ASTRONOMY PAPERS

Reprinted from

THE ATMOSPHERES OF MARS AND VENUS
National Academy of Sciences—National Research Council
Publication 944

Appendix 1

DIRECT PHOTOGRAPHY IN THE EXPLORATION OF PLANETARY ATMOSPHERES

A. G. Wilson

A. Information Spaces

In the systematic study of planetary atmospheres one primary role of direct photography is in the detection and identification of the indigenous meteorological phenomena. In planetary astronomy the detection of an object is accomplished through inter-comparison of areal "information cells" with respect to their brightness, color, and variation in time. The detectable objects are those which emerge above the contrast and resolving power thresholds, and are contained in the light response, spectral, temporal, and angular ranges of the instrumental system. Identification as meteorological phenomena must be through comparison of the detected objects with familiar meteorological objects or events in the Earth's atmosphere.

There exists what may be termed a "similarity threshold," on one side of which phenomena detected on other planets may be identified with familiar terrestrial phenomena or recognized as extrapolations of terrestrial phenomena, but on the other side of which their identification, and even their reality, becomes speculative. As more detailed knowledge of other planets is collected, the base of the familiar against which comparisons are made will be broadened. It is epistemologically fortunate that the first planets to be explored, Mars and Venus, are quite similar to the Earth, allowing ready identification of many phenomena. The exploration of these planets should extend the base of the familiar and provide experience which will create a more advantageous similarity threshold.

Since direct photography must play a basic role in any exploration program whose end is the discovery, observation, and analysis of the meteorological processes on other planets, it will be useful to have a measure of the relative capabilities of various photographic systems for planetary atmosphere studies.

For this purpose the process of direct photography may be considered to be a function of five basic parameters or "dimensions." These are the two linear or angular dimensions of the region photographed, the brightness dimension (which is recorded as photographic density), the spectral dimension, and the temporal dimension. In each of these dimensions there are bounds which define the *range* of a system and a threshold which sets the *resolving power*, as described in Table I.

Although there is not complete symmetry in considering the parameters in this way, the viewpoint is useful in that it allows the construction of a five-dimensional "Instrument Information Space" —the extension of the space is determined by the ranges, the size of the information cells is determined by the resolving powers—which provide a set of figures of merit for evaluating the capabilities of photographic (and other instrument) systems.

The instrument information space for direct photographic observations of planets is defined by the following four factors:

(1) the optical and photographic parameters such as telescopic aperture, emulsion sensitivity, grain, contrast, etc.;

(2) relative motions of the instrument and field being photographed;

(3) the location of the instrument with respect to the planet being explored;

(4) considerations of technical and economic feasibility.

Table I

DIMENSIONS FOR PHOTOGRAPHIC SYSTEMS

Dimension	*Resolving Power*	*Range*	*Effective Range*
Areal (2 dimensions)	Angular resolving power set by telescopic and photographic parameters together with seeing and instrument stability limitations.	Determined by angular field of view.	Field of view, modified by aberrations of the optical system.
Brightness	The contrast, λ	Determined by signal-noise ratio and emulsion saturation.	Determined by a set of exposure times.
Spectral	The filter-emulsion-optical component band pass.	Determined by filter-emulsion combinations and optical and atmospheric transmission limits.	Sums of bands at which exposures are taken.
Temporal	Exposure time and/or frequency of exposure.	Span of observations.	Determined by number of exposures, divided by frequency of exposures.

Item (1) is fully discussed in many texts on photography and optics. (v., e.g., J. Strong, *Procedures in Experimental Physics,* Prentice-Hall, 1945, or G. de Vaucouleurs, *"Planetary Astronomy from Satellite Substitute Vehicles,"* Chap. II, AFMDC–TR–60–6.) Item (2) refers to the dynamic stability of the telescope-camera system, periods of natural oscillations, guidance, relative motions caused by planetary rotations, movement of the instrument carrier, etc. Item (3) takes into account the limiting effects of seeing, sky brightness, and spectral transmission properties of the Earth's atmosphere, also the effects of distance to the planet on the linear field of view and linear resolving power. Item (4) involves the sizes of instruments which may be practically carried in balloons, placed in orbit about the Earth, or carried in fly-by probes and planetary landing capsules. It also involves the cost of constraints of each system and the extent of program economically feasible with each system.

The information space required for the photographic study of a given phenomenon must provide sufficient data for determining the size, structure, position, and movement of the phenomenon (as, say, a storm), the life span, rates of growth, decay, and other changes, the season and frequency of occurrence, and such physical quantities as brightness and color. An information space adequate for this purpose may be called a *description information space* for the phenomenon. Since it is redundant to observe all parts and all features of a phenomenon to the same degree of detail, it is evident that the description information space for the study of the structure and behavior of a phenomenon will best consist of a *set of instrument system information spaces.*

Two basic problems thus arise: First, the defining of the description information space for the phenomenon, and, second, the selection of the set of system information spaces which must economically (with respect to time, energy, dollars) span the description information space.

To provide illustrations of these concepts and to provide an example of the data which determine a description information space, and also what instrument system information spaces would best cover the description space, it is useful to consider the present situation of knowledge concerning the planet Mars—which is by far the best observed of the planets—gathered over the past 70 years by visual and photographic exploration from the Earth's surface.

In the following section, the phenomena that have been ascertained to exist on Mars are listed together with what is generally known quantitatively about these phenomena.

B. Photographically Observed Phenomena on Mars

One of the first constraints on our knowledge of Mars is the limiting resolving power to which the planet has been observed.

At the most favorable oppositions, Mars' closest approach to the Earth is about 5.6×10^7 km, or 150 times the distance to the Moon. The angular diameter at this time is only 25 seconds of arc compared with 31 minutes for the Moon. Since the average resolving power limit set by the atmospheric seeing is about 1 second of arc and the resolving power of the eye is of the order of 1 minute of arc, the amount of information on an average photograph of Mars is about the same as that which the naked eye receives from the Moon—roughly 10 bits ($\log_2 10^3$). However, there do exist a few photographs of Mars taken at instants of excellent seeing that contain perhaps 100 times this information. Because of the eye's ability to accommodate to seeing effects, it is able to do even better than the photographic plate and many observed Martian phenomena, such as the canals, lie in the region of information space beyond the 16.5-bit ($\log_2 10^5$) level of the best photographs and limited by the eye's capability. Under best conditions there is perhaps another augmentation factor of ten or so.

The present best linear resolution, corresponding to an angular resolving power of 0.1 second of arc, is about 30 km on Mars. This is for point phenomena like oases. For linear phenomena like canals, provided they are long enough, the resolution may be less than 5 km. Visual observers feel that an additional increase in resolution by a factor of 10 would give an information "break through" with regard to knowledge of Martian phenomena similar to the revelation of craters and mountains on the Moon that came with the first telescopes.

A second bound on our present knowledge of Mars is brought about by the temporal resolution with which the planet has been observed. The temporal range and resolving power of information space for Mars is more difficult to estimate. Photographic records go back to about 1890, but the observational coverage of Mars has been limited mostly to the few weeks before and after opposition, when the angular diameter is greater than about 15 seconds of arc. Unfavorable oppositions have for the most part been inadequately observed, partly because Mars is smaller and partly because these oppositions occur during the bad observing season for most observatories. The observations have had adequate temporal resolving power to determine the Martian seasonal changes, (for the southern hemisphere of Mars, spring is 146 Earth days, summer 160 days, autumn 199 days, and winter 182 days), but have been grossly inadequate for synoptic studies of the Martian atmosphere. An aggregation of all photographic observations might give an average temporal resolving power of one day for the six weeks preceding and following the most favorable oppositions since 1909.

During the oppositions of 1954 and 1956 the International Mars Committee organized a worldwide photographic patrol. The 1954 coverage, which was the all-time best, is shown graphically by Table II and Table III taken from the 1954 report of the Mars Committee. (Mitchell, R. I., *The 1954 International Mars Photographic Patrol,* Mars 1954, Report of the International Mars Committee, Lowell Observatory, Flagstaff, 1955.) On the left of the tables are given the Earth dates, on the right are given the Martian seasonal dates, it being southern hemisphere spring.

The photographic coverage of Mars with multicolored filters or color photography has been experimental rather than systematic. Except for the blue-yellow-red patrols of one or two observatories at recent oppositions no complete photographic comparative color record exists.

Table II

OBSERVATIONS OF MARS 1954

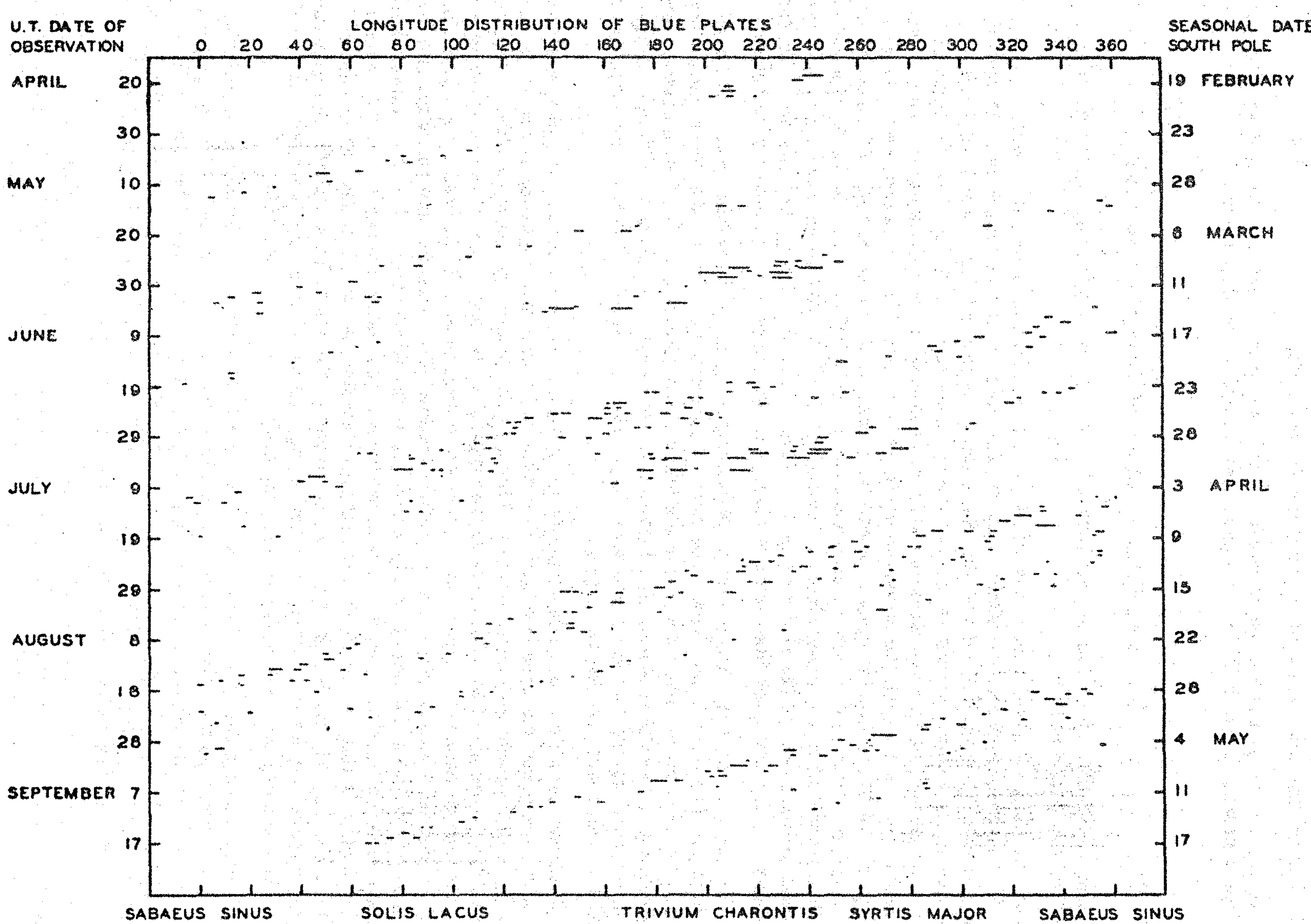
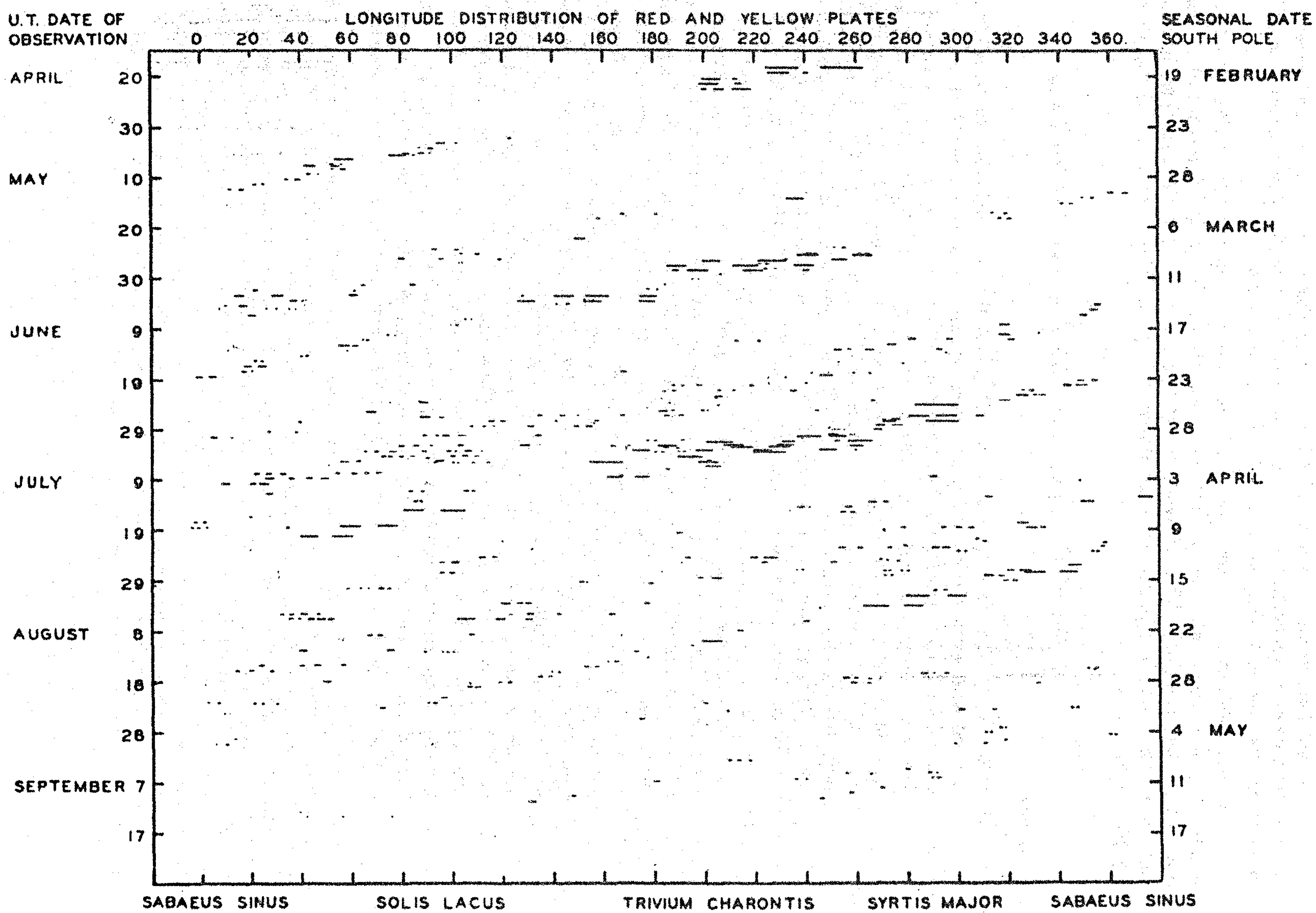


Table III

OBSERVATIONS OF MARS 1954



Faster and finer-grain emulsions have become available in recent years, and experiments with the contrast parameter have been made. Image tubes rather than direct photography, however, show the greatest promise for exploiting what can be done with contrast.

For *direct photographic* information purposes, Martian phenomena can be classified with respect to position, size, color occurrence, duration, and rate of change (with respect to formation, dissipation, size, color, position). The following list is a summary of generally available observational knowledge concerning Martian phenomena, omitting theoretical inferences and interpretations. (v., also, Chapter III on Mars.)

C. List of Photographically Observed Martian Phenomena

1. *Surface*

(a) *Polar Caps.* The caps appear to be formed in the seasonal autumn while they are largely covered with white clouds or fog. Toward the end of seasonal winter these clouds disappear, and a dark fringe appears on the edge of the caps which then begin to recede in size. The decrease continues through the seasonal spring, the dark fringe being widest when the melting rate is fastest. (Quantitative data on melting rates consisting of the size of four cap areas with corresponding dates, given by Pickering, 1924.) The South Cap is centered on long. 40°, lat. 83°; maximum size: to 45° lat.; minimum size: can disappear completely. It exhibits rifts in spring (Mountains of Mitchell), also occasional bright spots near edge. The North Cap at maximum size extends to 57° lat.; minimum size: 300 km (1° = 57 km).

(b) *Dark Markings.* Termed maria. For the most part, located in the southern hemisphere, have been carefully mapped and named. Cover about ⅜ of the surface area, are mostly permanent, but additional dark areas appear from time to time lasting for a few years. In 1954 a new area "size of Texas" (± Pecos County) northeast of the Syrtis Major was observed which had developed since the last observations in 1952. Seasonal color changes occur moving from the polar caps toward the equator in seasonal spring. Rate of advance of the change is about 45 km/day. The color change is regarded by most visual observers as from gray (or blue gray) to brown or violet. Dark areas are faint in seasonal winter.

(c) *Bright Areas.* Called "deserts." Are of a general orange or ochre color. Cover about ¾ of the surface area of Mars; are static. Observations of limb indicate that no abrupt heights on Mars exceed 2500 ft. (Lowell).* There are areas that become temporarily whiter from time to time. Hellas for example is usually whiter than the other desert areas.

(d) *Canals.* Controversial network of fine linear markings, many permanent. Some observers have mapped over 400, a fifth of them being double, widths under 25 km. Show color changes similar to dark markings. Seasonal color change moves more slowly in canals (18

---

* Remark by *C. W. Tombaugh*: Wilson quotes Lowell's value of 2500 feet as the level of abrupt heights of terrain on Mars. I strongly disagree with Lowell's value. It should be remembered that at the time of greatest phase defect (when the best opportunity to see relief occurs) Mars is at twice its oppositional distance. Also, the terminator is, even then, far from the center of the disk, so that any horizontal distance of a cast shadow is foreshortened by a factor of about two. This means that the shortest perceptible horizontal distance in the vicinity of the terminator is $30 \times 4 = 120$ km on the basis of 0.1 arc-second of resolution. (v. Chapter I, Table 2.) I would say that 0.15 arc-second is a more realistic limit of resolution. Then the smallest perceptible terminator resolution would be 180 km, and the smallest perceptible height would be 4.7 km = 15,500 feet. (Assume a projection, or cliff, casting a shadow on a smooth plain, and the rays at the edge of the shadow tangent to the surface.) But this kind of resolution comes only in fleeting glimpses. An observer would have to confine his attention to a few favored candidate areas on Mars with alert attention, hoping for the superb glimpse to occur before the planet's rotation carries the local area out of the opportune circumstance, which would occur within a few minutes. I would estimate that the amount of time that there is atmospheric seeing of the high quality required to attain this goal is less than one per cent of the time that the planet is within three hours of the observer's meridian. No good telescope of 36-inch aperture or greater is available to such an interested observer. Let us come back to Lowell's case. His 24-inch refractor is an excellent instrument. I have looked at planetary detail with this instrument for a total of some 500 hours.

km/day) than on maria. Largest of canals have been photographed. Some visual observers claim under best seeing conditions canals are resolved into broken linear dark markings.

