What are the Empirical Bases of the Hubble Law?

Konrad Rudnicki
Jagiellonian University Observatory
Krakow, Poland

Contrary to what is commonly believed, the Hubble Law is historically of theoretical origin. The strict redshift-distance proportionality was never deduced from observations. The interpretation of redshifts in terms of a Doppler effect is based on theoretical premises.

A Brief History of the Formulation of the Hubble Law

The positive correlation between the distance of a galaxy and its redshift was actually known before Hubble, *i.e.* before the scale of extragalactic distances was established. Stellar magnitudes of what were then called "extragalactic nebulae", or their angular sizes, were used as indicators of *relative* distances. In the first paper by Edwin Hubble devoted to the redshifts of nebulae (Hubble 1929), the relation between distance and redshift is given in a linear form, and it is this form that has come to be known as the Hubble Law. Only a few people were actually witness to the prehistory of the formulation of the law.

In fact, like his predecessors, Hubble first tried to find a polynomial form suitable for a regression curve of redshifts on the distance axis. After first acquainting himself with de Sitter's model of the Universe (de Sitter 1917), he abandoned the terms in other powers of distance, and accepted the linear form (Gates 1962). Of course any correlation, and any empirical function, can be represented in a linear form as a first approximation.

The Hubble Law was viewed by many as an observational confirmation of General Relativity in the early years. Even today, second-rate popular treatments often make this claim. However, it later became clear that the linear expansion of the Universe in most relativistic models (*e.g.* Friedman 1922, 1924) results not from General Relativity but from the mathematical assumption of uniformity and isotropy. This assumption was later called the Generalized Copernican Cosmological Principle. Every model of the Universe based on this assumption, independent of accepted physical theory, must obey the linear Hubble Law. On the other hand, relativistic models constructed without the Generalized Copernican Cosmological Principle do not necessarily fulfill the Hubble Law, as for the example Kurt Goedel's model (1949). This shows that the Hubble Law is not related to General Relativity or to any other physical theory, but to a cosmological principle, a mathematical assumption based on philosophical conviction.

The strength of the Generalized Copernican Principle was not realized at all for a long time. This fact is also reflected in its alternate name: the *Weak* Cosmological Principle. Contrary to its modest name, it creates a very specific property of models; it allows only systematic radial motions of the substratum, the velocity of these motions being proportional to the distance as seen by any (real or imagined) observer located on any particle of the substratum. The proportionality constant can theoretically take on all possible values, and may even be zero. It is the task of observation to establish a numerical value, which may be a function of time.

The Attitude of Early Cosmologists toward the Hubble Law

It must be said to Hubble's and his collaborators' credit that when they formally adjusted the redshift-distance relation to the de Sitter Model, they were aware that the observational data did not necessarily have to be regarded as confirmation of the expansion of the universe. They merely regarded expansion as the simplest of many possible hypotheses.

In one of the first papers devoted to this problem, Milton L. Humason (1931) writes:

> *It is not at all certain that the large red-shifts observed in the spectra are to be interpreted as a Doppler effect, but for convenience they are expressed in terms of velocity and referred to as apparent velocities.*

Edwin Hubble and Richard C. Tolman (1935) wrote the following about the redshift-distance relation:

> *The most obvious explanation of this finding is to regard it as directly correlated with a recessional motion of the nebulae, and this assumption has been commonly adopted in the extensive treatments of nebular motion that have been made with the help of the relativistic theory of gravitation and also in the more purely kinematic treatment proposed by Milne. Nevertheless, the possibility that the redshift may be due to some other cause, connected with the long time or distance involved in the passage of light from nebula to observer, should not be*

> *prematurely neglected, and several investigators have indeed suggested such other causes, although without as yet giving an entirely satisfactory detailed account of their mechanism.*
>
> *Until further evidence is available, both the present writers wish to express and open mind with respect to the ultimately most satisfactory explanation of the nebular red-shift and, in presentations of purely observational findings, to continue to use the phrase "apparent" velocity of recession. They both incline to the opinion, however, that if the red-shift is not due to recessional motion, its explanation will probably involve some quite new physical principles.*

In the above statements, not only is the nature of the redshifts considered to be uncertain, but, even assuming the Doppler interpretation to be correct, the authors do not see any need to connect it with General Relativity. They refer to a very general kinematic theory put forward by Milne (1935a). Moreover, in the same paper, they propose tests of the nature of galaxy redshifts that might be performed by future investigators. Later, however, these tests regrettably came to be viewed as too primitive. The last sentence of the paper contains the following statement:

> *It... seemed desirable to express and open-minded position as to the true cause of the nebular redshift...*

It is appropriate to add here that the theory of E.A. Milne and its formulae were developed from deep-seated kinematic considerations, without recourse to the assumption of the existence of "laws of nature" or appeal to any specific theory of gravitation, but the results were shown to be capable of a simple dynamical interpretation on an action-at-a-distance view of Gravitation (Milne 1935a). Milne was the first person (Milne 1932) who was courageous enough to raise doubts as to the validity of relativistic cosmology; he set forth his reasons in detail a few years later (Milne 1935a).

