactions involving neutrinos (electron type neutrinos and antineutrinos), any asymmetry in the abundance of neutrinos relative to antineutrinos (lepton asymmetry) can affect the BBN abundances. Once again, the ^{4}He abundance is especially sensitive to this deviation from the standard model. The parameter ξ_e provides a measure of the lepton asymmetry (in a similar manner, the baryon density parameter, $\eta_B \equiv 10^{-10}\eta_{10}$, measures the baryon asymmetry). However, if $\xi_e \approx \eta_B \approx 6 \times 10^{-10}$ (see below), there would be no measurable effect on BBN. Only for $|\xi_e| \gtrsim 0.001$ will a lepton asymmetry significantly modify the standard BBN-predicted primordial abundances of the light nuclides.

It is easy to understand the qualitative trends in Fig. 2.10 without having to delve deeply into the details of the standard BBN ($N_\nu = 3$ ($S = 1$), $\xi_e = 0$) calculation. Weak interactions among neutrons, protons, electrons, positrons, and neutrinos (electron neutrinos and antineutrinos) interconvert neutrons and protons. As the neutron is more massive than the proton, the proton is favored over the neutron and the ratio of neutrons to protons, by number, is always $\leq$1. At the time BBN begins in earnest, this ratio is $(n_n/n_p)_{BBN} \approx 1/7$. As the reactions leading to the production of ^{4}He, the most tightly bound of the light nuclei, are very fast compared to the universal expansion rate, virtually all the available neutrons are incorporated into ^{4}He. As a result, the ^{4}He abundance (by mass) is $Y = \frac{2n_n}{n_n+n_p} \approx 1/4$, very nearly independent of the baryon density as may be seen in Fig. 2.10.

As the ^{4}He abundance is so closely tied to the neutron-to-proton ratio at BBN, it is sensitive to the competition between the weak interaction rate and the universal expansion rate [281]. ^{4}He provides an early Universe chronometer.

$$\Delta Y = 0.16(S - 1) \approx 0.013\Delta N_\nu. \tag{2.5}$$

The primordial abundance of ^{4}He also probes a universal lepton asymmetry. For $\xi_e > 0$, there are more ν_e than $\bar{\nu}_e$, so that the abundance of neutrons relative to protons is reduced, reducing the BBN-predicted primordial abundance of ^{4}He [265, 281].

$$\Delta Y \approx -0.23\xi_e. \tag{2.6}$$

As may be seen from Fig. 2.10, Y is a very slowly varying function of the baryon (nucleon) density. For $\eta_{10} \approx 6$, $N_\nu \approx 3$ ($S \approx 1$), and $|\xi_e| \lesssim 0.1$, a very good fit to the BBN-predicted primordial abundance of ^{4}He is [281, 528]

$$Y_P = 0.2485 \pm 0.0006 + 0.0016(\eta_{10} - 6) + 0.013\Delta N_\nu - 0.23\xi_e. \tag{2.7}$$

In contrast to ^{4}He, as the less tightly bound nuclei of D and ^{3}He are being burned to produce ^{4}He, their relic abundances are sensitive to the baryon density at BBN with $(D/H)_P \propto \eta_{10}^{-1.6}$ and $(^3He/H)_P \propto \eta_{10}^{-0.6}$. D and ^{3}He are primordial baryometers. Notice that while ^{4}He, the second most abundant element to emerge from the early Universe, has a BBN-predicted abundance, by number, of order 10% of that of hydrogen, the abundances of D and ^{3}He are predicted to be smaller than that of hydrogen by some 4–5 orders of magnitude.

The only other nuclide produced in an astrophysically interesting abundance is ^{7}Li, whose abundance is smaller than those of D and ^{3}He by another five orders of magnitude. The reason for the very small abundance of ^{7}Li traces to the bottleneck at ^{4}He due to the gap at mass-5: there is no stable nucleus at mass-5. Coulomb-suppressed reactions of ^{4}He with the much less abundant D, ^{3}H, ^{3}He nuclei, guarantees that the primordial abundances of the heavier nuclides are strongly suppressed. Those few reactions that do jump the mass-5 gap lead to mass-7, producing ^{7}Be and ^{7}Li. Later in the evolution of the Universe, ^{7}Be captures an electron and decays to ^{7}Li, the only surviving mass-7 primordial nuclide.

