17

GUSTAV A. TAMMANN*

The cosmological constants

The evolution of the large scale of the Universe is but one chapter of cosmology, yet it is fundamental to all of its aspects. The understanding of the large-scale evolution has once been described as the "Search for two numbers" (Sandage 1970). The first of these numbers is the Hubble constant H_0 which measures the present expansion rate. The second number was thought to account for the gravitational deceleration of the Universe. But since evidence has been found for an additional acceleration, the original deceleration parameter has been split into a decelerating term Ω_m and an accelerating term Ω_Λ. Ω_m is a measure of the gravitating mean matter density of the Universe, while Ω_Λ corresponds to the energy density of Einstein's "cosmological constant" Λ.

The three quantities H_0, Ω_m, and Ω_Λ fix the expansion age T of the Universe and determine also whether the Universe will eventually recollapse or expand forever.

In principle the values of H_0, Ω_m, and Ω_Λ can be measured in the present Universe as well as the age t of the oldest objects which – as an essential test – must be smaller than T.

Yet the cosmological constants are elusive. Even the concerted effort of many of the largest optical and radio telescopes has not led to their satisfactory determination. As a consequence observations from space were added wherever possible. In fact the determination of the cosmological constants was a driving force for some astronomical satellites.

Several satellites will be mentioned in the following which have contributed to our dawning understanding of the cosmological parameters, but none has been as important as the Hubble Space Telescope (HST).

The initial motivation for astronomers to go into space were the wavelengths inaccessible from the ground. The reward was overwhelming with the detection of X-ray and γ-ray sources in the early 1970s (see Chapter 12) and the subsequent results in ultraviolet spectroscopy (see Chapter 13). But optical observations above the atmosphere were also of fundamental interest offering almost 10 times smaller seeing disks and consequently the detection of roughly 100 times fainter objects than from the ground. Individual stars in crowded regions would become accessible and much greater detail would be visible in extended objects, as first pointed out by Hermann Oberth in 1923. However, this enormous gain was offset in part by the size limitations of a telescope in space. Plans for what was to become the 2.4 m HST for imaging and spectroscopy in the optical and near UV developed during the 1960s and were finally approved by the US Congress in 1977. ESA joined the venture with a 15% share, which turned out to be of great benefit to astronomy in Europe. The start of HST, planned for 1983, was delayed by the Space Shuttle accident until 1990. Since then, and particularly since 1993, that is, after the repair mission for the defective mirror, HST has not only brought spectacular progress in astronomical research, but has also stirred an enormous public interest in astronomy.

When plans were made in 1979 (Macchetto *et al*. 1979; Longair and Warner 1979) for future research with HST, the author was ambivalent about its benefits for the determination of H_0, being convinced that the main uncertainties of the time were due to statistical problems (Malmquist bias; Section 1.3). He was wrong by not foreseeing that standard supernovae of type Ia are so powerful standard candles that they can beat all statistical selection effects, and that HST can provide Cepheid distances to their nearest representatives and thus calibrate their true luminosity. It was clear, however, already in 1979 that the same supernovae can be observed with HST at cosmologically significant redshifts and that this is the route to determine the cosmological parameters q_0 (or Ω_m) and Λ.

* Universität Basel, Switzerland

Gustav A. Tammann, The Century of Space Science, 373–397

1 THE HUBBLE CONSTANT H_0

Since the Big Bang space has expanded. During the early stages the expansion was governed by complex processes, but after an inflationary period the Universe turned into a constant growth rate carrying with it all matter and the emerging galaxies. As a consequence any observer in the Universe has the impression that all galaxies are being carried away, and the faster they move the larger the distance. Yet strictly speaking the distant galaxies do not recede with any velocity, but space between the observer and the galaxy is stretched. In spite of this the spectra of nearby galaxies are interpreted as if they had recession velocities. This is permissible here because the determination of the *present* expansion rate, that is, the Hubble constant H_0, must be carried out at small distances where the look-back times are short.