(e) *Oases.* Roughly circular dark markings usually at intersections of canals, diameters of the order of 150 km. Some 200 have been reported.

2. Atmosphere

(a) *Clouds.* Three cloud species, yellow, blue, and white, exist. (White clouds may be distinct species or only thicker blue clouds.) White or blue clouds may be observed on any part of the disk but are concentrated toward the limb. Angular sizes up to 45° in areographic coordinates (about 3000 km) extending along the limb are observed. On the morning side of the disk, clouds may extend almost to the noon meridian, on the afternoon side rarely over 45° from the terminator. During several periods of observation, morning clouds were photographed only over maria, afternoon clouds only over deserts. Clouds were regularly observed on the limb but not on the terminator at 40° phase, indicating that observed limb clouds may be an observational foreshortening appearance of atmospheric haze rather than a distinct physical phenomenon. White (or blue) clouds may form in less than 24 hours and last over two weeks. White clouds are most common at aphelic oppositions. A total of about two dozen measurements of cloud movements available. Speeds up to 35 km/day have been computed. Yellow clouds are usually associated with perihelic opposition. Major storms involving yellow clouds occurred near the 1924 and 1956 perihelic oppositions, resulting in the covering of the entire planet for several days with a yellow pall. After 1956 storm, polar cap reappeared quickly, dark markings more slowly.

(b) *Blue Haze.* Thin haze covering entire planet rendering surface features invisible in photographs taken in blue light ($\lambda <$ 4330 Å). Haze dissipates from time to time particularly near oppositions. Haze can dissipate or form in 3 or 4 hours. Clearings may be planet-wide or cover as small an area as ⅛ of the disk.

3. Other Phenomena

Large W-shaped cloud observed in 1926, 1954, and 1958 associated with oasis-canal network in Tharsis region (Tithonius Lacus), rotates with surface. Cloud bands parallel to equator on blue photographs and Y-shaped haze patterns on blue photographs from time to time. Short duration bright flash reported 1954 by Saheki.

From this outline of observed phenomena on Mars, it may be assumed with a fair measure of confidence that all contrasting daytime Martian phenomena photographable in the λ3500 to λ6500 range whose extents are greater than 50 km, which may be observed near opposition, and whose temporal durations exceed two or three days, are now known. This may be taken as the present description information space of Mars.

D. The Scale of Atmospheric Phenomena

It is also evident that the available *quantitative* observations with regard to sizes, rates of motion, rates of growth and decay and time spans of Martian atmospheric phenomena is very sparse. This deficiency is caused by three principal factors: (1) the low resolving power of photographs as limited by the seeing, (2) lack of observations taken often enough over long enough time spans, and (3) incomplete use made of the observations which do exist. This third point is attributable both to the fact that much data has not been published or made generally available and to the fact that a

---

(Cont.) This telescope is equipped with an iris diaphragm which the observer can conveniently regulate from an aperture of 24 inches down to 6 inches. In my experience, I have never been able to gain in finer detail with apertures larger than 20 inches. The secondary chromatic aberration becomes severe enough to spoil any view the seeing might allow. Lowell and his associates could not possibly have seen any Martian detail smaller than 0.2 arc-second. Therefore, the smallest possible horizontal distance he could have observed in the vicinity of the terminator would be 240 km. The smallest vertical height that could be detected would be 27,500 feet. Lowell was certainly wrong by a factor of 10. Thus, this means of mapping relief is beyond ground-based capabilities.

great many photographs have not been completely reduced. The present situation is such that the data required for synoptic studies of the Martian atmosphere are not available. This means that much of the analytical work which has been done so far has necessarily been top heavy with assumptions and overinterpretations of the available data.

What then is a description information space adequate for synoptic studies of a planetary atmosphere? In order to answer this question it is necessary to know the linear and temporal resolutions necessary to describe meteorological phenomena.

On the Earth the smallest atmospheric phenomena of meteorological significance are perhaps tornadoes and thunderstorms. These have a spatial extent of the order of from two to five km. From atmospheric events of this scale, sizes range up to planet-wide circulation patterns. What spans of observations and temporal revolving power should be used in order to observe adequately these atmospheric phenomena of various sizes? The answer depends on the lifetime of the phenomena and their rates of evolution.

Figure 1 shows a relation between average sizes and lifetimes of four types of terrestrial atmospheric phenomena: tornadoes (A); thunderstorm cells (B); hurricanes (C); and cyclonic storms (D). The approximate average size in kilometers is plotted as the ordinate and one-tenth the average lifetime as abscissa. It is assumed that a generally useful interval between observations of such phenomena is about one-tenth their lifetime. Thus, a tornado should be photographed every three minutes, a hurricane observed every 12 hours, etc.

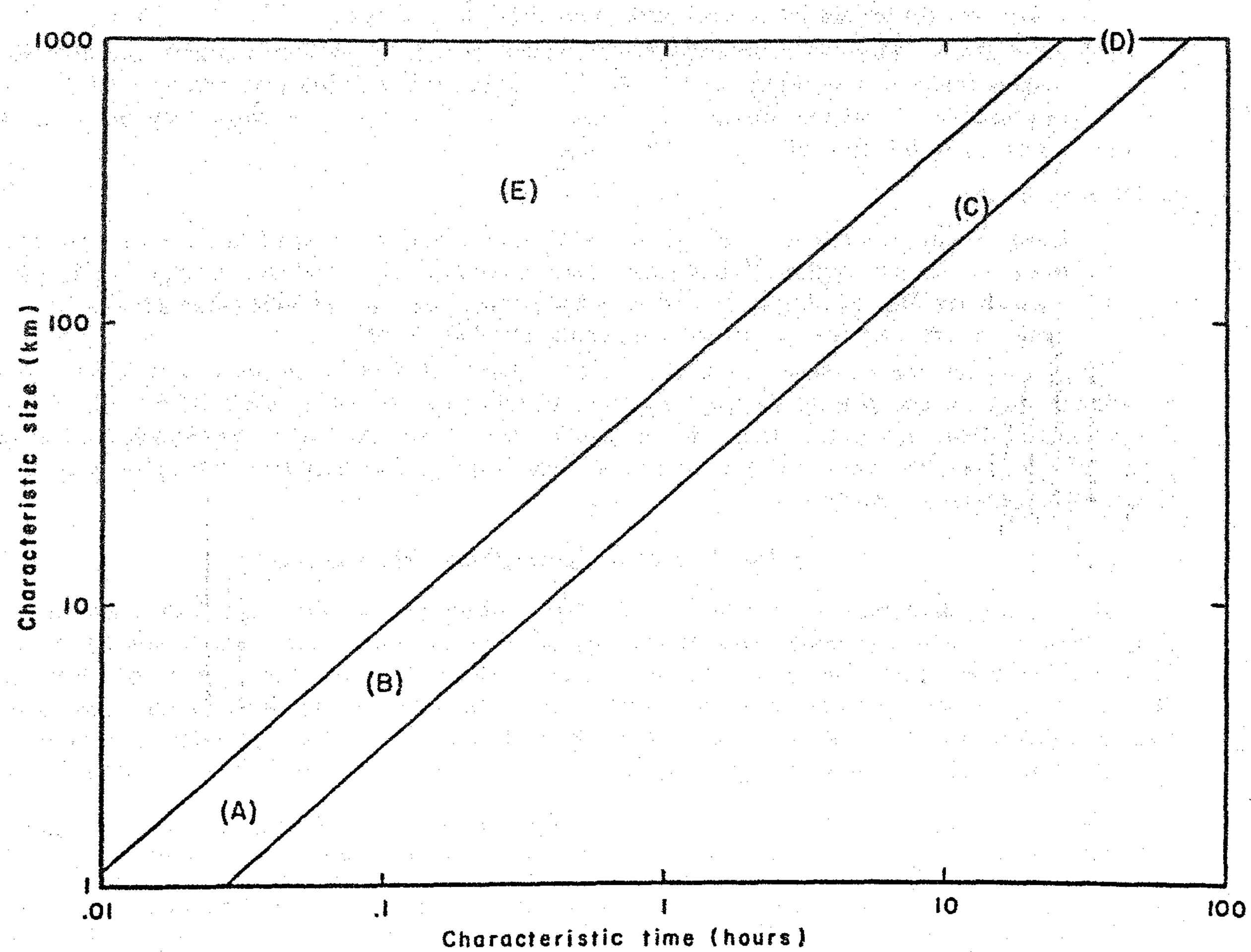


Figure 1

It is thus seen that the frequency with which a phenomenon must be observed to obtain an adequate description information space depends on the spatial extent of the phenomenon, Fig. 1 giving the relation between extents and frequencies for terrestrial atmospheric phenomena. The time span for the observations should be about ten times the plotted characteristic time.

It is reasonable to assume that in the atmosphere of Mars (and other planets) a similar relation of the type

$$s = a\,t^b$$

holds between a characteristic size s and a characteristic time t for several classes of atmospheric phenomena. There is no *a priori* reason for assuming, however, that the coefficients a and b in the equation are the same as for the Earth, though this may prove to be the case. But as a starting point for the determination of the suitable instrument information spaces for Mars, the relations of Fig. 1 can be used to give the temporal resolutions and spans required to describe the meteorological phenomena of various sizes.

An interesting exception to relations of the type illustrated in Fig. 1 is given by (E), the plotted position of the characteristic size and time of a *thunderstorm complex.* Such a complex consists of a great many thunderstorm cells and may be considered as an *aggregate phenomenon* rather than a simple phenomenon.

In the exploration of another planet, exceptions to characteristic time vs. characteristic size relationships may indicate the presence of such aggregate phenomena and suggest areas of investigation which call for observations with higher resolutions. This concept also indicates how, when certain basic relationships have been determined, the epistemological resolving power of the exploration may be extended beyond the actual resolving power of the observational instruments.

To aid in determining the set of instrument information spaces that would optimally span a description information space for synoptic studies of the Martian atmosphere, comparative instrument information spaces are given in Table IV for a 60-inch telescope located on the Earth's surface, a 20-inch telescope in a balloon at 30,000 meters, and a 5-inch telescope in a fly-by probe 40,000 km and 1,000,000 km from Mars.

In Table IV the characterizing parameters are listed on the left column; the other columns show their relative ranges.

The exposure time is limited on the short end by signal-to-noise ratio and on the long end by emulsion saturation, atmospheric turbulence, sky brightness and the relative motions of the object and camera.

The matter of economics of ballon flights is complex. Flights for isolated purposes such as obtaining physical data in spectral ranges inaccessible on the surface of the Earth, or taking photographs with a finer resolving power than can be made from the surface, can be readily justified. But the relative cost of using balloons for patrol purposes, such as the monitoring of atmospheric changes, as compared with surface observations, must be questioned. The atmosphere does not present a filter against detection of changes in phenomena that are already observable from the Earth's surface.

E. Conclusions

About all that can be learned without great effort and expense from the Earth's surface concerning *static* phenomena on the nearer planets has been learned. It would be fruitless, for example, to try to continue to make marginal gains in resolving power against the obstacles of seeing when telescopes above the atmosphere can give orders of magnitude improvements.

Many plates of planets taken at various observatories over the past half century have not been reduced quantitatively to obtain the physical data which they contain. In view of the obvious superiority for planetary studies of observations or physical measurements made from outside the Earth's atmosphere, it would not be worthwhile to reduce most of this plate material.

Table IV

COMPARATIVE SYSTEMS INFORMATION SPACES FOR PHOTOGRAPHY OF MARS

	For Mars at Most Favorable Opposition: Distance 56 $\times$ 10^6 km		5-inch Telescope in Fly-By Probe at 10^6 km from Mars	5-inch Telescope in Fly-By Probe at 40,000 km from Mars
	(A) 60-inch telescope at Earth's surface	(B) 20-inch telescope, balloon-mounted, elevation 30,000 meters		
Angular Field	Entire planet = 25 seconds of arc.	Entire planet = 25 seconds of arc.	Entire planet = 20 minutes of arc.	Entire planet = 10° of arc.
Angular Resolving Power	Theoretical optical, 25 km; seeing limited, 140 to 560 km.	Theoretical optical, 75 km; guidance limited, 3 km.	5 km.	0.2 km.
Spectral Range	0.32μ to 0.9μ.	0.29μ to limit of emulsion.	Limited only by Martian atmosphere and emulsion.	
Exposure Time	Signal : Noise limit to seeing and emulsion saturation limits. Rotation of Mars limit, 0.01 sec to 2 min.	Signal : Noise limit to emulsion saturation. Rotation limit, guidance limits, 0.01 sec to 2 min.	Limited by the velocity of the probe relative to the surface of Mars, guidance limits.	
Frequency of Exposures	Any frequency up to reciprocal of exposure time.	Same as surface except for uneconomical frequency intervals.	Up to reciprocal of exposure time.	
Number of Exposures	Limited by economic and data processing factors.	Limited by economic factors governing number of flights.	Limited primarily by data storage and transmission capability; and possibly by frequency of exposures and allowable span of observations.	
Span of Observations	Limited by planetary configurations and sky brightness. Possible up to 45° from Sun.	Daytime observations possible up to 5° from Sun.	Determined by orbital parameters of probe.	

Nonetheless, a large and important role remains for surface observations. (v. Chapter V. C.) Essentially no quantitative work has been done on the *changes* in planetary phenomena. Dynamic planetary phenomena (atmospheric phenomena for the most part) have not been observed frequently enough or over long enough time spans to afford any but the most vague ideas of their properties.

This situation indicates that intensified uniform observations from the surface of the Earth with good instruments, geographically spaced to give a complete coverage of Mars and complemented with an efficient data distribution and reduction facility, is absolutely essential to fill the present gap in our knowledge of planetary atmospheres. It would also be most important to re-examine and reduce according to a standard procedure those existing photographic observations which are suitable for studies of changes. Particularly, this is important for studies of secular changes, for which they are our only source of data.

The following programs for *direct photography* can be recommended:

1. *From the Earth's Surface.*
   - (a) Study of existing plate material for data on secular changes.
   - (b) Study of existing plate material for *quantitative* data on atmospheric dynamics.
   - (c) Continuing program of observations for data on the dynamics of atmosphere, motions of clouds, storms, the formation and decay of phenomena. These observations should include time-lapse photographs, taken with various colors and polarizations, which can be differentially superimposed to study various aspects of changing phenomena.

2. *From Balloons*
   - (a) High-resolution direct photographs.
   - (b) Short-term time-lapse photographs in color, for observing spans up to 12 hours.

3. *From Probes*
   - (a) Reserved for detection of phenomena beyond present resolutions.

A. G. Wilson 6/19/68

PULSARS -- A SUMMARY

A new large radio telescope operating at 81.5 MH_3 was put into use by the Mullard Radio Astronomy Observatory of the University of Cambridge in July of 1967. The aerial consists of a rectangular array containing 2048 full wave dipoles arranged in 16 rows of 128 elements. The array is 470 meters (E-W) by 45 meters (N-S). This telescope was built to investigate the angular structure of compact radio sources using the scintillation caused by the interplanetary medium. A weekly survey of the sky between the declination zones -08° and +44° using this new telescope resulted in the detection of four very weak pulsating signals at fixed declinations and right ascensions. Systematic investigations of these signals were started in November of 1967 and the first publication of discovery appeared in Nature, vol. 217, p. 709, February 24, 1968. The observed properties of these sources -- now called "Pulsars" -- are summarized in the table.

No distances have been determined, but the observing of a doppler shift reflecting the earth's orbital motion places the pulsars definitely outside of the solar system. From frequency dependence of signal retardation and the value of interstellar electron density,the pulsars are estimated to be over 150 light years distant.

The precise periods afford many applications -- determination of the A. V., galactic rotation and magnetic field, time service, space navigation, etc.

PULSARS

Designation	Position	Galactic Coordinates	Pulse Period	Pulse Duration	Mean Flux Density
	α (1950.0) δ (1950.0)	λ II b II	seconds	milliseconds	@ 81.5MH3
CP.0834	$08^h\ 34^m\ 07^s$ + 07° 00'	220°	1.273 764 200 ± 300	∿ 35 ± 4	0.3
CP.0950	$09^h\ 50^m\ 29^s$ + 08° 10'	230° 44°	0.253 065 000 ± 100	∿ 15 ± 4	0.8
CP.1133	$11^h\ 33^m\ 32^s$ + 17° 00'	240° 70°	1.187 909 280 ± 150	∿ 35 ± 4	0.3
CP.1919	$19^h\ 19^m\ 37^s$ + 21° 47'	56° 4°	1.337 301 092 ± 2	∿ 37 ± 4	0.4

NOTES:

a) Mean flux density in units 10^{-26} watts m^{-2} H_3 -1

b) The pulsars have been observed at frequencies from 75.3 to 2700 MH_3

c) The fine structure of the pulses follows in general the pattern: single pip (CP.0950), double pip (CP.0834 and CP.1133), and triple pip (CP.1919).

d) The pulse duration, d, seems approximately to follow the law $d^2 = AT$ where T is the pulse period and A is about one millisecond.

PULSARS: BIBLIOGRAPHY TECHNICAL PAPERS

Nature	v.217	2/24/68	p.709-713	Hewish, Bell Pilkington, Scott, Collins "Observation of a Rapidly Pulsating Radio Source"
Nature	v.217	3/9/68	p.907-910	Ryle, Bailey, "Optical Identification of the First Neutron Star?"
			p.910-912	Davis, Horton, Lyne, Rickett, Smith, "Pulsating Radio Source at $\alpha = 19^h\ 19^m$, $\delta = +22°$"
I.A.U.Circular	No.2064	3/29/68		Hewish
Nature	v.217	3/30/68	p.1222-1227	Saslaw, Faulkner, Strittmatter, "Rapidly Pulsating Radio Sources"
			p.1227-1228	Ostriker, "Possible Model for a Rapidly Rotating Radio Source"
Ap.J.Letters	v.152	April'68	L71-L75	Thorne, Ipser, "White Dwarf and Neutron Star Interpretations of Pulsating Radio Sources"
Nature	v.218	4/13/68	p.123-124	Hoyle, Norlikar, "Pulsed Radio Sources"
			p.124-126	Lyne, Smith, "Linear Polarization in Pulsating Radio Sources"
			p.126-129	Pilkington, Hewish, Bell, Cole, "Observations of Some Further Pulsed Radio Sources"
			p.129-130	Bailey, Mackay, "Accurate Position of a Second Pulsed Radio Source"
Nature	v.218	4/20/68	p.227-229	Moffet, Ekers, "Detection of the Pulsed Radio Source CP.1919 at 13 am"
			p.229-230	Radhukpishnan, omesaroff, Cooke, "Measurements on the Period of the Pulsating Radio Source 1919"

Nature	v.218	4/20/68	p.230-232	Scott, Collins, "Further Observations of Pulsating Radio Sources"
Science	v.160	4/26/68	p.415-416	Duthrie, Sturch, Hafner, "Optical Pulse of a Periodic Radio Star"
			p.416-419	Drake, "Pulse Structure of the Pulsating Radio Source in Vulp"
Nature	v.218	4/27/68	p.326-330	Lyne, Rickett, "Measurement of Pulse, Shape, and Spectra of the Pulsating Radio Source"
I.A.U.Circular	No.2072	5/3/68		Moffet
Science	v.160	5/3/68	p.503-507	Drake, Gundemann, Jauncey, Comella, Zeissig, Craft, "The Rapidly Pulsating Radio Source in Vulpecala"
Science	v.160	5/17/68	p.758-760	Drake, Craft, "Pulse Structure of Four Pulsars"
			p.760	Tanenbaum, Zeissig, Drake, "Parameters of the Plasma Affecting the Radiation of Pulsar 1"
Science	v.160	6/7/68	p.1104-1105	Papalioliog, Carleton, Horowitz, Liller, "Optical Search for Pulsations of CP 1919"

AN EMPIRICAL RELATION IN PULSAR PERIODS

Pulse periods of the four known pulsars have been determined to an accuracy of the order of one part in 10^7 (Ref. 1, 2). The pulse durations cannot be nearly so accurately measured, but the mean pulse durations of sets of superimposed pulses can be estimated to within three or four milliseconds (Ref. 3, 4). To within the accuracies of these estimates a single simple empirical relation appears to hold between the pulse periods, T, and the mean pulse durations, $<d>$, for each of the four observed pulsars,

$$(1) \qquad AT = <d>^2, \qquad \text{where } A = 10^{-3} \text{ seconds.}$$

A comparison of values is shown in the table where $<d_o>$ is the approximate *observed* mean duration and d_c is the value given by Eq. (1). All values are in seconds.