There is no need to quote all the early scientists who entertained doubts about the "simplest interpretation" of the Hubble Law. We should, however, mention Fritz Zwicky, who, several decades later, continued to use the symbol V_s for redshifts of extragalactic objects (*i.e. symbolic* velocity expressed in km. s^{-1}) instead of V_r (*radial* velocity expressed in the same units).

These examples show that, almost from the time the Hubble Law was enunciated, its interpretation as a confirmation of Friedman-type models of the Universe was readily accepted by persons less familiar with astronomy, but certainly not by some of the more reputable scientists.

Three Facets of the Hubble Law Problem

Before we can accept the Hubble Law as an observational confirmation of the relativistic general expansion of the Universe, the following independent proofs are required:

1. Proof that the redshift-distance $(1+z,r)$ relation is strictly linear, not only to a first approximation (of course, corrections are allowed for the curvature of space or variation of the Hubble constant with time if one wants to support a relativistic or similar model). It must be remembered, as every student of mathematics knows, that any three points can be connected with a straight line, provided the line is thick enough.
2. Proof that the observed redshifts are caused mainly by the Doppler effect, while other causes (*e.g.* gravitational effects) remain insignificant for large distances).
3. Further, a separate proof is needed that the Universe conforms with a Friedman model, and that its instability is the physical cause of the expansion.

Performing the first of these proofs would be of great importance in itself, since it would qualify the redshift as a simple, secure indicator of extragalactic distances, and not just in a statistical sense. This would have enormous practical significance. A large part of modern discussions about the Hubble Law is devoted to this practical problem. But the proof would also be very important for the physical understanding of the Universe. If it turns out that the effect is strictly proportional to the distance independently of the nature of the objects, we could conclude that the phenomenon is due either to space or to time by itself. If, for example, we decide to attribute the affect to intergalactic matter, we should conclude that this matter is distributed completely uniformly over the entire Universe accessible to our observations. (It is worth noting that this implication is not valid in reverse. The redshift may be caused solely by space or time, even if it is not strictly proportional to distance, provided space-time is not uniform.)

It may be almost impossible to carry out Proof No. 2, even if Proof No. 1 turns out negative, but it should not be rejected *a priori*. There is still a possibility that redshifts are Doppler in origin, but connected with the nature and history of bodies, not space or time. There could be some additional (unknown) effects which make bodies moving toward us with high velocities invisible. Another possibility would be to accept a real but irregular expansion of the Universe. It may be mentioned here that it was Milne (1935a) who showed that a phenomenon of general expansion observed in terms of radial velocities is not equivalent to pure physical expansion of a system of celestial bodies. The formulae connecting redshifts with the spatial motions of celestial bodies are quite complicated even on the assumption that only systematic, rather than peculiar, motions are observed. In this case, however, one should bear in mind that no real Universe that showed systematic velocities with distances but revealed no strict proportionality to distance could ever be reconciled with either the Generalized Copernican or the Perfect Cosmological principles.

Proof No. 3 can be accepted without No. 1 or No. 2 in only one case, namely when we suppose a static relativistic model of the Universe. Under these circumstances, a Universe model may be constructed according to General Relativity even if the redshifts are neither Doppler in origin nor strictly proportional to distance. Of course, other physical theories must be added to explain the nature of the redshifts in this case.

Summing up then, the Hubble Law, with its simplest explanation as a confirmation of the Friedman model is not a monolithic, indivisible statement. Its three main components—the phenomenological part, the Doppler explanation and its application as an argument confirming the Friedman model—are logically independent, and may be confirmed or falsified, accepted or rejected, independently of one another.

The General Problem of Proofs for Cosmological Statements

Let us assume that all three of the proofs mentioned in the previous section could be carried out some day, and that the Hubble Law and its usual interpretation are proven valid in all areas of the Universe accessible to our observations. Even in this hypothetical case the law would not be proven in general, only in the area included within the cosmological horizon. Only if one could prove that no cosmological horizon existed and that the entire Universe is accessible to our observations (which is not impossible, but still highly unlikely), could the Hubble Law or any other law concerning the entire Universe be proved observationally. Observational testing of cosmological models means testing inside the cosmological horizon only, and is by no means equivalent to a definite proof.