If light travels through an expanding space its waves will be stretched and when it arrives at the observer it will appear redder than at the time of emission. Also the spectral lines of the galaxies will be shifted to the red. These lines can be ascribed to certain neutral or ionized elements and each appears in the laboratory at a fixed wavelength λ_0. If that particular line appears in the galaxy spectrum at wavelength λ one defines a redshift z of the galaxy as $z \equiv (\lambda - \lambda_0)/\lambda_0$. The redshift is determined by the stretch factor the Universe has experienced during the light travel time, that is, $z \equiv (R_0/R_e) - 1$, where R_0 is the present radius of the Universe (more exactly its radius of curvature) and R_e is the radius at the time of emission. As long as z is small, say $z \lesssim 0.2$, the redshift can also be interpreted with sufficient accuracy as a Doppler effect. The recession velocity v is in that case simply given by $v = cz$, where c is the velocity of light (in $\mathrm{km\,s^{-1}}$).

To pin down the expansion rate of the Universe one introduces the Hubble constant H_0 which is defined as

$$H_0 \equiv \dot{R}_0/R_0 = v/r_0 \ (v \ll c) \tag{1}$$

where v is the "recession velocity" of a galaxy and r_0 is its distance (in Mpc). Thus the units of H_0 are $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (the units are not repeated in the following), but its dimension is simply $(\mathrm{time})^{-1}$.

It should be noted that H_0 is constant in the sense that it is the same anywhere in the present Universe, but it has a different value at other times. This is obvious because the distance r is constantly increased by the expansion of the Universe and consequently the Hubble parameter H must decrease with time.

Equation (1) gives the false impression that the determination of the Hubble constant H_0 is simple. One just determines the redshift z of the spectrum of any given galaxy, and hence its recession velocity v, and divides the latter by the galaxy's distance. Indeed the determination of the redshift poses no problem. The crux, however, is introduced by peculiar motions which are superposed on the regular Hubble flow. They are caused by the gravitational pull of galaxies and clusters of galaxies and of any additional dark matter. For instance our Local Group of galaxies is a sufficient density enhancement to have turned its original expansion into a contraction. Our nearest neighbor, the Andromeda galaxy, M31, has thus a *blueshift* and approaches us at about $100\,\mathrm{km\,s^{-1}}$. The largest peculiar velocity we partake of is a $630\,\mathrm{km\,s^{-1}}$ velocity which manifests itself as a dipole anisotropy of the cosmic microwave background (CMB, see below).

The peculiar velocities cause a dilemma for the determination of H_0. On the one hand one wants to determine it locally to find the *present* value H_0; on the other hand one must determine it at sufficiently large distances, where the recession velocities are significantly larger than any possible peculiar motions. The best compromise is to determine H_0 between recession velocities of 10 000 and $30\,000\,\mathrm{km\,s^{-1}}$. But the determination of the corresponding distances poses formidable problems.

With this requirement in mind a new strategy has developed to determine H_0. The idea is to map the Hubble flow out to the necessary distances by means of *relative* distances of a set of galaxies, for instance from "standard candles" of constant luminosity, and then to determine the true distances of one or a few nearby galaxies of the same set. The latter step has been greatly facilitated by HST. The route is described below.

1.1 The Hubble diagram of standard candles

The expansion of the Universe implies that the recession velocities v increase linearly with the photometric distance. Astronomers measure, in general, photometric distances in terms of the distance modulus $(m - M)$, which is defined as $(m - M) \equiv 5\log\, r_{\mathrm{Mpc}} + 25$, where m is the apparent magnitude of an object and M its absolute magnitude, that is, the apparent magnitude the object would have if seen from a distance of 10 pc. Objects with (nearly) constant absolute magnitude are called "standard candles." The apparent magnitudes m of standard candles are a measure of their *relative* distances, or more exactly of their $\log r$ values.

A plot of $\log v$ *v.* the apparent magnitude m of extragalactic standard candles constitutes the so-called "Hubble diagram." The standard candles concentrate in a linearly expanding Universe along a straight line of slope 0.2. Systematic deviations from this line beyond $\sim 30\,000\,\mathrm{km\,s^{-1}}$ are caused by the deceleration or acceleration of the Universe (provided that the apparent magnitudes are corrected for the z-dependent photometric K-term).