OBJECT	T	$<d_o>$	d_c
CP. 0834	1.273764	0.035	0.0357
CP. 0950	0.253065	0.015	0.0159
CP. 1133	1.187909	0.035	0.0345
CP. 1919	1.337301	0.037	0.0366

The parameter A with the dimensionality of time defines a nearly constant *characteristic time* for objects of the pulsar class.

The correct identification of the observed periods with limiting periods ($\sqrt{3\pi/GP}$) for various classes of bodies is critical for formulation of the right pulsar model. The pulse period itself is too short for white dwarfs. The periods d and A are both consistent with the limiting period of neutron stars.

REFERENCES:

1. J. D. H. Pilkington, A. Hewish, J. J. Bell, T. W. Cole, Nature, vol. 218, p. 126, April 13, 1968.
2. A. T. Moffet, I. A. U. Circular No. 2072.
3. A. G. Lyne, B. J. Rickett, Nature, vol. 218, p. 326, April 27, 1968.
4. F. D. Drake, H. D. Craft, Jr., Science, vol. 160, p. 758, May 17, 1968.

Albert Wilson
Douglas Advanced Research Laboratories
5251 Bolsa
Huntington Beach, California 92647

A Dynamic Parallax of M 31

Albert G. Wilson

Mt. Wilson and Palomar Observatories
California Institute of Technology
Carnegie Institution of Washington

Abstract

A dynamical method, which is entirely free of any luminosity criteria, can be employed for determining the distance to the spiral galaxy M 31. It is assumed that spirals like our own galaxy and M 31 can be adequately represented in their first order dynamic features by Oort's model of a set of concentric spheroids. An equilibrium condition within the rotating spheroids gives a relation between the distance to the system; the velocity of rotation at a given angular distance from the center, the degree of flattening of the spheroids, and the density of the spheroids. On the assumption that the densities in the solidly rotating spheroids are of the same order in M 31 and our galaxy, a distance of 430 kiloparsecs (1,400,000 light years) is derived. This value is in good agreement with the latest photometric moduli, which double the earlier values of the distance.

A Dynamic Parallax of M 31

Albert G. Wilson

Mt. Wilson and Palomar Observatories
California Institute of Technology
Carnegie Institution of Washington

The concatenated set of distance criteria required for the law of red shifts is calibrated initially against the distances to nearby galaxies. It is, therefore, important to estimate these distances by as many different methods as possible. The most reliable methods have proved to be those based on photometric moduli of suitable types of variable and high luminosity stars. But methods depending on luminosity measurements are subject to uncertainties arising from unknown absorptions and uncertainties in the absolute magnitudes. Consequently, a method of approximating the distance to an external galaxy which is independent of luminosity considerations would be of interest.

Oepik (Ap. J. vol. 55, p. 406, 1922) has proposed a method which is based on dynamical features of a galaxy but which still involves the use of certain luminosity assumptions. However, by use of data now available, it is possible to estimate the distance of M 31 without recourse to the luminosity features.

Oort (B.A.N. vol IX, no. 338., p. 193, 19 and Ap. J. vol. 116, p. 233, 1952) has found that the salient dynamical properties of our galaxy are exhibited by a model which consists of a set of superimposed concentric spheroids, each possessing a uniform density throughout. The principal part of the galaxy, the part containing most of the mass, is adequately represented by two spheroids: A nuclear high density spheroid of low eccentricity and a flatter spheroid of lower density which extends almost to the neighborhood of the sun. To a first order of approximation these spheroids appear to rotate as solid bodies.

This configuration is stable and a simple equilibrium condition holds within the spheroids. Let V_c represent the circular velocity at a distance r from the axis of rotation. Let e_1, ρ_1, a_1, and e_2, ρ_2, a_2 be the eccentricity, density, and semi-major axis of the inner and outer spheroids respectively. Then on the equatorial plane, for $a_1 \leq r \leq a_2$

$$(1) \qquad \frac{V_c^2}{r} = 2\pi G r \left[\rho_1 J(e_1) S\left(\frac{a_1 e_1}{r}\right) + \rho_2 J(e_2) S(e_2)\right]$$

where $J(x) = \frac{\sqrt{1-x^2}}{x^3}$ and $S(x) = \sin^{-1}x - x\sqrt{1-x^2}$

and G is the gravitational constant.

If r and a_1 subtend angles α and α_1 respectively when viewed from a distance D, then equation (1) can be rewritten

$$(2) \qquad D = \frac{1260\, V_c}{\alpha\sqrt{\rho_1 J_1 S\left(\frac{\alpha_1 e_1}{\alpha}\right) + \rho_2 J_2 S(e_2)}}$$

where the units are D in kiloparsecs, V_c in kilometers/second, α and α_1 in seconds of arc, and ρ in solar masses per cubic parsec. Equation (2), thus gives the distance, D, to a galaxy of the type described, in terms of observables and the densities ρ_1 and ρ_2 of the two spheroids.

The researches of Baade and Mayall (Pub. Obs. Univ. of Mich., vol.X, 1951) indicate that our own galaxy and M 31 are quite similar. Both are Sb type spirals and have similar rotational properties. The dynamical features of M 31 can also be adequately represented by an Oort type model, and equation (2) may therefore be used for estimating its distance.

Babcock (L.O.B., 498, 1939) and Mayall (loc.cit.) have studied the motions of M 31 spectroscopically. The nuclear spheroid appears to extend only 4 or 5 minutes of arc from the axis. The second spheroid, as shown by a radially increasing approximately linear law of rotation, extends about 68 minutes from the axis. The measures in the outer spheroid are made on emission nebulosities. Since these objects are confined to the equatorial plane and rotate in nearly circular orbits, the measured velocities may be used for V_c in equation (2) with no further corrections than those for the relative motion of M 31 as a whole and for the inclination of the equatorial plane to the line of sight.

The ratio of the sizes of the two spheroids in M 31 is such that the function $S\left(\frac{a_1 e_1}{a}\right)$ is negligible except for a few minutes of arc immediately beyond its boundary. Therefore if V_c is taken as the maximum rotational velocity, which occurs at 68 minutes of arc, the first term in the radical can be dropped. This value is 331 km/sec when corrected for tilt.

The eccentricity of the second spheroid, e_2, can be derived from isophotal contours made by Hiltner and Williams (Pub. Obs. Univ. Of Mich., vol. VIII, no.7, 1941). When corrected for a tilt of 15°, e_2 has a mean value of 0.957. This leaves as the only unknown in the right member, the density ρ_2 which is unobservable. On the basis of the other similarities between M 31 and our galaxy, a reasonable first value to assume for the density would be that ascribed by Oort to the second spheroid in our own galaxy, a value about 2.15 times the density in the solar neighborhoor or 0.172 solar masses per cubic parsec. With the above values, the distance to M 31 comes out about 430 kiloparsecs.

Oort and his co-workers (M.N. vol. 106, p. 159, 1946) have shown that the gas densities conducive to the formation of dust particles and leading to absorption coefficients of the same magnitude as those observed, lie in a critical range centered about the value 0.13 solar masses per cubic parsec. Dust does not appear in the central regions of galaxies, but as manifested by the presence of spiral arms, first occurs in the second spheroids in both M 31 and our galaxy. This fact makes it again reasonable to assume that the gas densities in the second spheroids of both galaxies have had the same values - in the critical range - during their histories. Although there is no a priori reason for requiring the star densities (and hence the total densities) to be equal if the gas densities are equal, the hypothesis that galaxies acquired their present dynamical structures when still gaseous, (as applied to ellipticals by Belzer, Gamow and Keller (Ap.J. vol. 113, p. 166, 1951)), would be consistent with equal total densities in zones of equal gas densities.

The above value of 430 kpc is in good agreement with the latest photometrically determined distances to M 31. Baade announced at the recent I.A.U. meeting in Rome that the zero point of the Cepheid Period-Luminosity Law should be made 1.5 magnitudes brighter. This doubles the old distance of 230 kpc. (Baade, Ap. J. vol. 100, p. 137, 1944). Measurements of Thackery in the Small Magellanic Cloud bear out this correction.

The hypothetical character of the dynamical method of estimating the parallax vitiates assigning a probable error to the result. The agreement with photometric results should not be taken as a confirmation of any of the individual assumptions embodied in the method, but rather as a suggestion that a dynamic method for estimating parallaxes of galaxies might be further explored with profit as the observational data becomes available.

Dynamic Parallaxes of Extra-Galactic Nebulae

Albert G. Wilson
Lowell Observatory

Because of the many uncertainties present in the determination of the distances of extra-galactic nebulae from photometric criteria, it is desirable that independent methods of distance determination be developed. Recently a correction of the value of the absolute magnitudes of the classical cepheids led to a change in the value of the distance to M 31 by a factor of approximately two. This new distance of M 31 and the distance of any other galaxy in which the magnitudes of cepheids can be measured will probably now be free of most errors, except possibly unknown absorption effects. But distances to galaxies in which the magnitudes of cepheids cannot be measured will remain uncertain until secondary photometric criteria such as those of novae, brightest stars, and total magnitudes of the galaxies themselves can be accurately calibrated. This calibration process is a lengthy one and it will probably be several years before a definitive photometric modulus of the Virgo Cluster, which is important as the base from which more distant studies must be made, will be available. Hence a method of determining distances which is independent of luminosity criteria and which can be used immediately for distances out to the Virgo Cluster, is adequately justified. Such a method is outlined in the following:

It will be assumed that certain galaxies such as Sa's and Sb's, can be approximated by a set of concentric homogeneous spheroids. Oort has shown [1] that the salient dynamic features of our own galaxy can be well represented by a set of seven such spheroids, with the principal part of the galaxy adequately represented by the two innermost spheroids. To a first approximation these spheroids appear to rotate as solid bodies; and

though assumptions of uniform density and solid body rotation are not exact, they are good approximations and are dynamically consistent. It can be shown that, at distances from the center of about twice the semimajor axis of the inner ellipsoid, its effect may be neglected and the dynamic manifestations in the outer parts of the second spheroid resemble those of the second spheroid taken alone. Further, the effect of the outer spheroid on the properties of the inner can, for present purposes, be allowed for by a modification of the density. Thus, either the inner or outer spheroid may be employed in the following considerations:

An equilibrium condition between the eccentricity, e; angular velocity, ; and density, ; may be derived for homogeneous solidly rotating spheroids. This condition, due to Maclaurin, is:

1) $$\frac{\quad}{2 \quad G} = \frac{3 - 2e}{e} \quad 1 - e \quad \sin \quad e - \frac{3}{e} (1 - e) = F (e)$$

Where G is the gracitational constant.

From equation 1), since $= \frac{v}{D}$, where v is the linear velocity of rotation at a distance from the center, we can derive for the distance,

2) $$D = 1.26 \quad \frac{v}{\quad}$$

Where D is in megaparsecs, v in kilometers/second, in seconds of arc, and in solar masses per cubic parsec.

The quantities in the right member of equation 2) are observables except the density, . The nearly linear relation between v and manifested by the spectra of most galaxies allows these quantities to be replaced in many cases by the observed inclination of the spectral lines. Even in the case of rotational data derived from measured velocities of individual emmission objects, it is only necessary to have knowledge of the spectra over an

interval sufficient to establish a mean slope.

The ellipticities of galaxies are best observed by an isophotometer. It is well known that ellipticities vary with the size of the isophote, and there is enough change to vitiate the entire method, except that the function F(e) is very insensitive to e. In fact, for the entire range of ellipticities ordinarily encountered F(e) varies by less than four per cent.

The matter of the density is more difficult. Since it is not an observable, and since at present there is not enough data to explore possible supplementary relationships which could be applied, as is the mass luminosity law for dynamic parallaxes of double stars, it is necessary to introduce a further assumption that limits the present application of the method to giant Sa and Sb systems, viz. the densities of such systems are the same as in our own galaxy. The striking result which emerges from this assumption is that these giant systems (such as M 31, M 81, NGC 4594, etc.) all possess nearly the same absolute magnitude (about -19.1). Or conversely, the assumption of equal densities is valid, if the brightest galaxies are bounded by a limiting luminosity.

It is interesting to compare the values for the distances of galaxies derived under the foregoing assumptions from equation 2) with the new photometric distances.

As the first example, consider M 31 itself. Using the outer spheroid delineated by Babcock's 2 curve showing a linear rotation based on the spectra emission objects, there is a velocity of 328 km/sec at 68 minutes of arc from the center (corrected for a 15° tilt). From Hiltner and Williams isophotal data [3], the ellipticity of the principal spheroid is about 1:3. Finally assuming the density in this spheroid to be the same as in Oort's [1]

second spheroid in our galaxy, we have = 0.218 solar masses/cubic parsec. These values give D = 487 kiloparsecs, corresponding to a value of 23.4 for the distance modulus. Baade [4] now gives 23.5 or 501 kiloparsecs as the best photometric distance to M 31. (This includes his correction of 0.4 magnitude for absorption in our own system). The difference of 2.8% is perhaps much too satisfactory considering the uncertainties involved.

The available rotational and isophotal data is limited, but Slipher's unpublished spectra of NGC 4594 affords us the opportunity to derive a preliminary distance to the Virgo Cluster. This is best done by getting the distance of NGC 4594 relative to M 31. In view of the above considerations with regard to the ellipticities and densities for giant spirals, their relative distances will be in proportion to the inclination of their spectral lines.

$$D_{4594} = \frac{0.487 \quad {}_{4594}}{{}_{M\,31}} \text{ megaparsecs}$$

Since the spectra of more distant galaxies show only the rotation of the nuclear regions from inclinations of absorption lines, the 's must be compared with the rotational conditions at the nucleus of M 31. Pease [5] and Babcock [2] both find near the center of M 31 from the absorption spectra a value for $_{M\ 31}$ of 0.467 km/sec per second of arc.

Slipher's value of for NGC 4594 is 5.0 km/sec per second arc, leading to a distance of about 5.2 megaparsecs, or a value of 28.6 for the modulus. Baade's tentative correction for galaxies beyond the local group, as reported by Sandage at the last AAS meeting, is -2.1 magnitudes. Correcting Hubble's old modulus of 26.7 for the Virgo Cluster, we get 28.8. Again the agreement is not unsatisfactory. These results suggest that the doubling of distances found necessary in the local group is inadequate for more remote objects.

The factor may be closer to two and one-half.

The mean value of the red-shift for the Virgo Cluster is 1100 km/sec. The derived value of Hubble's constant from the above distance is 211 km/sec/megaparsec, corresponding to 4.6 x 109 years (Hubble's original value was 525 km/sec/megaparsec). This new value based on the dynamic method lies in the interval 125 to 276 km/sec/megaparsec given by Sandage as the probable bounds for Hubble's constant based on present photometric data.

However tentative any of these results may be at the present time, there is certainly encouragement from the agreement of results and the dynamic method might be profitably refined as more observational data becomes available. But not only is the dynamic method of value in that it affords an independent check on photometric distances, but it allows possible study of the relative distances of unresolved galaxies which could never be treated by the classical methods.

[1] OOrt, Ap.J. v. 116, p. 233ff, 1952

[2] Babcock, L.O.B., 498, 1939

[3] Hiltner and Williams, Pub. Obs. U. of Mich., v. 8, No. 7, 1941

[4] Baade, Symposium on Astrophysics, U. of Mich., p. 17, 1953

[5] Pease, P.N.A.S., v. 4, p. 21ff, 1918

Dynamic Parallaxes of Extra-Galactic Nebulae

Albert G. Wilson
Lowell Observatory

Because of the many uncertainties present in the determination of the distances of extra-galactic nebulae from photometric criteria, it is desirable that independent methods of distance determination be developed. Recently a correction of the value of the absolute magnitudes of the classical cepheids led to a change in the value of the distance to M 31 by a factor of approximately two. This new distance of M 31 and the distance of any other galaxy in which the magnitudes of cepheids can be measured will probably now be free of most errors, except possibly unknown absorption effects. But distances to galaxies in which the magnitudes of cepheids cannot be measured will remain uncertain until secondary photometric criteria such as those of novae, brightest stars, and total magnitudes of the galaxies themselves can be accurately calibrated. This calibration process is a lengthy one and it will probably be several years before a definitive photometric modulus of the Virgo Cluster, which is important as the base from which more distant studies must be made, will be available. Hence a method of determining distances which is independent of luminosity criteria and which can be used immediately for distances out to the Virgo Cluster, is adequately justified. Such a method is outlined in the following:

It will be assumed that certain galaxies such as Sa's and Sb's, can be approximated by a set of concentric homogeneous spheroids. Oort has shown [1] that the salient dynamic features of our own galaxy can be well represented by a set of seven such spheroids, with the principal part of the galaxy adequately represented by the two innermost spheroids. To a first approximation these spheroids appear to rotate as solid bodies; and

though assumptions of uniform density and solid body rotation are not exact, they are good approximations and are dynamically consistent. It can be shown that, at distances from the center of about twice the semimajor axis of the inner ellipsoid, its effect may be neglected and the dynamic manifestations in the outer parts of the second spheroid resemble those of the second spheroid taken alone. Further, the effect of the outer spheroid on the properties of the inner can, for present purposes, be allowed for by a modification of the density. Thus, either the inner or outer spheroid may be employed in the following considerations:

An equilibrium condition between the eccentricity, e; angular velocity, ; and density, ; may be derived for homogeneous solidly rotating spheroids. This condition, due to Maclaurin, is:

1) $$\frac{\quad}{2\quad G} = \frac{3 - 2e}{e} \qquad 1 - e \qquad \sin\ e - \frac{3}{e}(1 - e) = F(e)$$

Where G is the gracitational constant.

From equation 1), since $= \frac{v}{D}$, where v is the linear velocity of rotation at a distance from the center, we can derive for the distance,

2) $$D = 1.26 \ \frac{v}{\quad}$$

Where D is in megaparsecs, v in kilometers/second, in seconds of arc, and in solar masses per cubic parsec.

The quantities in the right member of equation 2) are observables except the density, . The nearly linear relation between v and manifested by the spectra of most galaxies allows these quantities to be replaced in many cases by the observed inclination of the spectral lines. Even in the case of rotational data derived from measured velocities of individual emmission objects, it is only necessary to have knowledge of the spectra over an

interval sufficient to establish a mean slope.