There are two possible ways to approach this simple truth. The first is to take the position that every science must be based solely on experiment and observation. Since no observations or experiments or observations can be performed beyond cosmological horizons, no cosmology, nor even a science about the entire Universe is possible. We are restricted to practicing astronomy only. The second approach is to realize that in every science we rely on some form of thinking in any case. No science is possible if it is based on experiment and observation alone, without thinking. In this respect, cosmology may be said to differ only quantitatively from other sciences. Certainly, in cosmology we enter a realm where thinking plays a much greater role than in the other sciences. In fact, as soon as we can think about something, we already know at least a little bit about it.

The Status of Cosmological Principles

The products of our thinking that make it possible for us to know something about the entire Universe we call Cosmological Principles.

The Hubble Law is a consequence of the acceptance of the Generalized Copernican Cosmological Principle, or its offshoot, the Perfect Cosmological Principle. Thus, the Hubble Law is connected more with the way we think than with observations. To put this in a paradoxical, but perhaps clearer way: If we could find any sure and positive reason in our logical thinking to accept the Generalized Copernican or the Perfect Principle, we should accept the Hubble Law even when all the observations argue against it. We should then consider these observations as evidence of an exceptional status for the part of the Universe in which we live, rather than as evidence against the Hubble Law.

In fact, both the cosmological principles mentioned above are very primitive indeed. Both were introduced merely to fill in—in the simplest way—the blanks in our map of the Universe, blanks covering the areas situated beyond the cosmological horizon. If one accepts, for example, the Perfect Cosmological Principle, one has the sense of knowing everything about the entire Universe. One knows not only how the Universe looks in every direction and at all distances, but even how it looked in any past epoch and how it will look in any epoch in the future. Simply put: It looks and has always looked the same as it looks here and now. It is very tempting indeed to accept this very powerful cosmological principle.

I shall not discuss each of these cosmological principles or investigate the general problems connected with them. Fifty years ago, only one cosmological principle was known. Five years ago, there were five. Now, the Anthropic Principle is claimed by some to be a cosmological principle. When we study some lesser known publications, it seems that at least fifteen different principles have been formulated, all laying claim to cosmological validity. I have already discussed this problem at some length (Rudnicki 1989), and I intend to return to it in the light of more recent publications at some point. Here, however, I only wish to express my conviction that the development of our thinking about the Universe is still at a very primitive level indeed. I hope that with further development, we shall learn to use better techniques for thinking and understanding the Universe as a whole than we do today with our cosmological principles.

Conclusions

Returning to the problem of the Hubble Law, we have established that it is a product of the Generalized Copernican Cosmological Principle.

It follows logically and mathematically from this principle. To accept one or another cosmological principle today is a matter of belief. By the same token, accepting the Hubble Law and its current interpretation can only be a matter of personal choice.

To avoid any misunderstanding, let it be said that I make no claim to the effect that the Hubble Law is not a good phenomenological approximation to the redshift-distance relation. Nor do I insist that the Universe is not expanding or that the Big Bang did not take place. I consider all three to be possible hypotheses. But I do claim that other, diametrically opposed, hypotheses are possible as well. Even if all the observational data were in agreement with one set of hypotheses, I would like to consider all the alternative explanations as possible as long as available observations can be explained by them as well. That the majority of cosmologists today share a certain conviction is by no means an argument for its truth. The verification of the Hubble Law and its interpretation was at the start, and remains today, an open problem.

Several observational facts can be quoted as arguments against acceptance of the Hubble Law as something more than a first approximation. Many hypotheses can be quoted as alternative explanations of the observed redshifts. To present these observations and theoretical considerations is the task of the other speakers in this Workshop.

References

Friedman, A., 1922. *Zeits. f. Phys.*, **10**, 377.

Friedman, A., 1924. *Zeits. f. Phys.*, **21**, 326.

Gates, H.S., 1962. *Private communication confirmed by F. Zwicky and other senior astronomers in Pasadena*.

Goedel, K., 1949. *Rev. Mod. Phys.*, **21**, 447.

Hubble, E., 1929. *Mt. Wilson Communic*. **No. 105.**

Hubble, E. and Tolman, R.C., 1935. *Astroph. Journ.*, **82**, 302.

Humason, M.L., 1931. *Astroph. Journ.*, **74**, 35.
Milne, E.A., 1932. *Nature* **,130**, 9.
Milne, E.A., 1935a. *Mon. Not. Roy. Astr. Soc.*, **95**, 560.
Milne, E.A., 1935b. *Relativity, Gravitation and World Structure*, Oxford Press.
Rudnicki, K., 1989. *APEIRON* **,4**, 1.
de Sitter, W., 1917. *Proc. Acad. Wetensch. Amsterdam*, **19**, 1217.