2.7.1.2 Primordial Abundances

Inferring the primordial abundances of the light nuclides from present-day observations of a variety of astronomical objects (stars, H II regions (regions of ionized gas), neutral gas) involves a complex interplay between physics, astrophysics, and astronomy. At each step in the process, statistical errors as well as systematic uncertainties may arise. While the former may be reduced by acquiring large amounts of data, the latter are, by their very nature, difficult to quantify and more data may, or may not, lead to their reduction. These uncertainties must be kept in the forefront of any confrontation between theory (the BBN-predicted abundances) and observation (the observationally inferred primordial abundances). While the bad news, at present, is that limited data sets plague the determination of the primordial abundances of D, ^{3}He, and ^{7}Li, and systematic errors are a cause for concern in determining the primordial abundances of all the light nuclides. For a detailed discussion of current data and the problems and uncertainties associated with inferring the primordial abundances from the observational data, the reader is referred to [528] and references therein. Here, I summarize the results that emerge from that analysis.

2.7.1.3 Deuterium

Because of its simple post-BBN evolution and its notable dependence on the baryon density parameter, Deuterium is the baryometer of choice. As a result of its very weak binding, whenever gas containing D is cycled through stars, deuterium is destroyed. As a result, $(\mathrm{D/H})_\mathrm{P} \gtrsim (\mathrm{D/H})_\mathrm{OBS}$. The abundances of the "heavy" nuclei (the so-called "metals": C, N, O, ...) provide a measure of the amount of gas that has been processed through stars. In the limit of low metallicity, the observed deuterium abundance should provide an accurate probe of its primordial abundance. Deuterium is best observed by its absorption spectrum as light passes from a background light source (e.g., a QSO) through intervening, neutral gas. The data from seen lines of sight through high-redshift, low-metallicity QSO absorption line systems [277] leads to an estimate [404] of

$$y_\mathrm{DP} \equiv 10^5(\mathrm{D/H})_\mathrm{P} = 2.70^{+0.22}_{-0.20}. \tag{2.8}$$

For standard BBN, this abundance corresponds to a baryon density parameter $\eta_{10} = 6.0 \pm 0.4$. This BBN-inferred baryon density is in excellent agreement with that inferred from observations of the CMB temperature fluctuation spectrum [524], $\eta_{10} = 6.1 \pm 0.3$, which provide a measure of the universal baryon density some 4×10^5 years after BBN. The standard model is consistent with observations of relics from the Universe at a few minutes and a few hundred thousand years after the beginning of the expansion.

2.7.1.4 Helium-3

^{3}He is observed in Galactic H II regions via the emission from the spin-flip transition at 3.46 cm (the analog of the 21 cm line in hydrogen) from singly ionized ^{3}He. Unfortunately, these Galactic H II regions contain gas that has been processed through several generations of stars and the evolutionary correction required to infer the primordial abundance introduces model-dependent, systematic uncertainties into the inferred primordial value. Following the suggestion [23] that the ^{3}He abundance inferred from observations of the most metal-poor (least processed) Galactic H II regions provide an estimate of (or, an upper bound to) the primordial abundance,

$$y_3 \equiv 10^5 (^3\mathrm{He/H})_\mathrm{P} = 1.1 \pm 0.2. \tag{2.9}$$

For standard BBN, this ^{3}He abundance corresponds to a baryon density $\eta_{10} = 5.6^{+2.2}_{-1.4}$ which, while much more uncertain than that inferred from BBN and D, and from the CMB, is entirely consistent with each of them. Observations of D, ^{3}He, and the CMB provide consistent, independent support for the standard models of cosmology and particle physics.