The ellipticities of galaxies are best observed by an isophotometer. It is well known that ellipticities vary with the size of the isophote, and there is enough change to vitiate the entire method, except that the function F(e) is very insensitive to e. In fact, for the entire range of ellipticities ordinarily encountered F(e) varies by less than four per cent.

The matter of the density is more difficult. Since it is not an observable, and since at present there is not enough data to explore possible supplementary relationships which could be applied, as is the mass luminosity law for dynamic parallaxes of double stars, it is necessary to introduce a further assumption that limits the present application of the method to giant Sa and Sb systems, viz. the densities of such systems are the same as in our own galaxy. The striking result which emerges from this assumption is that these giant systems (such as M 31, M 81, NGC 4594, etc.) all possess nearly the same absolute magnitude (about -19.1). Or conversely, the assumption of equal densities is valid, if the brightest galaxies are bounded by a limiting luminosity.

It is interesting to compare the values for the distances of galaxies derived under the foregoing assumptions from equation 2) with the new photometric distances.

As the first example, consider M 31 itself. Using the outer spheroid delineated by Babcock's 2 curve showing a linear rotation based on the spectra emission objects, there is a velocity of 328 km/sec at 68 minutes of arc from the center (corrected for a 15° tilt). From Hiltner and Williams isophotal data [3], the ellipticity of the principal spheroid is about 1:3. Finally assuming the density in this spheroid to be the same as in Oort's [1]

second spheroid in our galaxy, we have = 0.218 solar masses/cubic parsec. These values give D = 487 kiloparsecs, corresponding to a value of 23.4 for the distance modulus. Baade [4] now gives 23.5 or 501 kiloparsecs as the best photometric distance to M 31. (This includes his correction of 0.4 magnitude for absorption in our own system). The difference of 2.8% is perhaps much too satisfactory considering the uncertainties involved.

The available rotational and isophotal data is limited, but Slipher's unpublished spectra of NGC 4594 affords us the opportunity to derive a preliminary distance to the Virgo Cluster. This is best done by getting the distance of NGC 4594 relative to M 31. In view of the above considerations with regard to the ellipticities and densities for giant spirals, their relative distances will be in proportion to the inclination of their spectral lines.

$$D_{4594} = \frac{0.487 \quad 4594}{M\ 31} \text{ megaparsecs}$$

Since the spectra of more distant galaxies show only the rotation of the nuclear regions from inclinations of absorption lines, the 's must be compared with the rotational conditions at the nucleus of M 31. Pease [5] and Babcock [2] both find near the center of M 31 from the absorption spectra a value for $_{M\ 31}$ of 0.467 km/sec per second of arc.

Slipher's value of for NGC 4594 is 5.0 km/sec per second arc, leading to a distance of about 5.2 megaparsecs, or a value of 28.6 for the modulus. Baade's tentative correction for galaxies beyond the local group, as reported by Sandage at the last AAS meeting, is -2.1 magnitudes. Correcting Hubble's old modulus of 26.7 for the Virgo Cluster, we get 28.8. Again the agreement is not unsatisfactory. These results suggest that the doubling of distances found necessary in the local group is inadequate for more remote objects.

The factor may be closer to two and one-half.

The mean value of the red-shift for the Virgo Cluster is 1100 km/sec. The derived value of Hubble's constant from the above distance is 211 km/sec/megaparsec, corresponding to 4.6×10^{9} years (Hubble's original value was 525 km/sec/megaparsec). This new value based on the dynamic method lies in the interval 125 to 276 km/sec/megaparsec given by Sandage as the probable bounds for Hubble's constant based on present photometric data.

However tentative any of these results may be at the present time, there is certainly encouragement from the agreement of results and the dynamic method might be profitably refined as more observational data becomes available. But not only is the dynamic method of value in that it affords an independent check on photometric distances, but it allows possible study of the relative distances of unresolved galaxies which could never be treated by the classical methods.

[1] OOrt, Ap.J. v. 116, p. 233ff, 1952

[2] Babcock, L.O.B., 498, 1939

[3] Hiltner and Williams, Pub. Obs. U. of Mich., v. 8, No. 7, 1941

[4] Baade, Symposium on Astrophysics, U. of Mich., p. 17, 1953

[5] Pease, P.N.A.S., v. 4, p. 21ff, 1918

Dynamic Parallaxes of Extra-Galactic Nebulae

Albert G. Wilson
Lowell Observatory

Because of the many uncertainties present in the determination of the distances of extra-galactic nebulae from photometric criteria, it is desirable that independent methods of distance determination be developed. Recently a correction of the value of the absolute magnitudes of the classical cepheids led to a change in the value of the distance to M 31 by a factor of approximately two. This new distance of M 31 and the distance of any other galaxy in which the magnitudes of cepheids can be measured will probably now be free of most errors, except possibly unknown absorption effects. But distances to galaxies in which the magnitudes of cepheids cannot be measured will remain uncertain until secondary photometric criteria such as those of novae, brightest stars, and total magnitudes of the galaxies themselves can be accurately calibrated. This calibration process is a lengthy one and it will probably be several years before a definitive photometric modulus of the Virgo Cluster, which is important as the base from which more distant studies must be made, will be available. Hence a method of determining distances which is independent of luminosity criteria and which can be used immediately for distances out to the Virgo Cluster, is adequately justified. Such a method is outlined in the following:

It will be assumed that certain galaxies such as Sa's and Sb's, can be approximated by a set of concentric homogeneous spheroids. Oort has shown [1] that the salient dynamic features of our own galaxy can be well represented by a set of seven such spheroids, with the principal part of the galaxy adequately represented by the two innermost spheroids. To a first approximation these spheroids appear to rotate as solid bodies; and

though assumptions of uniform density and solid body rotation are not exact, they are good approximations and are dynamically consistent. It can be shown that, at distances from the center of about twice the semimajor axis of the inner ellipsoid, its effect may be neglected and the dynamic manifestations in the outer parts of the second spheroid resemble those of the second spheroid taken alone. Further, the effect of the outer spheroid on the properties of the inner can, for present purposes, be allowed for by a modification of the density. Thus, either the inner or outer spheroid may be employed in the following considerations:

An equilibrium condition between the eccentricity, e; angular velocity, ; and density, ; may be derived for homogeneous solidly rotating spheroids. This condition, due to Maclaurin, is:

1) $$\frac{\quad}{2\quad G} = \frac{3 - 2e}{e} \qquad 1 - e \qquad \sin\ e - \frac{3}{e}(1 - e) = F\ (e)$$

Where G is the gracitational constant.

From equation 1), since $= \frac{v}{D}$, where v is the linear velocity of rotation at a distance from the center, we can derive for the distance,

2) $$D = 1.26 \ \frac{v}{\quad}$$

Where D is in megaparsecs, v in kilometers/second, in seconds of arc, and in solar masses per cubic parsec.

The quantities in the right member of equation 2) are observables except the density, . The nearly linear relation between v and manifested by the spectra of most galaxies allows these quantities to be replaced in many cases by the observed inclination of the spectral lines. Even in the case of rotational data derived from measured velocities of individual emmission objects, it is only necessary to have knowledge of the spectra over an

interval sufficient to establish a mean slope.

The ellipticities of galaxies are best observed by an isophotometer. It is well known that ellipticities vary with the size of the isophote, and there is enough change to vitiate the entire method, except that the function F(e) is very insensitive to e. In fact, for the entire range of ellipticities ordinarily encountered F(e) varies by less than four per cent.

The matter of the density is more difficult. Since it is not an observable, and since at present there is not enough data to explore possible supplementary relationships which could be applied, as is the mass luminosity law for dynamic parallaxes of double stars, it is necessary to introduce a further assumption that limits the present application of the method to giant Sa and Sb systems, viz. the densities of such systems are the same as in our own galaxy. The striking result which emerges from this assumption is that these giant systems (such as M 31, M 81, NGC 4594, etc.) all possess nearly the same absolute magnitude (about -19.1). Or conversely, the assumption of equal densities is valid, if the brightest galaxies are bounded by a limiting luminosity.

It is interesting to compare the values for the distances of galaxies derived under the foregoing assumptions from equation 2) with the new photometric distances.

As the first example, consider M 31 itself. Using the outer spheroid delineated by Babcock's 2 curve showing a linear rotation based on the spectra emission objects, there is a velocity of 328 km/sec at 68 minutes of arc from the center (corrected for a 15° tilt). From Hiltner and Williams isophotal data [3], the ellipticity of the principal spheroid is about 1:3. Finally assuming the density in this spheroid to be the same as in Cort's [1]

second spheroid in our galaxy, we have = 0.218 solar masses/cubic parsec. These values give D = 487 kiloparsecs, corresponding to a value of 23.4 for the distance modulus. Baade [4] now gives 23.5 or 501 kiloparsecs as the best photometric distance to M 31. (This includes his correction of 0.4 magnitude for absorption in our own system). The difference of 2.8% is perhaps much too satisfactory considering the uncertainties involved.

The available rotational and isophotal data is limited, but Slipher's unpublished spectra of NGC 4594 affords us the opportunity to derive a preliminary distance to the Virgo Cluster. This is best done by getting the distance of NGC 4594 relative to M 31. In view of the above considerations with regard to the ellipticities and densities for giant spirals, their relative distances will be in proportion to the inclination of their spectral lines.

$$D_{4594} = \frac{0.487 \quad 4594}{M\ 31} \text{ megaparsecs}$$

Since the spectra of more distant galaxies show only the rotation of the nuclear regions from inclinations of absorption lines, the 's must be compared with the rotational conditions at the nucleus of M 31. Pease [5] and Babcock [2] both find near the center of M 31 from the absorption spectra a value for $_{M\ 31}$ of 0.467 km/sec per second of arc.

Slipher's value of for NGC 4594 is 5.0 km/sec per second arc, leading to a distance of about 5.2 megaparsecs, or a value of 28.6 for the modulus. Baade's tentative correction for galaxies beyond the local group, as reported by Sandage at the last AAS meeting, is -2.1 magnitudes. Correcting Hubble's old modulus of 26.7 for the Virgo Cluster, we get 28.8. Again the agreement is not unsatisfactory. These results suggest that the doubling of distances found necessary in the local group is inadequate for more remote objects.

The factor may be closer to two and one-half.

The mean value of the red-shift for the Virgo Cluster is 1100 km/sec. The derived value of Hubble's constant from the above distance is 211 km/sec/megaparsec, corresponding to 4.6 x 109 years (Hubble's original value was 525 km/sec/megaparsec). This new value based on the dynamic method lies in the interval 125 to 276 km/sec/megaparsec given by Sandage as the probable bounds for Hubble's constant based on present photometric data.

However tentative any of these results may be at the present time, there is certainly encouragement from the agreement of results and the dynamic method might be profitably refined as more observational data becomes available. But not only is the dynamic method of value in that it affords an independent check on photometric distances, but it allows possible study of the relative distances of unresolved galaxies which could never be treated by the classical methods.

[1] OOrt, Ap.J. v. 116, p. 233ff, 1952

[2] Babcock, L.O.B., 498, 1939

[3] Hiltner and Williams, Pub. Obs. U. of Mich., v. 8, No. 7, 1941

[4] Baade, Symposium on Astrophysics, U. of Mich., p. 17, 1953

[5] Pease, P.N.A.S., v. 4, p. 21ff, 1918

OLBERS' PARADOX AND COSMOLOGY

A. G. Wilson

October 1965

P-3256

OLBERS'PARADOX AND COSMOLOGICAL MODELS

A. G. Wilson

The RAND Corporation,* Santa Monica, California

Dr. Watson used to say that nothing provided his friend Sherlock Holmes with more satisfaction than the challenge of determining a stranger's business, background, habits, and history from only a few seconds of observation. There is an intriguing example of a similar challenge in astronomy -- a challenge which I am sure would have delighted Holmes had he encountered it. I refer to the possibility of deducing the nature of the entire universe, including that it is expanding, from an observation which consists of no more than looking up at the sky at night and noting that it is dark. Watson might refer to this startling set of deductions as the case of the Paradox of the German Physician. The year was 1826, the place was Hamburg. In a rather obscure journal called Bode's Jahrbuch, a local physician and amateur astronomer named Heinrich Olbers published an account of how he could use the brightness of the night sky alone as a fundamental clue from which he could derive the nature of the universe.

Although what has become known as "Olbers' Paradox" is somewhat dated, it is still of primary relevance to cosmology, and I introduce it not only to review it as an important contribution to modern cosmology, but also in order to make use of it for a generic classification of cosmological models.

---

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This paper was an invited lecture presented to the Los Angeles Astronomical Society at the Griffith Observatory, June 14, 1965.

Olbers' deductions provide an ideal example of the application of the scientific method. First, a hypothesis is assumed. It is then tested by a simple observation, and if necessary, rejected. Alternatives are then formulated, and the possibilities consistent with observation are narrowed down. Finally, additional observations for determining the validity of the remaining hypotheses are designed. Olbers did not succeed in carrying through the entire argument, but in hindsight we can see how he could have constructed the argument -- or rather perhaps how Holmes might have constructed it.

What was this piece of cosmological detective work done by Olbers? In brief, if the stars are more or less of the same intrinsic brightness and are distributed more or less uniformly, we would expect two things: First, that their apparent luminosities would vary inversely with the square of their distances ($1/d^2$) and second, since the volume of a spherical shell is proportional to $4\pi d^2$, that the number of stars in a given distance interval would vary directly with the square of the distance. Hence the contribution to the brightness of the sky of all the stars in a shell at a given distance is essentially a constant, independent of distance (d^2 in numerator cancels d^2 in denominator). The total brightness of the sky is then given by the sum of the contributions of each shell. If b is the contribution of brightness per square degree from one shell, the brightness from n shells will be, nb. Hence for a uniform universe, n is very large and the sky brightness should be infinite. More precisely, because some of the stars will intercept and block off the light from more distant stars, the sky should be as bright as an average star or near in brightness to the disk of the sun. But a brief glimpse at the night sky shows it to be dark, not bright like the sun. Hence, Olbers concluded there is a paradox. We assume the simplest thing about the universe -- namely, that it is everywhere the same as it is locally. We reach a conclusion in contradiction with the simplest observation -- namely, that the sky is dark. The assumption then must, in some way, be wrong.

In order to determine the reason or reasons for this paradox, let us examine in more detail the assumptions -- both overt and tacit -- that Olbers really made.

Olbers, in effect, made five basic assumptions that are summarized as follows:

(1) The physical laws derived from terrestrial experience apply throughout the universe.

(2) The universe is homogeneous when viewed on a large scale.

(3) The universe is unchanging in time when viewed on a large scale.

(4) There are no major systematic motions in the universe.

(5) There is no interstellar fog absorbing starlight.

The first assumption is quite reasonable, and there is much evidence to substantiate it. Spectral investigations of stars, nearby and remote, show that their chemistry and physics is the same as the chemistry and physics we are familiar with locally. Stars are composed of the same atoms possessing the same energy levels as terrestrial atoms. This is not only true of stars in our galaxy but is also true of stars in other galaxies. In fact, the same resonance line of hydrogen seen in the ultraviolet spectrum of the sun has recently been detected in the most remote object known in the universe -- the Lyman α line in the quasar 3C9 whose redshift is greater than 2.[1] Furthermore, investigations of binary stars show the same radiation laws and gravitational laws that apply in the solar system are valid outside the solar system. The universality of local physical laws is also confirmed by observations of radio astronomy.

Olbers' second assumption means that the types of stars, their luminosities, and their average separations are, on the whole, the same everywhere. Today in discussing cosmological models we prefer to take <u>galaxy</u> as the basic cosmic unit rather than <u>star</u>, but the concept of homogeneity is the same.

Olbers' third assumption was not explicit. We know that because of the finite velocity of light, we see the more distant parts of the universe as they were in earlier times. Olbers' homogeneity assumption that the distant parts of the universe are the same as the nearby parts thus has implicit in it the concept of invariance of stellar luminosities in time.

Olbers' fourth assumption is implicit in the static nature of his analysis. No stars move between his shells of different distance. Finally Olbers assumed there was no absorbing material present that would reduce the light received from the stars, whatever their distance.

On the basis of these five assumptions, Olbers readily derived by the line of reasoning of adding the contributions of all shells, that the brightness of the sky should be something of the order of the brightness of the solar disk. Since this violently contradicts the observational facts, one or more of the five basic assumptions must be wrong. (Arguments might be made that a non-Euclidean geometry would account for the paradox. But the modifications of replacing the d^2 of Euclidean models with some other function of d also cancel from numerator and denominator and result in contributions from each shell that are again independent of distance.) Olbers concluded that it was the fifth assumption that must be in error. There probably existed unobservable interstellar material which diminished the flux of radiation and cut the sky brightness down to the value observed. Olbers was happy with this explanation and dropped the question, taking up the astronomical fad of the times -- comet chasing. Consequently, as Bondi points out,(2) Olbers missed the opportunity to have made the prediction of the age -- the expanding universe. Let us imagine how Sherlock Holmes would have persisted to a correct solution.

Actually, if the first four assumptions are valid, then the fifth assumption can have nothing to do with the paradox. If absorbing matter were present during the very long time allowed in the third assumption, this matter would have reached thermal equilibrium and would reradiate as much as it absorbed so that all absorbing matter would become as bright as the stars. The resolution of the paradox must accordingly depend on the error of one or more of the first four assumptions. Let us next investigate assumption number (4): There are no major systematic motions in the universe.

If we are to retain the first three assumptions, the question arises: What systematic motions are compatible with homogeneity and the preservation of homogeneity in time? The only motions possible which preserve homogeneity are those along the lines connecting all pairs of bodies that are also proportional to the separation of the bodies. This is more clearly seen when we consider that motion which preserves homogeneity must preserve the ratios of the distances between every pair of bodies.* Motions between two bodies proportional to their separation and along the line joining them results in either a uniform contraction or a uniform expansion of the system. The rate of expansion or contraction is given by the ratio of relative velocity to distance of separation and is the same for all pairs of bodies. Under assumptions one and two this rate may vary in time, but if assumption three is valid, this rate must be constant.**

Given then, that the only motions compatible with homogeneity and the preservation of homogeneity are a uniform expansion or contraction, we must determine whether either of these motions can resolve Olbers' Paradox.

If an emitting source of light is moving away from the observer, the total number of photons received per unit time is reduced in comparison with the number received from the same source when it is relatively stationary. That is to say, the total intensity of the contribution to the sky brightness of an expanding shell of stars is less than the contribution from a stationary shell by a factor depending on the velocity of recession of the shell, viz. $(1 + z)^{-2}$ where z is the redshift $\delta\lambda/\lambda$. Since the more distant shells will be

*Motions induced by gravity are proportional to $1/\sqrt{d}$ and thus generate inhomogeneity; centrifugal force results in velocities proportional to center distance, but these motions are not along lines joining all pairs of bodies and the result is a flattening of the system.

**A uniform contraction or expansion will, of necessity, increase or decrease the mean density of matter and will, therefore, be contrary to assumption number (3). Thus in order to preserve the validity of number (3), some additional assumption such as annihilation or creation of matter must be introduced. A weaker assumption than number (3), which we might call (3a), would require only that homogeneity be preserved in time. Assumptions (1), (2), and (3a) are sufficient conditions for uniform expansion or contraction.

receding faster, their contribution is more greatly reduced. Accordingly, when we sum the contributions from each shell, we find that instead of n times b, we have $b_1 + b_2 + b_3 + \ldots b_i$, with each successive b smaller than the preceding. In the language of mathematics the series converges. It is no longer infinite but has a bounded sum. The sum should equal the observed value of the sky brightness, which would have different values depending on the details of the models. If, on the other hand, the universe were contracting, each successively more distant shell would contribute greater light than in the static universe, and the paradox would be still unresolved.