Since Hubble and Humason (1931), the Hubble diagram has provided the fundamental evidence for the linear expansion of the Universe. The ever-increasing persuasive power of the Hubble diagram from Humason *et al.* (1956) to Sandage (1966) and Kristian *et al.* (1978), who used first-ranked cluster galaxies as standard candles, has shattered all attempts to interpret redshifts as a non-cosmological effect.

1.1.1 The Hubble diagram of supernovae of type Ia

Following an early suggestion of Kowal (1968) it has become increasingly clear that supernovae of type Ia (SNe Ia) are the most powerful standard candles to date. This can be physically rationalized because a SN Ia outburst is caused each time a white dwarf is pushed over its Chandrasekhar limit by a mass-shedding companion.

For 35 SNe Ia in the distance range $1200 < v \lesssim 30\,000$ km s^{-1} good B and V magnitudes at maximum light are known, for 29 of them also $I_{\rm max}$ magnitudes. The SNe Ia, listed by Parodi *et al.* (2000), are selected to have $B_{\rm max} - V_{\rm max} \lesssim 0.06$, after correction for galactic reddening to guard against absorption within their parent galaxies and to exclude a few intrinsicly red objects with peculiar spectra. Two blue SNe Ia with peculiar spectra at early phases (SN 1991 T, 1995 ac) are also excluded. The selected objects of the sample all have normal spectra as far as is known (Branch *et al.* 1993).

The 35 SNe Ia define Hubble diagrams in B, V, and I with a very small scatter of $\sigma_{\rm M} \lesssim 0^{\rm m}.2$. This means that SNe Ia are standard candles to better than $\pm 20\%$ of their luminosity. In spite of this the residuals about the Hubble line correlate with second parameters. The SNe Ia with slow decline rates Δm_{15} are somewhat brighter, where Δm_{15} is the magnitude decline during the first 15 days after B maximum. The relation between magnitude residual and decline rate is roughly $\delta m \propto 0.45 \times \Delta m_{15}$. Also the intrinsically bluer SNe Ia are somewhat brighter than their redder counterparts by about $\delta m \propto p \times \Delta(B-V)$, where p is $\sim$1.2–2.5 for I, V, and B magnitudes, respectively (for the exact relations, see Parodi *et al.* 2000). Corrected magnitudes $m^{\rm corr}_{B,V,I}$ are obtained by reducing all SNe Ia to the mean decline rate $\langle \Delta m_{15} \rangle = 1.2$ and to the mean color $\langle B-V \rangle = -0^{\rm m}.01$.

The corrected magnitudes $m^{\rm corr}_{B,V,I}$ define Hubble diagrams as shown in Figure 1. *Their tightness is astounding*. The Cerro Tololo collaboration, to whom one owes 70% of the photometry of the fiducial sample, quote a mean observational error of their $m_{\rm max}$ values of $\sim 0^{\rm m}.10$ and of their colors $(B-V)$ of $\sim 0^{\rm m}.05$ (Hamuy *et al.* 1996). This alone would suffice to explain the observed scatter of $\sigma_m = 0^{\rm m}.12 - 0^{\rm m}.13$. An additional error source are the corrections for galactic absorption which were adopted from Schlegel *et al.* (1998). In fact, if one excludes the nine SNe Ia with large galactic absorption corrections ($A_V > 0^{\rm m}.2$) the scatter decreases to $0^{\rm m}.11$ in all three colors. Two important conclusions follow from this. (1) If the total observed scatter of the Hubble diagrams is read vertically as an effect of peculiar motions, a generous upper limit is set of $\Delta v/v = 0.05$, which holds for the range of $3500 \lesssim v \lesssim 30\,000$ km s^{-1}. The (all-sky) distance-dependent variation