2.7.1.5 Helium-4

As gas is cycled through generations of stars in post-BBN chemical evolution, the abundance of ^{4}He increases. To minimize the model-dependent evolutionary corrections, the most valuable data are provided by observations of low-metallicity, extra-galactic H II regions where the presence of ^{4}He is revealed via the emission lines produced when ionized helium (and hydrogen) recombines. The good news is that there is a database of some 90 such H II regions, which are useful for minimizing the statistical uncertainties in the inferred value of Y_P. The bad news is that the detailed analyses of the physics and astrophysics of the formation and radiative transfer of the recombination lines in such H II regions have many systematic uncertainties. As a result, current estimates of Y_p vary from $Y_\mathrm{p} = 0.243 \pm 0.001$ [254], inferred from the study of some 80 H II regions, to $Y_\mathrm{p} = 0.249 \pm 0.009$ [380], or $Y_\mathrm{p} = 0.250 \pm 0.004$ [186], or $Y_\mathrm{p} = 0.248 \pm 0.003$ [399] inferred from studies of many fewer H II regions, but with an eye to deal more carefully with systematic corrections and their uncertainties. Following a critical review of these recent results, I have suggested [528] that the current data and analyses are consistent with a primordial abundance

$$Y_P = 0.240 \pm 0.006, \tag{2.10}$$

and a robust upper bound to Y_p of

$$Y_P < 0.251 \pm 0.002. \tag{2.11}$$

Notice that for either the BBN-predicted baryon density found using deuterium [528] or the consistent value inferred from the CMB [524], the standard BBN-predicted abundance of ^{4}He is $Y_p = 0.249 \pm 0.001$, consistent, within the uncertainties, with the primordial abundance inferred from the observational data. However, if these other data are ignored, then the baryon density inferred from standard BBN and ^{4}He alone, $Y_p = 0.240 \pm 0.006$, would be much smaller, $\eta_{10} = 2.8^{+2.0}_{-1.0}$ [528], hinting at a tension between D (and ^{3}He) and ^{4}He (and between the CMB and ^{4}He).

2.7.1.6 Lithium-7

Observations of lithium in the Sun and in the local interstellar gas are of little value in inferring the primordial abundance of ^{7}Li due to the large and highly uncertain evolutionary corrections required to connect the observationally inferred abundances to the BBN-predicted abundance. The only data of value available for inferring the primordial abundance are provided by observations of lithium on the surfaces of the very oldest, most metal-poor stars in the Galaxy. Even these data may need to be corrected for the post-BBN evolution of ^{7}Li [15, 481]. And, it is clear that there are processes at work in these stars which, over their long lifetimes, may have modified their surface abundances via depletion, dilution, or gravitational settling. Ignoring these latter corrections, a primordial abundance [15]

$$[\mathrm{Li}]_P \equiv 12 + \log\,(\mathrm{Li/H})_P = 2.1 \pm 0.1 \tag{2.12}$$

is inferred [15, 481]. In contrast, using the BBN or CMB inferred baryon density, the primordial abundance is predicted to be $[\mathrm{Li}]_P = 2.63^{+0.07}_{-0.08}$ (D + BBN) or, $2.65^{+0.05}_{-0.06}$ (CMB + BBN) which, while consistent with each other, differ from the above estimate by a factor of ~ 3 or more. On the basis of lithium alone, the standard BBN-predicted baryon density would be $\eta_{10} = 4.0 \pm 0.6$ [528].

However, in an attempt to estimate the correction to the observed lithium abundance due to gravitational settling, observations of stars in a Globular Cluster (of the same age and metallicity) have led to a (model-dependent) higher estimate [292] of

$$[\mathrm{Li}]_P = 2.54 \pm 0.10 \tag{2.13}$$

which, within the errors, is entirely consistent with the standard BBN prediction.

Thank you Gary. Now Keith will enter more deeply into the observational tests of the BBN.

2.7.2 Tests of Cosmological Nucleosynthesis

Dear Keith *(Olive)*, in which way observations of element abundances can probe the BBN theory? Could you explain why, and what are the most significant tests used in present day cosmology?

Unfortunately, we can not observe element abundances directly at the time of BBN. Abundances are observed at much later epochs, after stellar nucleosynthesis has commenced. The ejected remains of this stellar processing can alter the light element abundances from their primordial values, and also produce heavy elements such as C, N, O, and Fe ("metals"). Thus one seeks astrophysical sites with low metal abundances, to measure light element abundances that are closer to primordial. For all of the light elements, systematic errors are an important and often dominant limitation to the precision of the primordial abundances.