We may then conclude, if assumptions (1) and (2) are valid, and if (3) is valid in either its stringent form or as assumption (3a), the universe must be expanding. In other words, if the universe is everywhere the same as we know it "locally," and if the sky is "dark," the universe must be expanding. This argument has here been presented as a qualitative argument only. The conclusions depend, of course, on quantitative justification.

Today we know the universe is expanding. The observational work of Slipher, Hubble, and Humason established this during the third decade of this century. The expansion, however, was not theoretically predicted, except contemporarily with its observational discovery, and this through the relativistic model of de Sitter and not on the basis of Olbers' Paradox. Thus exactly 100 years later (in 1926), the magnitude--redshift relation observationally refuted assumption number (4), and the paradox found a possible explanation.

It is a temptation now to conclude that the culprit is assumption number (4) and that the remaining three assumptions are valid. But even if assumption number (4) quantitatively can account for the darkening, the logic of the paradox does not permit this conclusion. Any of the first three assumptions, singly or in combination, may be wrong. Holmes would say that number (4) may have had an accomplice in darkening the sky.

Let us then look at the other assumptions and see what the implications of their validity might be. Bondi, Gold, and Hoyle, the creators of continuous creation, or the steady-state model as it is usually called, have given prestigious names to assumptions (2) and (3). Number (2), the assumption that the universe is homogeneous when viewed on a large scale, is called the Cosmological Principle. It is the result of a metaphysical flight from geocentricity and anthropocentricity. It may be paraphrased to state that the universe would appear the same to all observers no matter where they are located. With this assurance, we may assume that our view of the universe is typical and extrapolate with impunity. To number (3), the assumption that the universe is unchanging in time when viewed on a large scale, they give the name Perfect Cosmological Principle. This may be paraphrased to state that the universe would appear the same to all observers no matter where they observe or when they make their observation. If one assumes (1), (2), and (3) to be valid, together with the expansion of the universe, the cosmological model one comes up with is inevitably the steady-state model. Holmes would conclude, if number (4) had no accomplices, the universe conforms to the steady-state model. In this model the decreasing density caused by expansion is compensated for by the creation of new matter. Hence all properties, including density, remain constant in time.

If number (3) happens to be an accomplice to number (4), we could not be living in an unchanging universe such as that described by the steady-state model. For instance, if by looking back in time, the stars are systematically fainter, then evolutionary effects would also be contributing to the altered value of brightness of the night sky. This combination of (1) and (2) true and (3) false leads to a family of so-called evolutionary relativistic models, popularly referred to as "Big Bang Models." Unlike the unique steady-state model, there are a great many possible evolutionary models; some with positive curvature closed like a sphere; others with negative curvature infinite and open like a saddle. This is not the occasion to describe the detailed properties of these models. We are only trying to point out broad generic differences in the models.

In the models which assume the validity of number (2), the homogeneity is interpreted in such a way as to allow the actual distribution of matter in the universe to be approximated by a uniform perfect fluid.

This approximation affords a mathematical simplification of the relativistic field equations that are otherwise nearly intractable. Later we shall return to the question of homogeneity and uniform perfect fluids, but let us now conclude the classification of models which may be derived from the Olbers' assumptions.

What about assumption number (1), the universality of the laws of physics, as we observe them here and now? There have been cosmological models which infer the variation of basic physical parameters, such as the gravitational constant, G. Examples are the models of Dirac and Jordan which would be classified as holding assumption (1) as false.

It was pointed out earlier that the laws of chemistry and physics seem to be invariant in space and time. But observation has also shown us that the laws of physics may not be extrapolatable in scale, e.g., classical mechanics is not valid on atomic scales. It should not be surprising that classical mechanics may fail on a cosmic scale. One rationale for applying relativistic mechanics to cosmology rests on this question; however, we must remember that the scales over which relativistic mechanics is valid also have not yet been established. The three Schwarzschild tests apply to scales on the order of stellar diameters and planetary orbits. The proper mechanics valid for cosmic distances and times is still open to exploration.

We can summarize the eight possible combinations of the three remaining assumptions in the following classification scheme.

GENERIC CLASSIFICATION OF EXPANDING COSMOLOGICAL MODELS

1. Universality of physical laws
2. Cosmological principle
3. Perfect cosmological principle

True	False	Model
1,2,3	-	Steady-State
1,2	3	Evolutionary
1,3	2	inconsistent*
2,3	1	inconsistent*
1	2,3	Lambert-Charlier
2	1,3	Dirac-Jordan
3	1,2	inconsistent
-	1,2,3	?

*If (3) is valid, the validity of (1) and (2) are implied.

The validity of the Perfect Cosmological Principle implies the validity of the Cosmological Principle since all instants of time, including the present, are governed by number (3). The validity of number (3) also rules out secular changes of the laws of physics. Hence, number (3) being valid implies that number (1) and number (2) are both valid. We are accordingly left with five cosmological possibilities in an expanding universe. Of these, if (1), (2), and (3) are all false, then scientific cosmology becomes excessively speculative and uncertain. This leaves, in addition to the systems already described, the combination: (1) true; (2) and (3) false.

Considering this case, we ask what kind of universe would result if assumption (1) is valid, and yet every observer at every time would not necessarily observe the same thing (e.g., assumptions (2) and (3) false). One answer is found in the work of C. V. L. Charlier, a Swedish mathematician and astronomer. Charlier, in 1921,[3] proposed a solution to Olbers' Paradox not necessarily involving expansion, but rather that the matter of the universe was distributed into aggregates and clusters of aggregates. We said earlier that if we were to repeat Olbers' analysis today, instead of taking a star as our fundamental unit, we would take a galaxy but that the Olbers' argument would go through in the same way. This is true with one modification: Namely, if we use galaxies instead of stars as the fundamental building blocks of the universe, then Olbers' arguments lead to a brightness of the night sky equal not to the surface brightness of the sun but to the surface brightness of the Milky Way, or an average galaxy. Now this is quite a drop in surface brightness -- something like a factor of $1/10^{13}$. In other words, if instead of assuming, as Olbers did, that the stars are uniformly distributed, we clump the stars together into galaxies and then assume the galaxies to be uniformly distributed in the same manner as were the stars in the original analysis, we find that, in effect, we have reduced the brightness of the sky from about -28 magnitudes per square degree to +6 magnitudes per square degree. When we look at the night sky we see this order of brightness only in the direction of the Milky Way. Elsewhere it is much darker. It occurred to Charlier that if the clustering process of stars into galaxies effected this reduction in sky brightness, then if the clustering process were continued, that is, that the galaxies were aggregated into clusters, that there would be an even further reduction in the brightness of the sky. In fact, if we replace galaxies by clusters of galaxies (such as the Coma cluster) as the basic building blocks of the universe, we can reduce the brightness of the sky still further by a factor of about 1/100. In other words, we can make the sky as dark as we please and yet continue to have an infinitely large number of stars, galaxies, clusters, etc., provided that instead of distributing the elements uniformly, we distribute them in an hierarchy of aggregates. Thus we have an alternative way of resolving Olbers' paradox. Further, it should be remarked that a universe in which matter is distributed in aggregates and clusters of aggregates would be classified as assumptions (2) and (3) false. The cosmological principle could be generalized as valid "in the large," provided the sequence of aggregates terminates. Alternatively a cosmological principle of the form:

> All observers located at the centers of n^{th} order clusters would view the universe the same

could be introduced.

This type of universe was first proposed in 1750 by Johann Heinrich Lambert, an Alsatian physicist noted for his work in diffusion of light. Lambert reasoned simply by analogy. He noticed that the satellites of Jupiter and Saturn formed miniature solar systems, with the planets playing the role of the sun. He then considered that the sun with the planets revolving about it, might behave in an analogous way to the satellite systems. He speculated that the sun might be in a planet-like orbit revolving around some distant center of gravitational attraction. Further, this center in turn would be in a planet-like orbit revolving around some even more remote center, etc. Lambert designed a complete universe of the hierarchal type on the basis of this analogy. This is reminiscent of the hierarchy of epicycles used in the Ptolemaic system, which may explain, in part, why hierarchal models have not been seriously considered in modern cosmological thinking.

We have seen that Olbers' paradox may be explained on the basis of the expansion of the universe or the evolution of the universe. A hierarchal universe is not needed to resolve the paradox. Expansion being an observationally established fact confirmed by the law of red-shifts, most cosmologists have felt that it is unnecessary to postulate a Lambert-Charlier type of cosmology, especially since the explanation of sky darkening by expansion or evolution is far simpler than the concept of a hierarchal universe. As a consequence, we find the main-stream of cosmological thought centered on the various types of cosmological models involving expansion which may be classified in the manner we have described.

Recently Harrison[4][5] has questioned the magnitude of the effect of expansion on sky brightness and concluded that expansion accounts for only a small portion of the light attenuation needed to resolve Olbers' Paradox. If, as Harrison claims, assumption number (4) cannot quantitatively remove the difficulties, then the darkness of the sky must be explained as due to either an evolutionary effect or hierarchal structure or both. This is necessary whether the universe be expanding or static. Evolutionary models may be hierarchically structured, although the aging of stars alone can resolve Olbers' Paradox. Steady-state models, in which new stars are continually replacing old ones, cannot appeal to aging. If expansion alone is insufficient to account for darkening, then steady-state models must be hierarchically structured.

The question we wish to consider in the remaining part of this lecture is whether or not a hierarchal type of universe is consistent with present physical and astronomical observations. In his books, Flights from Chaos and Of Stars and Men, the American astronomer, Harlow Shapley, describes the hierarchal way in which the matter of the universe is organized. Shapley's classification of material systems is shown below. If one begins with the fundamental particles of physics, electrons, protons, etc., the first order aggregates of these particles are the atoms. The atoms in turn are the building blocks of the molecules (the next higher order of aggregate in the hierarchy). The list shows the different orders in Shapley's system: atoms; molecules;

SHAPLEY CLASSIFICATION OF MATERIAL SYSTEMS

-5	+2 Satellitic Systems
-4 Corpuscles (Fundamental Particles)	+3 Stars and Star Families
-3 Atoms	+4 Stellar Clusters
-2 Molecules	+5 Galaxies
-1 Molecular Systems	+6 Galaxy Aggregations
+1 Meteoritic Associations	+7 The Metagalaxy
	+8 The Universe: Space-Time Complex
	+9 .....

molecular systems, including crystals and colloidal systems; meteoritic associations, built up from molecular systems; satellite systems; stars; star clusters; galaxies; clusters of galaxies; the metagalaxy; and the universe: each level being an aggregate or set whose elements are in turn the aggregates of order one less. This classification shows us that in the scale interval of the universe with which we are familiar, the scale-wise structure is definitely hierarchal. We have no reason to assume that the largest aggregate that we now know is the largest which exists (saving the term universe for the last). Although arguments from analogy are often persuasive, arguments based solely on analogy cannot definitively establish whether the hierarchy continues to larger and larger aggregates, and there is probably no way to establish whether or not the universe is hierarchal *ad infinitum*. Proper questions for scientific investigation seem to be: Are there observational tests which we may apply to determine whether or not there exist higher orders of aggregates than the largest we now universally recognize, viz., clusters of galaxies? Do the various aggregates have properties in common, and is there some quantitative relationship between the aggregates?

First let us ask how does the concept of hierarchal distribution of matter differ from the uniform distribution of matter assumed in current cosmological models? We may illustrate the differences by considering a large crowd of people standing in a field. Assume the crowd is scattered in a more or less uniform manner over the entire field. This would mean that if we establish a net or reseau of squares over the field, each square being, initially, 100 feet on a side -- then, if we counted the number of people in each square we would find that, except for minor fluctuations, the number would be the same in each square. If we took a small square, 10 feet on a side, and counted the number of people in each square, we would find much larger fluctuations. On the other hand, if the square were larger than 100 feet, we would expect to find smaller fluctuations and a smoother distribution in the number of people in each square. The degree of homogeneity thus depends on the fineness of the net. The larger the squares or the lower the "resolving power," the more homogeneous the counts.*

Now let us consider a second case. Instead of our field being occupied by a crowd, let us assume that a regiment of soldiers is drawn up for review on the field. The soldiers are assembled in platoons; the platoons in companies; the companies in battalions, etc. These are all lined up in an orderly array; however, the spacings between soldiers in a platoon and the spacing of the platoons within the companies, and the companies in the battalions, etc., are not the same. We see in this situation that the results of counts may be quite different from the case of the crowd. If we take a set of 100-foot squares we may find that some squares contain no soldiers at all, because some squares are located between platoons or companies, whereas other squares contain a high density of soldiers. This distribution would be quite inhomogeneous according to the original definition. On the other hand, it may be possible in the regimental case, but not in the crowd case, to pick a size of square which gives no fluctuations; but if the reseau were translated a distance of half the side of the square, large fluctuations would occur. It is also possible to obtain no fluctuations for several different sizes of squares and yet to have fluctuations for intermediate sizes. Furthermore, there is a homogeneity not only of soldiers but also of platoons and companies. This hierarchal type of homogeneity is thus seen to have several properties different from the homogeneity of random distributions. One of the most

*The description given here is highly simplified. In practice, the numbers counted in each square are not compared with some average number, but with a most probable distribution, such as a Poisson distribution.

most important differences is that fluctuations depend not only on the size of squares but also on the positions (and orientations) of the squares. Another difference: As the size of the square is increased, we might find the fluctuations becoming small and then large and then small again; whereas in the random type of homogeneity, we would find that as we increased the size of the squares, the relative fluctuations would always decrease. We shall refer to the first type of homogeneity as random homogeneity and the second type as hierarchal homogeneity.

Before turning to whether or not a hierarchal homogeneous distribution of matter is consistent with current cosmological observations, it is of value to be aware of some of the differences between random and hierarchal homogeneity in the evaluation of various cosmological quantities. These include estimates for (1) density of matter in the universe, (2) distance moduli such as the Hubble parameter, (3) statistical counts of galaxies and clusters of galaxies, and (4) estimates of the size of the visible universe.

In the computation of the density of matter in the universe, if one assumes the universe is made up of clusters as the ultimate element and these are distributed in a random homogeneous matter, we estimate as an upper limit[6] the value of 6×10^{-28} gms per cm^3. However, if the universe is continuingly hierarchal, it is possible to show that the mean density may have a value considerably smaller than this. This may be seen readily from the regimental parade field analogy. If we estimate the density of soldiers by counting the number per unit area inside a platoon, we will have obtained too high a value, for we did not allow for the empty spacing between platoons. Thus our value for soldiers per square foot, while valid in the platoon, is not valid for the density of soldiers in a company and still less valid for the density of soldiers in a battalion, etc. Abell[6] has estimated an upper limit to the matter density of 10^{-29} gms/cm^3 on the basis of second order clusters. If hierarchization continues, the upper limit must be lower still.

In the evaluation of the Hubble parameter (i.e., the ratio of red-shift to distance), it is necessary to use galaxies whose distances can be determined from primary methods as by the Cepheid variables.

If it should occur that all the galaxies for which we can get a useful estimate of distance by such a primary method instead of being distributed in a random homogeneous manner, suffer from some systematic motion because they all belong to a gravitationally contracting or rotating system such as a local cluster, then the value of the Hubble parameter determined within such an aggregate might be in error. That something of this sort may actually be true has been suggested by de Vaucouleurs[7] who has noticed a difference between the redshift-magnitude law in the southern sky and in the northern sky. This anisotropy suggests that the nearby galaxies are participating in some peculiar motion, such as rotation or some mixture of expansion and differential rotation that has introduced an error into the estimate of the Hubble parameter.

A third disparity in assuming random or hierarchal homogeneity arises in counts of galaxies and other extragalactic bodies. For example, the assumption of random homogeneity for the distribution of radio sources runs into inconsistencies in interpreting counts of the numbers of radio sources of various apparent powers. This anomaly might be due to hierarchal effects in the distributions.

A fourth disparity of the assumptions of random homogeneity with hierarchal homogeneity can be illustrated by Fig. 1. Figure 1 shows schematically two possible universes with the horizon of the visible universe enclosed by the black circles. On the right side, the visible sample is a small percentage of the total universe. In this case if we replace the actual clumped distribution of matter with a fluid of uniform density, the relative error will be small. On the other hand, if the visible sample is a large percentage of the total universe (as represented on the left side of Fig. 1), the approximation of the relatively much larger clumps by a uniform fluid introduces relatively large error. We do not know *a priori* the ratio of the visible universe to the total universe, nor whether *smoothing* is justifiable.

Let us now return to the observational evidence for hierarchies. We mentioned earlier, in describing statistical tests which are used in astronomical counts, that when the size of the cells in a reseau is increasing, the relative fluctuations in the number of objects contained in any one cell in the case the distribution is purely random become

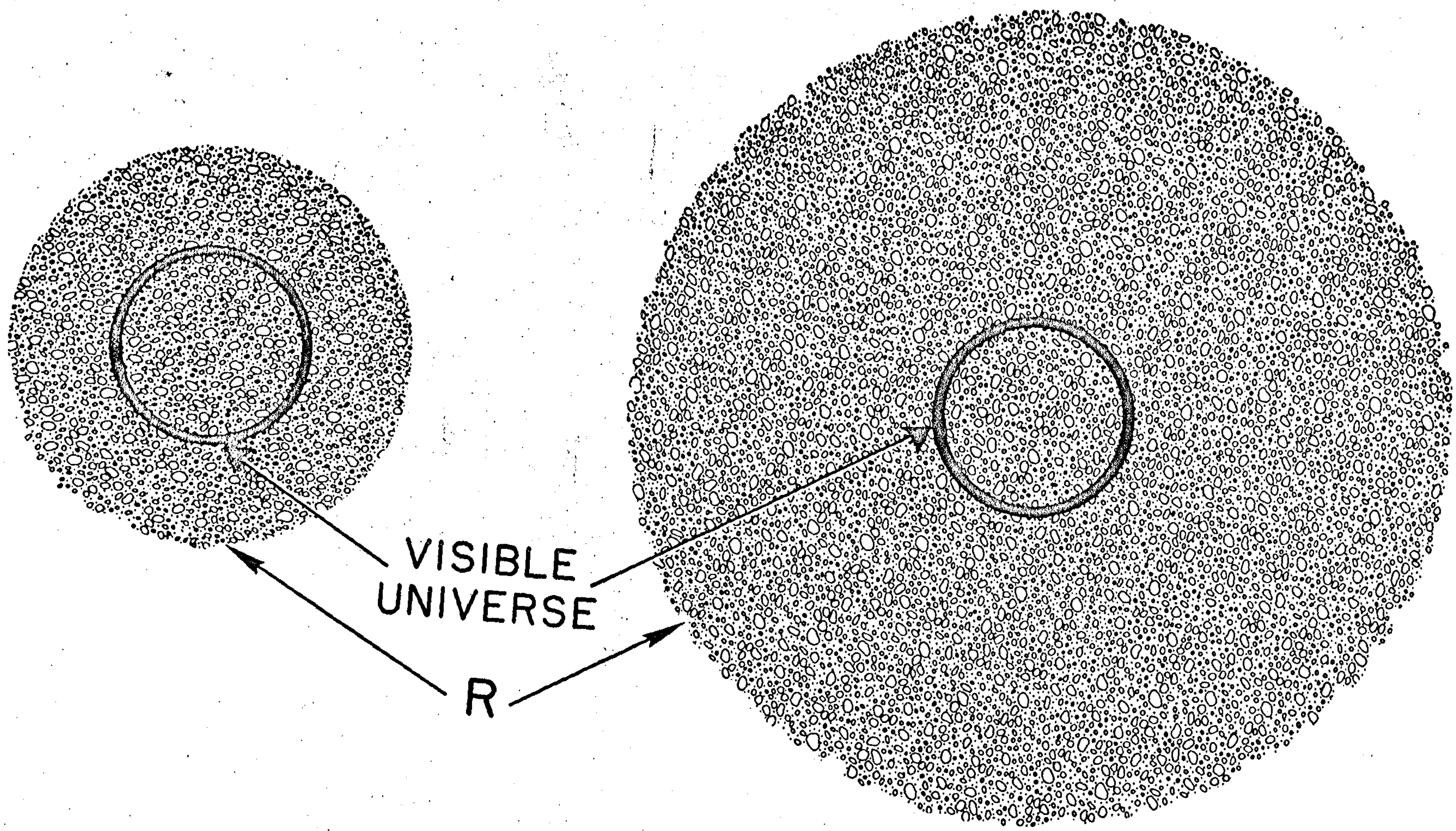


Fig. 1

smaller and smaller -- crowd case. On the other hand, if the distribution is nonrandom, the fluctuations in the number of objects contained in any one cell may go through maxima and minima as the reseau size is varied, maximum fluctuations occurring at certain critical sizes. Before it was established in the 1930's that clusters of galaxies existed as physical systems, they were considered to be statistical fluctuations in the distribution of galaxies. Tests of cell counts established the nonrandom nature of clustering.