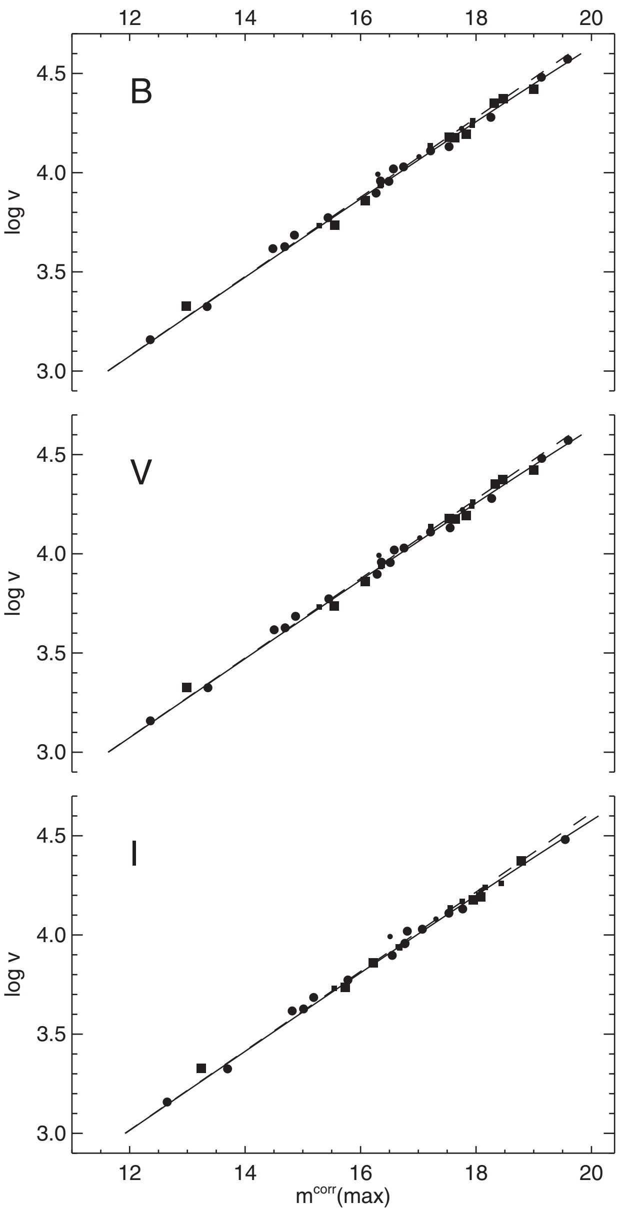

Figure 1 The Hubble diagrams in B, V (and I) for the 35 (29) SNe Ia of the fiducial sample with magnitudes $m^{\rm corr}$, that is, corrected for decline rate Δm_{15} and color $(B-V)$. Circles are SNe Ia in spirals, squares in E/S0 galaxies. Small symbols are SNe Ia whose observations begin eight days after B maximum or later. Solid lines are fits to the data assuming a flat universe with $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$; dashed lines are linear fits with a forced slope of 0.2 (corresponding approximately to $\Omega_{\rm m} = 1.0$ and $\Omega_\Lambda = 0.0$). (Parodi *et al.* 2000.)

of H_0 must be even smaller. (2) If, however, the scatter is read horizontally and if allowance is made for the observational errors of the apparent magnitudes (and for any peculiar motions) one must conclude that the luminosity scatter of blue SNe Ia, once they are homogenized in Δm_{15} and color, is smaller than can be measured at present. In other words, they are extremely powerful standard candles.

Linear regressions, which correspond closely to a Universe with $\Omega_{\rm m} = 1$, $\Omega_\Lambda = 0$, to the data in Figure 1 yield

$$\log v = 0.2 m_\lambda^{\rm corr} + c_\lambda \qquad (2)$$

with $c_B = 0.676 \pm 0.004$, $c_V = 0.673 \pm 0.004$, and $c_I = 0.616 \pm 0.004$.

It is easy to transform eqns (1) and (2), remembering $(m - M) = 5 \log r_{\rm Mpc} + 25$, into

$$\log H_0 = 0.2 M_\lambda + c_\lambda + 5 \qquad (3)$$

where M_λ is the *absolute* magnitude of SNe Ia in the appropriate passband. Thus the problem of determining the *large-scale* value of H_0 is reduced, with c_λ now being known, to measuring the absolute magnitude M_λ of one (or a few) *nearby* SNe Ia. It must be stressed that the recession velocity of the nearby calibrator(s) does not enter in any way. The solution for H_0 is therefore independent of any peculiar motions.