In recent years, high-resolution spectra have revealed the presence of D in high-redshift, low-metallicity quasar absorption systems (QAS), via its isotope-shifted Lyman-α absorption. These are the first measurements of light element abundances at cosmological distances. The six most precise observations of deuterium ([383] and references therein) in QAS give D/H = $(2.83 \pm 0.26) \times 10^{-5}$, where the error is statistical only. These measurements are clearly consistent with the CMB/BBN determined value of the primordial D/H abundance, which is predicted to be

$$(\mathrm{D/H})_p = 2.6 \pm 0.2 \times 10^{-5}. \tag{2.14}$$

^{4}He is observed in clouds of ionized hydrogen (HII regions), the most metal-poor of which are in dwarf galaxies. There is now a large body of data on ^{4}He and carbon, nitrogen, and oxygen (CNO) in these systems ([254] and references therein). The He abundance from this sample of 89 HII regions obtained $Y_\mathrm{p} = 0.2429 \pm 0.0009$ [254]. However, the recommended value is based on the much smaller subset of 7 HII regions, finding $Y_\mathrm{p} = 0.2421 \pm 0.0021$.

^{4}He abundance determinations depend on a number of physical parameters associated with the HII region in addition to the overall intensity of the He emission line. These include the temperature, electron density, optical depth, and degree of underlying absorption. A self-consistent analysis may use multiple ^{4}He emission lines to determine the He abundance, the electron density, and the optical depth. The question of systematic uncertainties was addressed in some detail in [379]. It was shown that there exist severe degeneracies inherent in the self-consistent method, particularly when the effects of underlying absorption are taken into account. The results of a Monte Carlo reanalysis [380] of NCG 346 [397, 398] showed that solutions with no absorption and high density are often indistinguishable (i.e., in a statistical sense they are equally well represented by the data) from solutions with underlying absorption and a lower density. In the latter case, the He abundance is systematically higher. These degeneracies are markedly apparent when the data is analyzed using Monte Carlo methods, which generate statistically viable representations of the observations. When this is done, not only are the He abundances found to be higher, but the uncertainties are also found to be significantly larger than in a direct self-consistent approach. The extrapolated ^{4}He abundance was determined to be $Y_\mathrm{p} = 0.2495 \pm 0.0092$. The value of η_B corresponding to this abundance is $\eta_{10} = 6.9^{+11.8}_{-4.0}$ and clearly overlaps with η_CMB. Conservatively, it would be difficult at this time to exclude any value of Y_p inside the range 0.232 – 0.258.

The systems best suited for Li observations are metal-poor Pop II stars in the spheroid of our Galaxy. Observations have long shown [525] that Li does not vary significantly in Pop II stars with metallicities $\lesssim 1/30$ of solar – the "Spite plateau". Recent precision data suggest a small but significant correlation between Li and Fe [480], which can be understood as the result of Li production from Galactic cosmic rays [172,571]. Extrapolating to zero metallicity, one arrives at a primordial value [481] $\mathrm{Li/H_p} = (1.23 \pm 0.06) \times 10^{-10}$.

The ^{7}Li abundance based on the WMAP baryon density is predicted to be

$$^7\mathrm{Li/H} = 4.3 \pm 0.7 \times 10^{-10}. \tag{2.15}$$

This value is in contradiction with most estimates of the primordial Li abundance. It is a factor of $\sim$3 higher than the value observed in most halo stars, and just about 0.2 dex over the value observed in globular clusters, ^{7}Li/H = $(2.2 \pm 0.3) \times 10^{-10}$ [57,58]. There are many possible sources for this discrepancy. Among them lie the possibility that some Li was destroyed or removed from the stellar surface. However, the lack of dispersion in the Li data limits the degree to which depletion can be effective. A very real systematic uncertainty stems from the assumed physical properties of the star that are used to derive the abundance from raw observations. Most important among these is the surface temperature of stars [174,353]. Nonstandard process may also play a role.