The next question is what is the evidence that clusters of clusters or second-order clusters exist as physical systems and are not merely statistical fluctuations in the distribution of clusters. The principal investigators of this problem have been Drs. G. O. Abell of the University of California at Los Angeles and F. Zwicky of California Institute of Technology. When the National Geographic Society--Palomar Sky Survey was begun in 1949, it was discovered that most high latitude fields show large numbers of clusters concentrated in small areas. Were these concentrations evidence of second-order clustering or optical illusions? Zwicky concluded that the existence of clustering of clusters is illusory. He argues that if clusters of clusters constituted physical systems acting under the law of gravity, then the dispersions in the mean velocities of the constituent clusters should be large. He finds no such relative motions -- only the velocities attributable to expansion. His statistical analysis shows that the clusters are randomly distributed and noninteracting. Furthermore, Zwicky asserts that Newton's inverse square law of gravity breaks down at distances exceeding 10^7 light years so that superclusters larger than this would be impossible. In addition, capture and accretion processes necessary to account for the formation of second-order clusters, require times far in excess of the currently accepted time scale.

Abell's studies on the other hand lead him to the conclusion that, even though some superclustering may be optical due to fortuitous alignments of clusters in the line of sight, second-order clusters do exist as physical systems. He points out that the observed scatter in the "magnitude--redshift" relation alone can account for the 10^3 km/sec to 3×10^3 km/sec dispersions predicted for second-order clustering.

His statistical investigations using reseaus of different sizes result in counts consistent with random distributions for small and large cells, but for cells of the size 5×10^7 parsecs he finds nonrandom distributions. Abell also has recently reported confirmation of this conclusion using Zwicky's own data from the Zwicky-Herzog-Wild catalogues. In addition to Abell's evidence, de Vaucouleurs' reported properties of the local supercluster substantiate the physical reality of a large organization containing several clusters as subunits.

What may be concluded from these seemingly contradictory findings? One intriguing speculative possibility comes to mind. Let us assume that both Zwicky and Abell are correct; that is, the distributions of clusters appear both entirely random, and the distributions are random except for a particular cell size. Is there necessarily an intrinsic contradiction? Recall that in the case of hierarchal homogeneity the detection of a nonrandom distribution depends not only on the size of the reseau but also on the position of the reseau. It is conceivable that Zwicky may have selected a set of reseaus of various cell sizes but with an origin that resulted in counts consistent with random distributions. Abell may have used the same set of cell sizes, but translated the whole net to an origin which permitted him to catch clusters of clusters when he tested with the right size cell. If this could be the case, then the apparent contradiction of results is itself confirmation of the existence of hierarchal homogeneity. We may then confidently conclude that hierarchization continues at least to the level of clusters of clusters of galaxies.

There are types of arrangements other than clustering into which matter is organized. Shapley's description of the various levels of organization of matter showed that sometimes regularized, spatial organization occurs, as for example, in molecules and crystals. Thus aggregates may be clusters of matter which are clumped together without any particular regularity such as a swarm of bees or they may possess regularized organization such as is exhibited in crystals. Keeping this in mind, in investigating higher orders of aggregates, we must not only look for clustering of aggregates as evidence of hierarchization but also for the possible existence of regularized arrangement.

Evidence for the possible existence of regularized structure on a cosmic scale was found by Wilson.[8] This was an observed regularity in the mean redshifts of clusters of galaxies which seemed unlikely for randomly distributed clusters, and suggests that hierarchal homogeneity is consistent with observations that reach to one billion light years.

In summary, we have reviewed Olbers' Paradox and noted how the assumptions he postulated in 1826 can be used to classify generically our current cosmological models. We have examined hierarchal homogeneity, which was first proposed by Lambert and Charlier, and find it could serve as a valid basis for new types of cosmological models. Whether or not the expansion of the universe can resolve Olbers' Paradox, the assumption of hierarchal distribution of matter for both expanding and static universes can account for the observed night-sky brightness. Although many of the questions raised here cannot be resolved until more data are available, we can no longer ignore the difficulties which have arisen from assumption of random homogeneity.

REFERENCES

1. Schmidt, M., Ap. J., 141, 1 April 1965, p. 1295.

2. Bondi, H., Theories of Cosmology, Advancement of Science, XII, No. 45, 1955, pp. 33-38.

3. Charlier, C. V. L., Arkiv för Matematik, Astronomi, och Fysik, Band 16, No. 22, 1921.

4. Harrison, E. R., Nature, No. 4955, 17 October 1964, pp. 271-272.

5. Harrison, E. R., Observatory, No. 944, February 1965, pp. 5-7.

6. Abell, G. O., Ann. Rev. of Astro. and Astrophys., 3, 1965, pp. 1-22.

7. De Vaucouleurs, G., Soviet Astronomy, 3, No. 6, pp. 897-1082.

8. Wilson, A. G., Proc. Nat. Acad. Sci., 52, No. 3, September 1964, pp. 847-854.

Hierarchical Structure in the Cosmos

Albert Wilson*

The primary focus of cosmological thought in the present century has been on interpreting the observations of the sample of the universe available to our telescopes in terms of a set of models based on various theories of gravitation; especially the General Theory of Relativity. The problem of the structure *of* the universe is customarily divorced from the problem of the structure *in* the universe. Theoretical cosmologists usually choose to explain the structure and behavior – past and future – of the universe with models that smooth out the distribution of matter in the universe, replacing the observed structured distribution of matter with a uniform homogeneous perfect fluid whose density varies in time, but not in space. However, the structure contained *in* the universe becomes difficult to relate to models constructed around smoothing postulates. This has resulted in separate theoretical approaches to the origin of the various structures in the universe. While most of these approaches have met with some success, they are inadequately related to one another and to cosmological theories.

The arbitrary separation of the structure and behavior of the universe from the structure and behavior of its contents may be expedient from the point of view of mathematical simplification, but it cannot be accepted as more than an exploratory strategy. The observational tests for discriminating between various cosmological models are difficult and marginal. Since several smoothed models are candidates for best fit to the observations, it is unfortunate that the large amount of information contained in the sub-structures of the universe cannot be used in testing these models. But until models that relate the properties of the sub-structures to the properties of the whole are employed, much information of potential cosmological value in sub-structure astronomical observations is not cosmologically useful.

**Douglas Advanced Research Laboratories, Huntington Beach, California, 92647.*

from Hierarchical Structures, eds. L.L. Whyte, A. Wilson and D. Wilson, pp 113-134. New York: American Elsevier (1969)

Paper read at Symposium on Hierarchical Structures in Nature and Artifact, November 1968.

So long as the cosmological problem has been approached through smoothing out the sub-structures, it is not surprising that little attention has been paid to the regularities that exist among the sub-structures. There are many features of the visible sample of the universe that suggest that the regularities in sub-structures which range over 40 orders of magnitude in size and 80 orders of magnitude in mass, are of central significance to the order and operation of the universe. The fact that these regularities may not be readily explainable in terms of existing physical theories, should not deter their examination. The object of this paper is to present an overview of the known structural regularities that link the properties of physical bodies across a hierarchy of levels from the atomic to the cosmic.

MODULAR HIERARCHIES

Because of the confusion created by the many uses of the term "hierarchy" some amplification concerning the sense in which hierarchy is used in astronomy and cosmology is needed. Astronomical usage, in general, employs "hierarchy" to mean a *set of related levels* where the levels may be distinguished by a size or mass parameter. Examples from the past include the hierarchy of spheres associated in ancient cosmographies with the various heavenly bodies beginning with the moon and continuing to the sphere of fixed stars, and the hierarchy of epicycles used by Ptolemy to account for observed planetary motions. Modern concepts of hierarchy in the cosmos began with the speculations of Lambert (1761) who extrapolated to higher order systems the analogy between a satellite system such as that of Jupiter and its moons and the solar system of the sun and its planets. Lambert speculated on a hierarchy consisting of a distant center about which the sun orbited as a satellite and an even more distant center about which the first center orbited, and on to more and more distant centers comprising larger and larger systems. To explain Olbers' and Seeliger's Paradox; Charlier (1908, 1922) posited a universe built up of a hierarchy of "galaxies." The first order galaxies were the familiar ones composed of stars, second order galaxies were aggregates of first order galaxies, third order of second order, and so on. Shapley (1930) pointed to the set of levels into which all matter appears to be organized extending from the sub-atomic particles to the "metagalaxies." Shapley's organization, like Charlier's, constructed the material bodies on any level from the bodies on the level next below. A hierarchy of this type which is of fundamental importance in astronomy we designate a *modular hierarchy*.

The central idea in a modular hierarchy is the *module* which is a structure or a system that may be regarded both as a *whole*, decomposible into sub-modules identified with a lower level, and as a *part* combinable into super-modules identified with a higher level. In astronomy, even though the modules on any level are not identical, the levels may be readily distinguished on the basis of the nature of the principal sub-modules out of which entities are directly composed. Thus, for organization in a modular hierarchy, open and globular star clusters and galaxies would be assigned the same level, all being aggregates of stars. Stars, planets, and moons, all built from atoms, would share the next lower level, while clusters of galaxies would be assigned the next level above. There are several other ways than that of a modular hierarchy for organizing cosmic bodies into levels. Some of these will be discussed later.

The term "module" being used here in this general sense need not be precisely defined, however, we may ascribe two fundamental properties to modules. First, a module possesses some sort of closure or partial closure (Wilson 1969). This closure may be topological, temporal, or defined by some operational rule as in group theory. Second, modules possess a degree of semi-autonomy with respect to other modules and to their context. These two properties appear to be common in all modular hierarchies.

In considering the origin of a modular hierarchy we may inquire at any level as to whether the size, the complexity, and the limits to the module are determined (1) totally by the properties of its sub-structures, (2) by its environment, or (3) by a combination of both module contents and context. And to these logical possibilities we must add a fourth: that the levels and modules in a hierarchical structure are determined by some principle or process that operates independently of all levels of the hierarchy. In this fourth case the *levels* of the modular hierarchy themselves become the *modules* on a single level of a meta-hierarchy. The various levels in the meta-hierarchy are an observable level, an energy or force level and a meta-relational level. As an example, we may think of the lines in the spectrum of an atom as an ordinary hierarchy (but not a modular hierarchy). The levels of the meta-hierarchy would be the spectral lines, the energy levels, and the mathematical law – such as the Balmer formula – that defines the sequence. It may be objected that this is but a representational hierarchy. But the essential point is that the levels are neither determined by the sub-levels nor the super levels, but by a set of eigen values that act as a causal meta-relation.

COSMIC-ATOMIC NUMERICAL RELATIONS

Let us now return to our specific example of a modular hierarchy: the levels of cosmic structure. Instead of assuming a two level model of the cosmos – the level of a homogeneous perfect fluid and the level of the universe as a whole – we shall attempt a multi-level view retaining the atomic, stellar, galactic, galaxy cluster and universe levels. Further, in view of the lacunae in our knowledge of physical processes governing "vertical" relations between levels, it is appropriate to work from observation toward theory. In doing this the steps we must take are somewhat analogous to those taken by Kepler and his successors in the investigation of planetary orbits. From the arithmetic ratios of various powers of the sizes and periods of planetary orbits, Kepler discovered his kinematical relations and from these later came Newton's formulation of the physical laws governing planetary motions. Thus while our ultimate goal is the formulation of the physical laws and processes governing the relations between the levels in the cosmic hierarchy, our immediate goal is much more modest. It is simply to display whatever quantitative regularities may exist between the fundamental measurements made on bodies at each cosmic level.

The properties of the arithmetic relations between fundamental atomic and cosmic constants is not new ground. It has received the attention of many leading physicists and astronomers. Eddington (1923, 1931a,b); Haas (1930a,b, 1932, 1938a,b,c); Stewart (1931); Dirac (1937, 1938); Chandrasekhar (1937); Jordan (1937, 1947); Schrödinger (1938); Kothari (1938); Bondi (1952); Pegg (1968); Gamow (1968); and Alpher (1968) all have developed the subject.

The central theme in the numerical approach to atomic-cosmic relations has been to identify quantitative equivalences between various dimensionless combinations of fundamental constants and whenever possible give them physical interpretations. The epistemological weakness in this approach is the shadow of chance coincidence that cannot be removed by any of the common tests of statistical significance. Confidence in the validity of the numerically indicated relations can only follow from successful predictions or the development of a consistent theoretical construct linked to well established physics.

The basic ingredients in the relational approach are the micro-constants, e, m_e, m_p, and h (the charge and mass of the electron, the mass of the proton, and Planck's constant) the meso-constants, c and G (the velocity of light and the gravitational coupling constant), and the macroparameters H and ρ_u (the Hubble parameter and the mean density of the universe). Recently determined valves of these constants are given in Table I. From these fundamental quantities several important dimensionless ratios may be formed. The values of the dimensionless quantities $\mu = m_p/m_e$ (= 1836.12); $\alpha = 2\pi e^2/hc$ (= 1/137.0378); and $S = e^2/Gm_pm_e$ (= $10^{39.356}$) may

Table I.
Values of Fundamental Physical and Cosmic Constants

Constant	Value (c.g.s.)	$\log_{10}$ (value)	Reference
e	4.80298×10^{-10}	−9.318489	1
m_e	9.10908×10^{-28}	−27.040526	1
m_p	1.67252×10^{-24}	−23.776629	1
h	6.62559×10^{-27}	−26.178776	1
c	2.997925×10^{10}	10.476821	1
G	6.670×10^{-8}	−7.176	1
H^{-1}	13×10^{9} years	17.613 seconds	2
ρ_u	10^{-28}	−28	3
a_o	5.29167×10^{-9}	−8.276407	1
r_e	2.81777×10^{-13}	−12.550095	1
α^{-1}	137.0388	2.136844	1
S	2.265×10^{39}	39.356	
μ	1836.12	3.263901	

From top: charge on electron, mass of electron, mass of proton, Planck's constant, velocity of light, Newton's gravitational constant, inverse Hubble parameter, mean density of visible matter in universe, Bohr radius, radius of electron, inverse fine structure constant, ratio of Coulomb to gravitational forces, ratio of proton to electron mass.

1. Cohen and DuMond (1965), 2. Sandage (1958), and 3. Allen (1963) p. 261.

be established in the laboratory. These are respectively, the ratio of proton to electron mass, the Sommerfeld fine structure constant, and the ratio of Coulomb to gravitational forces.[1]

When the two macro-parameters H and ρ_u are introduced, three additional dimensionless quantities may be formed. The first of these is the "scale parameter" of the universe (the product of the velocity of light, c, and the Hubble time H^{-1}), divided by the electron radius, c/Hr_e. The second is the "mass of the universe" expressed in units of baryon mass (where the scale parameter is taken as the radius of the universe), $\rho_u c^3/H^3 m_p$. The third is the dimensionless gravitational potential of the universe $GM_u/c^2 R_u = G\rho_u/H^2$. Using 75 km/sec/mpc as the present value of the Hubble parameter (Sandage 1968), and $10^{-28} g/cm^3$ for the mean density of matter in the universe (Allen 1963), we obtain:

$$c/Hr_e = 10^{40.64} \doteq 2\pi^2 S$$

$$G\rho_u/H^2 = 10^{0.05} \doteq 1.$$

$$\rho_u c^3/H^3 m_p = 10^{79} \doteq 2S^2$$

It is thus seen that to within small factors (whose exact value cannot be determined with the present precisions of ρ_u and H), the dimensionless cosmic quantities representing the potentital, size, and mass of the universe are closely equal to S^ν, where ν = 0, 1, and 2 respectively. The significant matter here is not the fact that the values differ from integral powers of S by factors

1 It has been recognized that S and α appear to be logarithmically related. As an example of an arithmetic equivalence presently lacking theoretical confirmation, we have $8\pi^2 S = 2^{1/\alpha}$ to within experimental uncertainties. If this equivalence is not a coincidence, it has several important implications. Bahcall and Schmidt (1967) have shown on the basis of 0 III emission pairs in the spectra of several radio galaxies with redshifts up to $\delta\lambda/\lambda = 0.2$ that α appears to have been constant for at least 2×10^9 years. The above equivalence, if non-coincidental, would imply that S has also been constant over this period. Hence if G has been changing with time, e^2 and/or m_p and m_e have also been changing, and if e^2 has been changing, so also has h and/or c. The gravitational constant may, indeed, be expressed in terms of other basic constants by the relation, $G = 8\pi^2 e^2/m_p m_e 2^{1/\alpha}$ (Wilson 1966).

as large as 2 or $2\pi^2$, but the fact that laboratory and observatory measurements of quite diverse phenomena when expressed in dimensionless form appear to approximate so closely some small power of the ratio of electric to gravitational forces. It is also interesting to note that the gravitational potential of the universe is near the Schwarzschild Limit, the theoretical maximum value for potential. These *quantitative* equivalences indicate that there probably exist basic causal *qualitative* relations between the structure of the universe and the properties of the atom and its nucleus (the question of the direction of causality being open).

So far the two levels represented by the atom and the universe as a whole have been shown to be derivable from integral powers of the basic dimensionless ratio S. Numerical relations of a similar type involving fractional powers of S were pointed out by Chandrasekhar (1937) to be related to other cosmic levels. Chandrasekhar formed the dimensional combination

$$M_\nu = \left(\frac{hc}{G}\right)^{\nu} m_p^{\,1-2\nu} \qquad (1)$$

having the dimensions of mass. He pointed out the case $\nu = 3/2$ occurring in the theory of stellar interiors, leads to $M_{3/2} = 5.76 \times 10^{34}$ grams, the observed order of stellar masses. This is also the upper limit to the mass of completely degenerate configurations.