1.1.2 Cepheids, HST, and the luminosity calibration of SNe Ia

With the luminosity calibration of a SNe Ia in mind, a small team was formed to observe with HST the Cepheids in IC 4182, home of the well observed SN 1937C. The Cepheids would give a good distance to that inconspicuous galaxy and hence fix the absolute magnitude M in the B and V band. Through eqn (3) this would lead to a high-weight determination of H_0.

The team consisted of Allan Sandage, Abhijit Saha, Lukas Labhardt, Duccio Macchetto, Nino Panagia, and the author. Their application for HST observing time was received with skepticism. Was the project not too easy for HST? Could it not be executed from the ground? It was a happy incidence for the team that the mirror of the most complex and most expensive telescope ever build produced miserable images when it was finally launched in 1990. IC 4182 became an ideal target. In fact the photometry of the extended stellar images on the CCD frames was easier than later, after the repair mission in 1993, when the crisp stellar images became smaller than the pixel size of the WFPC-2 camera, but, of course, offering higher resolution essential for more distant galaxies.

Cepheids are the most uncontroversial distance indicators in extragalactic astronomy. Their pulsational period is a function of their mass which in turn correlates with the luminosity. The corresponding log P–absolute magnitude M relation was discovered in 1912 by Henrietta Leavitt who studied the many Cepheids in the Large Magellanic Cloud (LMC) which all appear at the same distance. The slope of the log P–M relation for different passbands is today well determined from the LMC Cepheids. Remaining uncertainties concern the zeropoint of the relation and the effect of differences of the metallicity of the Cepheids in different galaxies. All authors agree that the metallicity effect is small, but they do not agree on the sign of the corresponding correction. Therefore an uncertainty of $0^{\rm m}.10$ will be included into the error budget.

At present it is customary to fix the zeropoint by adopting $(m - M)_{\rm LMC} = 18.50$. There is growing evidence that this value should be increased by $\sim 0^{\rm m}.05$–$0^{\rm m}.10$. Based on observations with the International Ultraviolet Explorer (IUE) and HST, Panagia (1998) has derived a *geometrical* distance of $(m - M)_{\rm LMC} = 18.58 \pm 0.05$ from the fluorescent ring of SN 1987A, a peculiar type II supernova in the LMC. Also the calibration of galactic Cepheids with the astrometric satellite HIPPARCOS tends to give a somewhat larger distance to the LMC (Madore and Freedman 1998; Feast 1999; Pont 1999; Groenewegen and Oudmaijer 2000) in agreement with RR Lyr stars and other distance indicators (Walker 1999; Gratton 2000; Sakai *et al*. 2000; Romaniello *et al*. 2000).

Smaller LMC moduli based on red-clump stars and RR Lyr luminosities, calibrated through statistical parallaxes, have comparatively low weight. Why two eclipsing binaries in the LMC give an unacceptably small modulus of 18.3, including HST observations (Fitzpatrick *et al*. 2000), is not understood at present.

DIVA, the planned next-generation astronomic satellite, is expected to measure very accurate parallaxes of 30–40 galactic Cepheids. This should settle for good the question of the zeropoint. ESA's very ambitious GAIA project would in addition measure the parallaxes of many supergiants and thus make a frontal attack on the distance of the LMC.

For the moment we will retain the customary zeropoint with $(m - M)_{\rm LMC} = 18.50$ and return to the question in Section 1.1.3.

The HST distance of the Cepheids in IC 4182 provided $M_{B,V}$ of SN 1937C which, inserted into eqn (3), gave a rather low value of $H_0 = 52 \pm 9$ (Saha *et al*. 1994). Essentially the same IC 4182 distance had been derived before from the ground from its brightest stars (Sandage and Tammann 1982), but the result had been criticized. Now, with the HST result the light curve of SN 1937C from archival material came under fire. The obvious step was to calibrate other nearby SNe Ia, which became feasible after the repair of HST.

The Century of Space Science
Bleeker, J.A.; Geiss, J.; Huber, M. (Eds.)
2001, XLIX, 1846 p., Hardcover
ISBN: 978-0-7923-7196-0