Detailed abundance observations of heavier elements play a key role in building a picture of the chemical history of the Universe. Indeed, chemical evolution at high redshift connects massive star formation in the early Universe to the epoch of reionization, the heavy element abundances of the oldest stars and the high redshift IGM, and the mass outflows associated with galaxy formation. Furthermore, predicted SN rates provide us with an independent probe of the early epoch of star formation. Combining abundance and SN rate predictions allows us to develop an improved understanding of both the cosmic star formation history and of the enrichment of the IGM, as well as to elucidate the nature of Population III ([120], see review of [102]).

Thank you Keith.

Dear Gary *(Steigman)*, how do the abundances of the light elements predicted by BBN compares with the information from CMB and LSS?

In Fig. 2.11 are shown the standard BBN-predicted values of η_{10}, inferred from a comparison with the observationally inferred abundances adopted before, along with their 2σ ranges. From deuterium alone we find $(\eta_{10})_\mathrm{D} = 6.0 \pm 0.6$ (at 2σ).

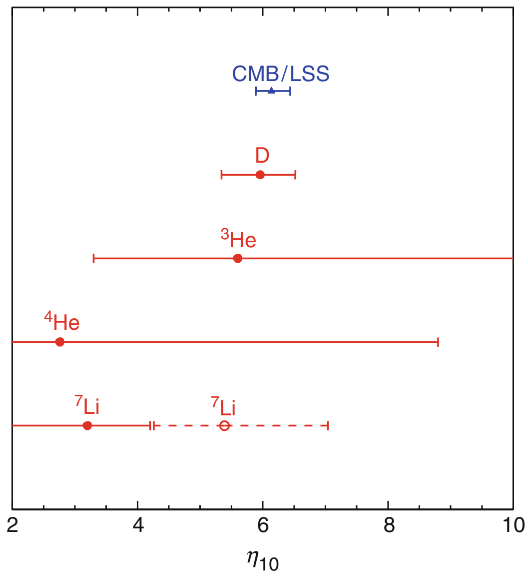

Fig. 2.11 The standard BBN-predicted values of η_{10}, along with their 2σ ranges, corresponding to the adopted primordial abundances (*filled circles*), along with the value inferred from cosmic background radiation and large scale structure data (CMB/LSS: *filled triangle*). The *open circle* and *dashed lines* correspond to the alternate lithium abundance discussed in the text

This value is in excellent agreement with that inferred from the more uncertain abundance of ^{3}He. While the central value of the ^{4}He abundance corresponds to a very different – much smaller – baryon density, as may be seen from Fig. 2.11, it is in agreement with the D, ^{3}He, and the CMB/LSS values at less than 2σ. However, even at 2σ, the lithium abundance adopted in (2.12) [15, 481] is inconsistent with these baryon density determinations, although the higher value corresponding to (2.13) [292] does agree with them. As the range of the density parameter covered in Fig. 2.11 is larger than the range of applicability of the analytic fit described earlier in (2.7), the specific values shown there are derived from a numerical BBN code. For the central value of the deuterium-predicted baryon abundance, the standard BBN-predicted helium abundance is $Y_p = 0.249$, only 1.5σ away from the observationally inferred value $Y_p = 0.240 \pm 0.006$. The lithium abundance poses a greater challenge; for $\eta_{10} = 6.0$, the standard BBN-predicted lithium abundance exceeds $[\mathrm{Li}]_P = 2.6$, which is far from the observationally inferred value of $[\mathrm{Li}]_P = 2.1 \pm 0.1$.

It is noteworthy that the two nuclides that may pose the most serious challenges to standard BBN (^{4}He and ^{7}Li) are those for which systematic corrections, and their corresponding uncertainties, have the potential to change the observationally inferred relic abundances by the largest amounts. The values of η_{10} corresponding to the alternative choices for ^{4}He and ^{7}Li considered above in (2.11) and (2.13) are in much better agreement with the D and ^{3}He determined baryon abundance: $(\eta_{10})_{\mathrm{He}} < 7.8^{+1.9}_{-1.5}$ and $(\eta_{10})_{\mathrm{Li}} = 5.4 \pm 0.6$, respectively.