But the Chandrasekhar relation (1) also gives the observed order of mass for other cosmic levels in addition to the stellar level although this is not justifiable theoretically. If values of ν of the form $(2 - 1/n)$ where n is an even integer 2, 4, 6, 8, . . . are selected, then the Chandrasekhar relation predicts a sequence of masses given in Table II that corresponds to those observed for the stellar, galactic, cluster, second order cluster, ...levels of cosmic bodies.[2]

Table II. Masses for Levels of Cosmic Bodies from the Chandrasekhar Relation

Level	n	ν	$\log_{10} M_\nu$ (grams)	$\log_{10} M_\nu$ (dimensionless)
stellar	2	3/2	34.766	58.543
galactic	4	7/4	44.523	68.299
cluster	6	11/6	47.775	71.552
2° cluster	8	15/8	49.401	73.178
3° cluster	10	19/10	50.377	74.153
.....	...	.....	.....	.....
Universe	∞	2	54.280	78.056

Using well known relations between fundamental constants, equation (1) may be rewritten in the form:

$$M_\nu = \left(\frac{2\pi m_e}{\alpha m_p} S\right)^\nu m_p = A^\nu S^\nu m_p \tag{2}$$

where $A = 0.4689$. Hence the masses of the bodies on various cosmic levels defined by $\nu = 1\frac{1}{2}, 1\frac{3}{4}, 1\frac{5}{6}, 1\frac{7}{8}, \ldots, 2$, are seen to be nearly equal to these respective powers of S times the proton mass.

2. If equation (1) is valid for all ν of this sequence, then clusters of higher orders could exist until the ratio of consecutive cluster masses becomes less than two. The first pair for which this happens is $\nu = 31/16$ and $\nu = 35/18$, i.e., 6° and 7° clusters. Observationally, although 3° order clustering has been suspected (Wilson 1967), not even the existence of 2° order clustering has been satisfactorily established. While *even* values of n give masses in good agreement with cosmic levels, the *odd* values do not appear to correspond to any long lived objects. Nonetheless, if there exist two species of body, with masses $10^8 \odot$ and $10^{13} \odot$, such bodies would correspond to $n = 3$ and 5 respectively.

There are additional relations between the measurements of cosmic physics and microphysics. The largest gravitational potentials that have been observed for each of four species of cosmic bodies (stars, galaxies, clusters and 2° order clusters) are given in Table III. The potentials for each species are derived in physically distinct ways. For stars, from eclipsing binary observations; for galaxies, from rotational dynamics; for clusters, from the virial theorem; and for second order clusters, from angular diameters, distances and galaxy counts. It is interesting and somewhat surprising that the maximum in each case is nearly the same, a quantity of the order of 10^{23} grams/cm. If, instead of c.g.s. units, masses are expressed in baryon mass units and radii in Bohr radius units, the dimensionless ratio, $M/R \div m_p/a_o$, is in each case closely equal to 10^{39}. Thus, the upper bound for the gravitational potential of these species of cosmic bodies seems to be σS where σ is a factor of the order of unity not determinable from the present precision of the observational data.

Table III. Maximum Values of Potentials

System	$\log_{10}[M/R]$ (c.g.s.)	$\log_{10}[M/R]$ (dimensionless)
Stars	23.27	38.8
Galaxies	23.6	39.1
Clusters	23.5	39.0
Second-Order Clusters	23.2	38.7

From $M/R \leqslant \sigma S m_p/a_o$, substituting $e^2/Gm_p m_e$ for S and $e^2/m_e \alpha^2 c^2$ for a_o, we obtain

$$\frac{GM}{c^2 R} \leqslant \sigma \alpha^2 \quad .$$

In other words, the dimensionless gravitational potential for these four species of cosmic bodies is bounded, not by the Schwarzschild limit, but by a bound α^2 times smaller. We thus see that not only the dimensionless microphysical quantity, S, but also the fine structure constant, α, emerges from cosmic measurement. (Another occurrence of α^2 in cosmic measurements derives from cluster redshifts (Wilson 1964).)

These results may be displayed graphically. Figure 1 is a small scale representation showing quantitative mass and size relations between atomic and cosmic bodies. The axes are logarithmic.

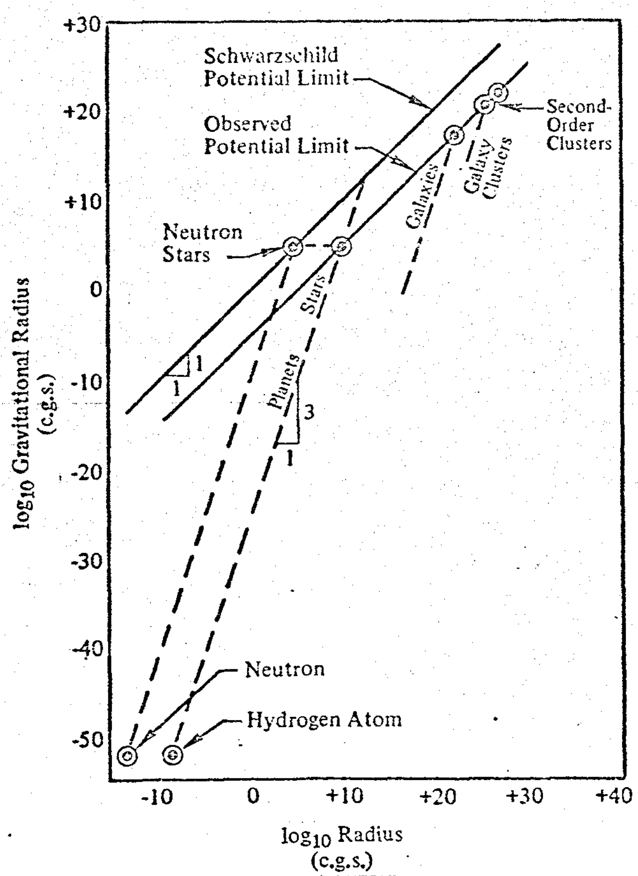


Figure 1. Mass and Size Relations Between Atomic and Cosmic Bodies

The abscissa represents the physical radius; the ordinate, the gravitational radius (GM/c^2). The upper 45 degree line is the Schwarzchild potential limit,

$$\frac{GM}{c^2 R} = \frac{1}{2},$$

the theoretical boundary separating the excluded region (upper left) from the allowable region for self-gravitating bodies. Such bodies as neutron stars, and presumably the universe itself lie on this limit. The lower 45 degree line is the observed or modular potential limit,

$$\frac{GM}{c^2 R} = \alpha^2,$$

marking the locations of the various cosmic bodies having the maximum observed potentials. All other stars, galaxies, clusters, etc., lie below this limit. The relation of the nucleus of the atom and the atom to the degenerate neutron star and the normal star is shown by the dotted lines of constant density (slope 3). Thus a neutron star has the largest mass with nuclear density allowed by the Schwarzschild limit. A normal main sequence star is seen to be limited to the same mass but is non-degenerate, lying on the line representing "atomic density." Thus, given the properties of the atom and the Schwarzschild limit, it is possible to derive the observed maximum mass for a star, but as with the Chandrasekhar relation, it is difficult to account for the locations on the diagram of the bodies of lower density (clusters, galaxies, etc.) and the fact that they are also bounded by the α^2 potential limit.

The parallel lines of equal density (slope 3) through the atom, planets and normal stars, the star clusters and galaxies, the clusters, etc., represent the levels of a modular hierarchy as previously described. These levels are thus definable by a discrete density parameter. Further, in consequence of the universal relation for gravitating systems, $\tau \propto \rho^{-1/2}$, relating a characteristic time to the density, the levels in the cosmic modular hierarchy are also definable in terms of a discrete *time* or *frequency* parameter. We shall return to this concept later.

MASS BOUNDS

In order to display the cosmic or upper portion of Figure 1 with more detail and to make comparisons with observations, the logarithms of observed masses (M) and potentials (M/R) of planets, stars, globular star clusters, galaxies, and clusters of galaxies have been plotted in Figure 2. The masses and potentials (Allen 1963) include maximum and minimum observed values and other representative values selected to show the domains occupied by the respective cosmic species.

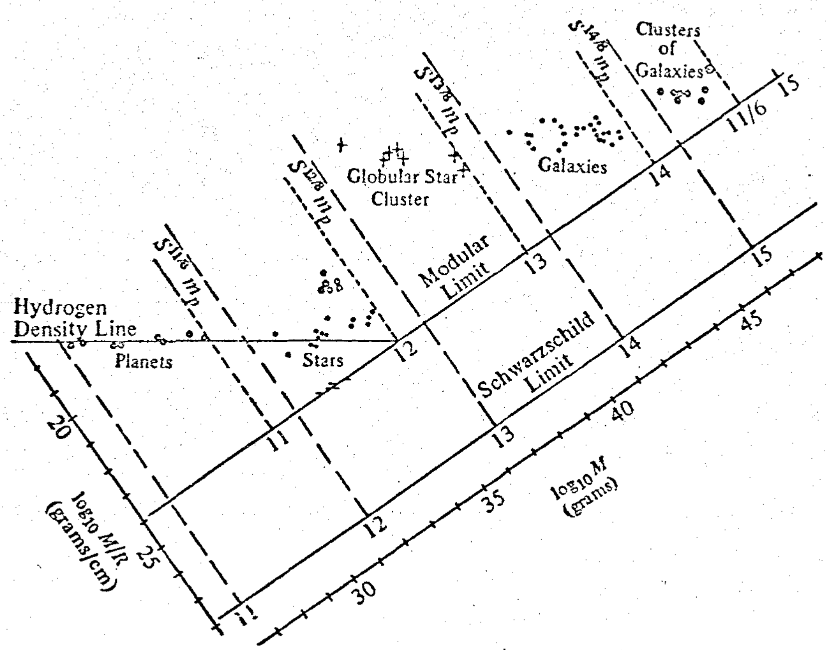


Figure 2. Mass Bounds of Cosmic Bodies

However, because of observational bias toward brightest and largest objects, the minimum observed values are not as representative of actual minimum values as the maximum observed values are of actual maximum values. Figure 2 is related to Figure 1 by an affine transformation (Figure 1 has not only been dialated, but has also been subjected to shear, reflection and rotation transformations). In Figure 2, the lines of constant density are shown horizontally so as to display the levels into which cosmic bodies fall when viewed as a modular hierarchy.

The supergiant stars lying above the mean stellar density level are shown as open circles, while the white dwarfs lying below the level near the modular potential limit are shown as dashes. The Schwarzschild Limit, $M/R = c^2/2G$ and the modular (or observed) limit, $M/R = Sm_p/a_o$ have a slope of 2/3 with respect to the horizontal equi-density lines. The short-dashed and long-dashed lines perpendicular to the Schwarzschild and modular limits are lines of constant mass. The set of short-dashed lines, extending only to the modular limit represent the sequence of masses $M_\nu = S^\nu m_p$, showing values of ν = 11/8, 12/8, 13/8, 14/8, and 11/6. The set of long-dashed mass lines, extending to the Schwarzschild Limit are located so as to pass through a sequence of points on the Schwarzschild Limit that have the same gravitational energy as the intersections of the $S^\nu m_p$ mass lines with modular limit. The pairs of intersections marked 14, 13, 12, . . .lie on lines of constant gravitational energy, $GM^2/R = S^\nu m_p(\alpha c)^2$. For identification, corresponding upper and lower bound intersections with the modular and the Schwarzschild Limits are marked with the *numerators* of the exponent ν. That is, 14 on the Schwarzschild Limit marks the lower bound of galaxies and corresponds to the upper bound $S^{14/8}m_p$ intersection with the modular limit.

The values of mass given by the Chandrasekhar relation (1) in Table II are the correct order of magnitude for the masses of stars, galaxies, and clusters. In Figure 2 it can be seen from the set of short-dashed lines of constant mass that the sequence of masses $S^{\nu}m_p$ are close in value to least upper bounds of the masses of planets, stars, globular star clusters, galaxies, and clusters of galaxies. Numerical comparisons of maxima are given in Table IV. In addition, the set of long-dashed lines are seen to be lower bounds, while probably not greatest lower bounds nonetheless close to the actual observed minimum values of the masses of the respective species of cosmic bodies. Numerical comparisons of minima are also given in Table IV where the lower bounds are the upper bounds diminished by $10^{3.9}m_p$. It can be shown that this value of maximum-minimum mass differential may be derived from "ν sequences" of maximum

Table IV. Observed and Calculated Mass Limits

Mass Limit	Planets	Stars	Globular Clusters	Galaxies	Galaxy Clusters
MAXIMUM					
	Jupiter	VV Cephei A	M22	M31	Local Super Cluster
Observed	30.279	35.225	40.14	44.8	48.3
Model	30.338	35.258	40.176	45.096	48.376
$S^{\nu}m_p$	$\nu = 11/8$	$\nu = 12/8$	$\nu = 13/8$	$\nu = 14/8$	$\nu = 11/6$
MINIMUM					
	Mercury	R CMa B	M5	NGC6822	
Observed	26.509	32.340	37.3	41.9	
Model	26.4	31.4	36.3	41.2	

All masses are given in $\log_{10}$ (grams). Upper bounds are given by $S^{\nu}m_p$, lower bounds by $S^{\nu}10^{-3.9}m_p$.

masses and gravitational energies, with the minimum mass being the least allowed by the Schwarzschild Limit for a given gravitational energy.

THE COSMIC DIAGRAM

The good agreement between the observed values for the masses and sizes of various species of cosmic bodies and the values given by sequences involving simple expressions containing fundamental physical constants indicates the probable validity of the gross features of the sequences. However, systematic errors and incompleteness in the observational data and the uncertainties intrinsic in establishing observationally least upper bounds and greatest lower bounds render it impossible, in the absence of a rigorous physical theory, to predict the exact form of the expressions and the values of the small factors (such as the 2π's, etc.) that should be included. We might, as an analogy, think of our discerning Kepler's Third Law in the form: periods squared are proportional to orbital diameters cubed without knowing the important constant of proportionality, $G(M_1 + M_2)$.

In the spirit of focusing on the major patterns that emerge from the present body of observations that are not likely to be seriously altered by refinements in observation, or even by discovery of new bodies, we represent the gross features of the structure in the universe in Figure 3. In this stylized representation, the cosmos is mapped on a rectangle whose length is the logarithm of the mass, $S^\nu m_p$, and whose hieght is the logarithm of the extension, $S^\eta a_o$. The masses and radii of various sub-components are related to values of ν and η. The hydrogen atom, mass m_p, and radius a_o, is located at the origin at H with $\nu = 0$, $\eta = 0$. The mass and radius of the universe are represented by the values $\nu = 2$, $\eta = 1$ at U. The modular and Schwarzschild potential limits are the upper and lower 45° lines respectively. The remaining observed bodies in the universe lie roughly within the three hatched bands, whose slope is that of constant density terminating at the modular limit. The bodies on the lowest and longest band have density of the order of one g/cm^3 and include asteroids, satellites, planets, and stars. This band terminates on the modular limit at $\nu = 3/2, \eta = 1/2$. With little mass overlap of the first sequence, the next sequence of bodies (star clusters and galaxies) begins near $\nu = 3/2$ and falls along an equi-density band reaching the modular limit at $\nu = 7/4$, $\eta = 3/4$. Above this point the observational uncertainties do not permit a definitive picture. It is not clear whether there exist two (or more) sequences of clusters of galaxies or only one.

A cluster sequence terminating at $\nu = 11/6, \eta = 5/6$ together with a second sequence of higher order clusters terminating at $\nu = 15/8$, $\eta = 7/8$ (as shown in Figure 1 and Figure 2) may fit observations better than the single sequence extending to $\nu = 15/8$, $\eta = 7/8$ shown in Figure 3. The resolution of this structure as well as whether still higher levels of clustering exist must be decided on the basis of future observations.

From the point of view of hierarchies, the levels occupied by cosmic bodies may be described either as *modular levels* (in the sense defined earlier), or as levels defined by a density

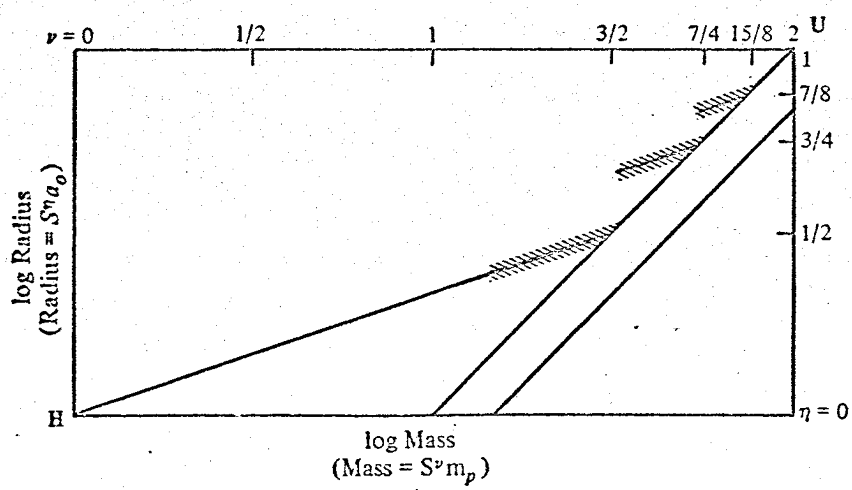


Figure 3. Cosmic Diagram

parameter, or its equivalent frequency parameter. In addition the structure may be "sliced" differently and the cosmic bodies may be allotted to distinct levels defined by a mass parameter. These levels are broad but on the scale of Figure 2 appear to be distinct.

INTERPRETATIONS

An intrinsic difficulty in relating empirical results (such as those displayed in Figures 2 and 3) to current physical theories is that numbers of the magnitude of S are not contained in any classical equations of physics. This difficulty has been expounded by Dirac (1938), Jordan (1947) and others. Eddington (1931) made attempts to derive the fundamental dimensionless constants from first principles, not, however, with complete success in reproducing the observed values. A theoretical understanding of the various observed relations between the different levels of cosmic structure – atoms, stars, galaxies, . . .the universe – is thus likely to come only after new theories of such concepts as time, degeneracy, and informational content of structure are available. At the present stage only some *speculative* suggestions can be made.

For example, the existence of *two* potential limits, the Schwarzschild and the modular, implying that the same extension ratio (the α^2 ratio of atomic to nuclear dimensions) holds between non-degenerate and collapsed configurations at stellar, galactic and cluster levels, suggests that through a generalization of the concept of degeneracy, the theorectical validity of equation (1) for all levels might be established. One might speculate that configurations at every level possess a collapsed or close packed state, and an extended state α^{-2} times larger. An alternate approach may be that the reflection of the α^2 ratio into higher levels of cosmic-structure is a cosmogonic vestige from a universe in a highly collapsed state. But whatever the cause of the modular limit, it must be regarded as an important observational feature to be accounted for by cosmological theories.

A second speculative suggestion is that in the sequence of powers of S that map observed mass configurations, we are encountering a resonance phenomenon. However, the fundamental and the overtones are exponentially related instead of being related in the manner of Pythagorean harmonics. This suggests kinship to the logarithmic time derived by Milne (1935) in his kinematic relativity. If we take as the basic gravitational frequency, the inverse Schuster period, $f_o = (Gm_p)^{1/2}/2\pi a_o^{3/2}$, then the overtones are given by

$$f_\nu = \frac{(GS^\nu m_p)^{1/2}}{2\pi(S^{\nu-1}a_o)^{3/2}} = f_o S^{3/2-\nu} \qquad (3)$$

where $\nu = 3/2, 7/4, 15/8, \ldots$.

Numerically, $f_{3/2} = f_o$, the frequency associated with the hydrogen-stellar line of Figure 3, corresponds to a period of about two hours; $f_{7/4}$, the galactic line corresponds to 10^6 years; $f_{15/8}$, the cluster line corresponds to 85 x 10^9 years; and f_2 corresponds to 10^{15} years. The cluster value is close to the period derived by Sandage for an oscillating universe. Viewed as a Hubble time, it corresponds to a value of H = 74.13 km/sec/mpc, in close agreement with the observed value of H = 75.3 km/sec/mpc derived from cluster distances (Sandage 1968).