Observations of the small temperature fluctuations in the cosmic background radiation and of the LSS they seeded currently provide the tightest constraint on the universal abundance of baryons $\eta_{\rm CMB/LSS} = 6.1 \pm 0.2$ [524], as shown in Fig. 2.11. The observationally inferred relic abundances of D and ^{3}He are in excellent agreement with the standard BBN predictions for this value/range of η_{10}. Depending on the outcome of various systematic corrections for the observationally inferred primordial abundances of ^{4}He and ^{7}Li, they may, or may not, pose challenges to standard BBN. If the tension between D and ^{4}He is taken seriously, it could be a sign of "new physics": $N_\nu \neq 3$ and/or $\xi_e \neq 0$. For example, for $\eta_{10} = 5.7$ and $N_\nu = 2.4$, the BBN-predicted primordial abundances of D and ^{4}He are now in perfect agreement with those inferred from the observational data and, also with that of ^{3}He [528]. However, the BBN-predicted lithium abundance remains very close to $[\mathrm{Li}]_{\rm P} \approx 2.6$, still a factor of $\sim$3 higher than that inferred from the observations. The same is true for the $\{\eta_{10}, \xi_e\} = \{6.0, 0.034\}$ pair [528].

2.7.2.1 At a Glance

According to the standard model of cosmology, the early Universe was hot and dense and, when it was a few minutes old, nuclear reactions among neutrons and protons synthesized astrophysically interesting abundances of the light elements D, ^{3}He, ^{4}He, and ^{7}Li. In the standard model, these relic abundances depend on only one free parameter, the baryon abundance (the baryon to photon ratio). Self-consistency of the standard model requires that there is a unique baryon abundance, consistent with the observationally inferred primordial abundances of these light elements. For D and ^{3}He, this is the case. As a bonus, this BBN-predicted baryon abundance is in excellent agreement with that inferred from the CMB/LSS. While the BBN-predicted abundance of ^{4}He may be somewhat higher than its observationally inferred value, within the observational errors, there is agreement. Three for the price of one; four, counting the CMB/LSS! However, the BBN-predicted relic abundance of ^{7}Li is a factor of three, or more, higher than its observationally inferred value. While this may provide a challenge to the standard model, it is not unlikely that the resolution of this challenge lies in the uncertain stellar physics associated with the evolution of the surface abundances of the oldest, most metal-poor stars in the Galaxy.

How may the observational verification of primeval nucleosynthesis be affected by post-BBN stellar nucleosynthesis and galactic evolution? Is this contribution known with the accuracy necessary for a robust verification of the cosmological model?

An essential, unavoidable step in comparing the predictions of primordial nucleosynthesis with the observational data is accounting for the chemical evolution of material that has been cycled through stars in the $\sim$14 Gyr since BBN was completed. Account for post-BBN evolution is not separate from, but is a crucial part of the analysis that leads us from the observational data to the inferred, relic abundances.

The post-BBN evolution of deuterium is straightforward, as whenever gas is cycled through stars, deuterium is entirely destroyed. Because of the very small binding energy of the deuteron, whenever deuterium is formed by nuclear reactions in the hot interiors of stars, it is immediately burned to tritium, helium-3, helium-4, and beyond. As a result, as the abundance of deuterium can only have decreased since BBN, *any* deuterium observed *anywhere* in the Universe, at *any* time in its evolution, provides a *lower* bound to the primordial abundance of D. For the same reason, by concentrating on those astrophysical objects (e.g., QSO Absorption Line systems) at high redshift and low metallicity, we can expect to measure an abundance nearly identical with the primordial value.

The post-BBN evolution of ^{3}He is much more complicated than that of D, because when gas containing ^{3}He is cycled through stars, some of the ^{3}He is burned away, some is preserved (not all layers of all stars are hot enough to burn ^{3}He) and, for some stars, new ^{3}He is produced. The result is that the extrapolation of the current data from chemically evolved H II regions in the Galaxy back to the early Universe is uncertain and model-dependent. It is for this reason that ^{3}He is usually given less weight in the comparison between theory and observation. Nonetheless, given the observed abundance of ^{3}He and our best estimate of its chemical evolution, theory and observation are in excellent agreement (see Fig. 2.11).

Stars burn hydrogen to helium (^{4}He). The abundance of ^{4}He observed in the post-BBN Universe has increased from its primordial value. Stars also synthesize the heavier nuclei, "metals" such as C, N, O, ... As the heavy element abundance (metallicity) increases in the course of stellar and galactic evolution, so, too, does the abundance of ^{4}He. Two options are available for accounting for – or avoiding – this inevitable correction required to pass from the observational data to the primordial abundance. One choice is to restrict attention to the very lowest metallicity regions observed and to *assume* that this will minimize the correction for post-BBN production of ^{4}He. The other is to use data from regions of all metallicities and to *extrapolate* to zero metallicity to find the BBN abundance. Each of these approaches has assets and liabilities and, each introduces its own uncertainties into the value of Y_p inferred from observations. These corrections, along with their attendant uncertainties, have been used to infer the primordial abundance of ^{4}He listed earlier.

Lithium is observed in the very most metal-poor, oldest stars in the Galaxy, stars with heavy element abundances lower than those in the Sun by factors of a thousand or more. It is expected that for these stars the observed lithium is completely dominated by the relic component from primordial nucleosynthesis. The problem, as discussed earlier, is that these oldest stars in the Galaxy have had the most time to modify their surface material, the material which is observed to infer the stellar lithium abundance. Post-BBN evolution of ^{7}Li is the least of our worries, compared to the uncertainties of stellar structure and evolution, in using the data to infer the ^{7}Li primordial abundance.

The bottom line is that for the two nuclides, D and ^{4}He, which are most valuable in testing the standard model, the post-BBN evolution is likely well-enough understood so that our conclusions are robust.

2.7.3 Alternatives to Standard BBN

Dear Gary *(Steigman)*, **do there exist reliable modifications and alternatives to the standard nucleosynthesis theory? Can the observational data be explained by alternative ideas?**

Alternatives to the standard model are limited only by the creativity and imagination of physicists and cosmologists. BBN is a pillar of modern cosmology, in that the first test any alternative theory must pass is that the correspondingly modified BBN needs to be in agreement with the observationally inferred primordial abundances presented earlier. Many new alternative theories never see the light of day because they fail this test. Nonetheless, there are still large classes of alternative theories that *may* be consistent with the predictions of BBN in the standard model and, therefore, consistent with the observational data.

For example, it could be that the early-Universe expansion rate, H, is modified compared to that in the standard model, as quantified by the expansion rate parameter, S, defined in (2.4) (or, by ΔN_ν). Although the standard value of $S = 1$ ($\Delta N_\nu = 0$, $N_\nu = 3$) is, within the uncertainties, consistent with the abundances, models of new physics or cosmology with nonstandard values are restricted to the range $1.6 \lesssim N_\nu \lesssim 3.3$. Models with N_ν outside this range are excluded.

Or, because of nonstandard physics, it could be that the lepton asymmetry of the Universe exceeds the baryon asymmetry by some nine orders of magnitude ($|\xi_e| \sim 0.1$). Again the effect of such an asymmetry would be to change the neutron-to-proton ratio at BBN and, therefore, to modify the BBN abundance of ^{4}He. This is allowed only for the asymmetry parameter restricted to the narrow range, $-0.027 \lesssim \xi_e \lesssim +0.086$. Models with ξ_e outside this range are excluded.

Thanks a lot Gary.
Together with the BBN, CMB studies provided up to now the stronger evidence in favor of the current cosmological scenario. Most of them come from the COBE mission. We now have the opportunity of speaking with the Nobel Laureate John Mather, who won the Nobel Laureate together with George Smoot for the fundamental cosmological results of the COBE mission. Here, we will ask him to review the characteristics of COBE and the importance of its discoveries.

2.8 CMB Observations and Main Implications

2.8.1 The COBE Legacy

Dear John *(Mather)*, **COBE opened the so-called era of precision cosmology with the up-to-now best measure of the CMB spectrum and discovered the CMB large scale anisotropy. Can you tell us about the scientific adventure of COBE? Why such a project has been so relevant in the context of physical cosmology in the beginnings of 1990?**

Springer

http://www.springer.com/978-3-642-00791-0

Questions of Modern Cosmology
Galileo's Legacy
D'Onofrio, M.; Burigana, C. (Eds.)
2009, XXX, 530 p., Hardcover
ISBN: 978-3-642-00791-0