If we take this equivalence between the ν = 15/8 cluster gravitational time and the observed cluster Hubble time, as additional corroboration of the valid representation of the cosmic diagram, then we infer that the visible sample of the universe, the "realm of the galaxies and clusters" is not the ν = 2 universe. The observations at the limits of our telescopes are describing the ν = 15/8 sub-structure and not the universe. Characteristic times of the order of 10^{10} years are those associated with the cluster level sub-structure. The characteristic gravitational time of the ν = 2 universe, on the other hand, is of the order of 10^{15} years. The appearance of a time of this magnitude brings to mind the controversy that waged in cosmology following the publication of James Jeans (1929) estimate of the dynamic age of the galaxy at 10^{13} years. The adherents of the "short time-scale," held the age of the universe to be but a few eons while those who subscribed to the "long time-scale," required an age of the order of 10^{13} years or greater. Since the galaxy could not be older than the universe, the issue was settled against Jeans. But if the few eons refers not to the universe but to the cluster level sub-structure, there is no *a priori* reason why the galaxy cannot be older than the cluster level sub-structure.

If the cosmic diagram suggests some form of resonance as the process of morphogenesis, then as sand collects at the nodes on a vibrating drum head, matter concentrates at nodes corresponding to the set of frequencies $S^{3/2-\nu} f_o$. This raises many physical questions. Most importantly, what is it that is pulsating or vibrating at these frequencies – some substratum, matter itself, or what? Analogies to familiar equations suggest that from the cosmic diagram, we have a set of eigen values representing mass levels, energy levels, or frequencies that are solutions to some "cosmic wave equation." Perhaps the first step toward a physical theory would be to derive such an equation.

REFERENCES

Allen, C. W. 1963. *Astrophysical Quantities.* 2nd ed. London: The Athlone Press.

Alpher, R. A., and Gamow, G. 1968. "A Possible Relation Between Cosmological Quantities and the Characteristics of Elementary Particles." *Proc. Nat. Acad. Sci* 61:363.

Bahcall, J. N., and Schmidt, M. 1967. "Does the Fine-Structure Constant Vary with Cosmic Time?" *Phys. Rev. Letters* 19:1294.

Bondi, H. 1952. *Cosmology.* Cambridge: Cambridge University Press.

Chandrasekhar, S. 1937. "The Cosmological Constants." *Nature* 139:757.

Charlier, C. V. L. 1908. "Wie eine unendliche Welt aufgebaut sein kann." *Arkiv för Matematik, Astronomi och Fysik.* Band 4, No. 24.

——. 1922. "How an Infinite World May be Built Up." *Arkiv för Matematik, Astronomi och Fysik.* Band 16. No. 22, pp. 1-34.

Cohen, E. R., and DuMond, J. W. N. 1965. "Our Knowledge of the Fundamental Constants of Physics and Chemistry in 1965." *Phys. Rev.* 37:537.

Dirac, P. A. M. 1937. "The Cosmological Constant." *Nature* 139:323.

——. 1938. "A New Basis for Cosmology." *Proc. Roy. Soc., Series A* 165:199.

Eddington, A. S. 1923. *The Mathematical Theory of Relativity.* Cambridge: Cambridge University Press.

——. 1931a. "Preliminary Note on the Masses of the Electron, the Proton, and the Universe." *Proc. Cambridge, Phil. Soc.* 27:15.

——. 1931b. "On the Value of the Cosmical Constant." *Proc. Roy. Soc., Series A.* 133:605.

Gamow, G. 1968. "Numerology of the Constants of Nature." *Proc. Nat. Acad. Sci.* 59:313.

Haas, A. E. 1930a. "Die mittlere Massendichte des Universums." *Anzeiger der Akad. Wiss. Wien.* 67:159.

——. 1930b. "Uber den möglichen Zusammenhang zwischen Kosmischen und physikalischen Konstanten." *Anzeiger der Akad. Wiss. Wien.* 67:161.

——. 1932. "Uber die Beziehung zwischen Krummungsradius der Welt und Elektronenradius." *Anzeiger der Akad. Wiss. Wien.* 67:91.

——. 1938a. "A Relation Between the Average Mass of the Fixed Stars and the Cosmic Constants." *Science* 87:195.

——. 1938b. "A Relation Between the Electronic Radius and the Compton Wavelength of the Proton." *Science* 87:584.

——. 1938c. "The Dimensionless Constants of Physics." *Proc. Nat. Acad. Sci.* 24:274.

Jeans, J. 1929. *Astronomy and Cosmogony.* Cambridge: Cambridge University Press.

Jordan, P. 1937. "Die Physikalischen Weltkonstanten." *Die Naturwissenscaften* 25:513.

——. 1947. *Herkunft der Sterne.* Stuttgart: Wiss, Verlag.

Kothari, D. S. 1938. "Cosmological and Atomic Constants." *Nature* 142:354.

Lambert, J. H. 1761. "System of the World." *Kosmologische Briefe.*

Milne, E. A. 1935. *Reliability, Gravitation and World Structure.* Oxford: Oxford University Press.

Pegg, D. T. 1968. "Cosmology and Electrodynamics." *Nature* 220:154.

Sandage, A. R. 1968. "Directors Report, Mt. Wilson and Palomar Observatories, 1967-1968." p. 33. Pasadena, California.

Schrödinger, E. 1938. "Mean Free Path of Protons in the Universe." *Nature* 141:410.

Shapley, H. 1930. *Flights from Chaos.* New York: McGraw-Hill.

Stewart, J. Q. 1931. "Nebular Red Shift and Universal Constants." *Phys. Rev.* 38:2071.

Wilson, A. G. 1964. "Discretized Structure in the Distribution of Clusters of Galaxies." *Proc. Nat. Acad. Sci.* 52:847-54.

——. 1966. "Dimensionless Physical Constants in Terms of Mathematical Constants." *Nature* 212:862.

——. 1967. "A Hierarchical Cosmological Model." *Astronom. J.* 72:326.

——. 1969. "Closure, Entity and Level" (this volume).

Reprinted from

THE ASTRONOMICAL JOURNAL VOLUME 70, NUMBER 2 MARCH 1965

Abstracts of Papers Presented at the 117th Meeting of the American Astronomical Society, held 28–31 December 1964 at Montreal, Canada

p 150

On Super-Organization Among Clusters of Galaxies. A. G. WILSON, *The Rand Corporation.*—The mean redshifts of clusters of galaxies do not appear to be distributed randomly. The small sample of available mean redshifts is consistent with the hypothesis that the clusters are located on a set of concentric shells which possess a definite relation between successive radii (Wilson, A. G., *Proc. Nat. Acad. Sci.,* **52,** 1964). If this distribution is real, the cosmological principle requires that apparent cluster distribution should be on concentric shells for all equivalent observers (i.e., observers located in or near a cluster). The actual spatial locations of clusters must then be at the intersections of the several sets of cluster-centered concentric shells. This requires structure in the *angular* distribution of cluster centers as seen by equivalent observers.

The investigation of regular structure in the distribution of clusters may be investigated further by combining the angular positions of clusters with the mean redshifts to generate additional "quasi-redshifts" by triangulation. If a linear redshift-distance relation and Euclidean space are assumed, the quasi-redshifts may be derived by the ordinary law of cosines. If these assumptions are valid and if the spatial distribution of clusters is regular the frequency distribution of quasi-redshifts should be a set of discrete peaks or resonances which represent the allowable separations between clusters.

The frequency histograms of the quasi-redshifts show a set of peaks distributed among a "noise" background. Statistical tests show that over fifteen of these resonances are not likely to be random fluctuations, (observed occurrence minus expected occurrence$>3\sigma$). It may be inferred that at least a subset of clusters manifests structured distribution. The noise may be due to breakdown of the Euclidean and linear-redshift approximations, to the coexistence of two or more independent organizations, and/or to actual random distributions.

It is further found that the ratios of the values at which some of the resonances occur are $3^{1/2}$, 2, $5^{1/2}$, $8^{1/2}$, $13^{1/2}$, suggesting the distance ratios which obtain for closely packed spheres.

The unlikelihood of the occurrence of these peaks and ratios in distances between clusters distributed in a random uniform manner suggests either that some form of super-organization exists among the clusters or that we are observing the vestiges of a structure whose angular and linear ratios have been preserved under a uniform and isotropic expansion from a time when the universe was in a highly compact stage. The latter hypothesis if physically consistent, would be corroborative of an oscillatory or other evolutionary model.

Alternatives to the vestige-hypothesis must account for an organization extended over 10^9 parsecs, the value bounding the separations of the clusters involved. It is difficult to explain such an extended organization without the introduction of physical communication processes not at present recognized.

THE ASTRONOMICAL JOURNAL

PUBLISHED BY THE AMERICAN INSTITUTE OF PHYSICS
FOR THE AMERICAN ASTRONOMICAL SOCIETY

VOLUME 71 1966 August ~ No. 1341 *NUMBER 6* p 402

Structural Parallels in Nondegenerate Cosmic Bodies. ALBERT G. WILSON, *Douglas Advanced Research Laboratories.*—Schwarzschild's exact solution of the Einstein field equations leads to the prediction of a bound to the ratio of the gravitational radius to the linear radius for all gravitating systems, namely,

$$2GM/c^2R<1.$$

It is observed that this ratio for each of four species of *nondegenerate* cosmic body—main sequence stars, galaxies, clusters of galaxies, second-order clusters—is bounded and that the bound is closely the same for each species: $2GM/c^2R \leqslant 10^{-4.3}$. ($M/R=10^{23.5}$ g/cm or 10^{39} with respect to the mass and radius of the hydrogen atom.) The ratio of the *observed* bound for nondegenerate bodies to the *Schwarzschild* bound is of the order of the ratio of atomic to nuclear dimensions.

Assuming the Schwarzschild bound governs totally degenerate matter, the upper limit to observed masses of stars may be explained as the result of the Schwarzschild limit forbidding a mass of greater than about 10^{34} g for a dense neutron fluid under initial conditions similar to those postulated in evolutionary cosmological models.

The $10^{-4.3}$ bound appears to play the limiting role for nondegenerate matter. The latter bound limits stellar matter under maximum nondegenerate density conditions to masses of about 10^{34} g, consistent with observed main sequence stellar masses.

A basic question is raised by the existence of the $10^{-4.3}$ bound for aggregates other than stars. Some generalized form of degeneracy for larger aggregates may be implied.

Reprinted from

THE ASTRONOMICAL JOURNAL VOLUME 72, NUMBER 3 APRIL 1967

Abstracts of Papers Presented at the 123rd Meeting of the American Astronomical Society, held 27–30 December 1966 at the University of California at Los Angeles, California

A Hierarchal Cosmological Model. ALBERT WILSON, *Douglas Advanced Research Laboratories.*—The observation of equal maximum values of gravitational potential for stars, galaxies, and clusters of galaxies (Wilson, A. G., *Astron. J.* 71, 402, 1966) suggests the existence of a universal potential bound governing gravitational stability. The assumption that systems whose potentials lie in the zone between the observed maximum value and the Schwarzschild limit ($10^{-4}<2GM/c^2R<1$), are unstable, whatever their densities or total energies, prohibits the stable existence of uniformly distributed matter of indefinite extent. Large masses in order to form stable systems must be structured hierarchically.

The existence of banded structures in the distributions of the redshifts of rich clusters of galaxies and radio sources (Wilson, A. G., *Proceedings of the 14th International Astrophysical Symposium, Liege, 1966*, to be published) indicates the existence of one or more possible additional members of a hierarchal structure which would be expected as a consequence of the assumed universal stability bound. Estimates of the potentials of these indicated super systems place them within the instability zone, consistent with, and possibly causally related to, the observed general expansion.

Anomolies in THE DISTRIBUTION OF LARGE REDSHIFTS

If we focus our attention on the values of redshifts which have been obtained from various cosmic objects that lie in the range of .015 or greater, we find that the distribution values show some rather remarkable clumpings. In fact it is readily shown that the distributions of redshifts are not random but tend to display bands and gaps. If one makes a division of the redshifts into two classes, those obtained for radio sources, and those obtained for clusters of the galaxies, one finds a rather remarkable relationship between the corresponding bands and gaps which occur in the two classes of redshifts, as shown in Figure 1. It is well known that at least half the radio sources lie in clusters of galaxies, and usually a large central galaxy of a cluster is itself a radio source. If we are to interpret the distributions shown in Figure 1 strictly on the basis that the redshift is a measure of distance in accordance with the Hubbell relationship, then we find that there exist gaps in space in which little or no cosmic matter exists. These gaps alternate with bands in which high densities of cosmic matter exist. That this interpretation is likely is substantiated in the case of the cluster distribution of redshifts by the fact that the richest clusters are found to lie in the center of the bands and the less rich clusters toward the edges. If thus we have a material distribution of this nature, we immediately see because of the fact that radio sources must be distributed cosmically the same way as cluster galaxies, that we cannot interpret the redshifts as adhering strictly to Hubbell's law, for the actual distances to the various radio sources and the actual distances to the various clusters lie in the same bands. There exists a systematic displacement in the redshifts of one species of object with respect to another. In other words, part of the redshift in one or both species must be attributable to other than a Hubbell redshift. Because of the possibility of studying the differential effects between the radio sources and the clusters, let us assume initially that the cluster redshift consist exclusively of a Hubbell redshift; that all other aspects, Einstein shift, proper motions, etc., are negligible; and let us consider that the excess displacement which appears in the redshifts of the radio sources is due to some other cause than the basic Hubbell or cosmic redshift.

Schmidt has shown that the appearance of emission lines in galaxies is correlated to the larger masses or smaller radii of galaxies. It seems reasonable, therefore, to conclude that the excess redshift may be attributable to an Einstein shift and that the radio sources displaying emission lines are more massive than the cluster galaxies, and having the same or smaller radii, are, therefore, likely to display an Einstein shift which is proportional to $\frac{M}{R}$. If we measure the amount of the displacement, assuming that the bands correspondingly labeled in Figure 1 are actually at the same distances, we find that ΔZ, the displacement in redshift between the radio sources and the clusters, increases as the cluster redshift to the 3/2 power. This ΔZ may be equated to $\frac{GM}{C^2R}$ where M is the mass of the radio galaxy and R is its radius. This may be solved for R and we find that the radius of the galaxy is changing with time. If the Einstein shift of the non-radio galaxies is negligible, we find not only that the radio sources are expanding with time but that the rate of expansion is decelerating. The 3/2 law suggests a possible alternative interpretation of the excess redshift, namely, that it may be due to a transverse Doppler effect and that we are observing a rotating system which is rotating in accord with a Keplarian 3/2 law.

On Expansion of Radio Sources

A. G. Wilson

~~A Note on the A-E Effect~~

In 1961 On the basis of ~~a~~ the sample of all available redshifts of galaxies in the Virgo Cluster, Holmberg (1) suggested that there existed a systematic difference in the redshifts derived from absorption lines and those from emission lines. A thorough analysis of Holmberg's sample together with additional redshifts by A. and G. de Vaucouleurs (2) concluded that the effect was not real and that Holmberg's result was due to the capriciousness of a small sample. This note reports further evidence of ~~xxxxxx~~ systematic differences in redshifts based on emission spectra and absorption spectra. The redshift samples discussed here - although again smaller than desirable - indicate a systematic discrepancy larger than attributable to observational errors.

The mean redshifts of clusters of galaxies are distributed in a manner more suggestive of some form of super-clustering than of uniform random distribution. Although the sample is small, the bands and gaps in the redshift distribution seem to reflect a real phenomenon (3,4). The distribution of cluster richness within a band, with the richest clusters tending to be centralized (Fig. 1), indicates that the bands are regions of high matter density and not just high cluster density. It is, therefore, not entirely surprising to observe that the redshifts of the radio sources show the same type of banded distribution. When the redshifts of clusters (Fig. 2) bands and gaps corresponding to those of the cluster distribution are noted. However, the radio source bands are systematically displaced to the red.

If we assume that there do exist large scale departures from homogeneity and that there are alternate high and low density regions with sizes of the order of mpc (H = 100 km/sec/mpc), we may expect the corresponding bands in Fig. 2 to arise from matter - whether in clusters of galaxies or in radio sources - in the same spatial neighborhood. On the basis of this assumption the redshifts from radio sources are systematically greater than the redshifts from clusters at the same distance. Assuming that the inner and outer band limits are not badly delineated by the present samples, the displacements ($z_{radio}-z_{cluster}$) of corresponding limits may be determined as a function of $z_{cluster}$ or z_{radio}. The log (displacement) - log ($z_{cluster}$) relation is shown in Fig. 3. This relation is well approximated by the expression

$$z_r - z_c = 0.5\ z_c^{3/2}$$

The line C-D in Fig. 2 shows the cluster redshifts as individually displaced by this relation. The correspondence with radio redshifts is remarkably close.

The correspondence of two samples not subject to the same selectivity factors, suggests that the effects of the capriciousness of small samples are sizeably reduced. If there exists a Holmberg displacement, the reality of the bands is statistically strengthened. If the bands are real, there is good evidence for a Holmberg displacement.

The Distribution of Large Redshifts

ROUGH DRAFT

There is ~~another~~ a phenomenon which suggests that there may exist differential expansion between galaxies which are radio sources and those which are not. It is observed that the distribution of the mean redshifts of clusters of galaxies is not random. Rather, they appear to be bunched, exhibiting a series of bands and gaps. On the basis of random null functions with uniform density and density proportional to the distance squared distributions, the observed gaps are highly improbable ($1:10^{-6}$). The distribution of the redshifts of radio sources shows a similar bunching into bands and gaps with an even more improbable ($1:10^{-8}$) origin from random fluctuations. (Slide 1)

1) Not all Hubble because of R.S. in clusters

2) If gravitation then expansion

There is no obvious selectivity factor operating in the choice of either clusters or radio sources which were in the observed samples which could account for these band-gap distributions. Furthermore, the fact that the clusters whose mean redshifts lie in the center portions of the bands are the richer clusters, suggests that we are probably looking at an actual band-gap distribution in cosmic matter. That is to say that the radiating matter in our part of the universe seems to be distributed in a series of shells or density waves.

If we assume the reality of these shells or waves delineating the distributions of matter and assume further that galaxies - whether radio sources or not - are similarly distributed, there appears to be a systematic displacement of the observed band patterns in the redshifts of the radio sources with respect to the band patterns in the mean redshifts of the clusters. If, as Dr. Arp has demonstrated, the radio sources are expanding, then there is an immediate explanation for this displacement bringing all the bands into a remarkable coincidence, and presenting strong confirmatory evidence for the hypothesis of the existence of a band-gap or density wave distribution in cosmic matter. (Or conversely, if we assume the density wave hypothesis as explaining the observed band-gap distributions in the redshifts, we have corraborative evidence for the differential expansion of the radio sources.)

Observed redshifts may contain three components: (1) a cosmic or Hubble component representing the general expansion, (2) a doppler component representing the local peculiar motion, and (3) an Einstein shift.

Neglecting the second component as being relatively small and likely to be similar for radio and non-radio galaxies, we may attribute the observed band pattern displacements in the total redshifts to a difference in the Einstein shifts of the two species of galaxies.

Assuming corresponding band centers and edges to have identical Hubble shifts (a factor arising from the limited size of the sample causes some uncertainties in the values of the band edges), for each of these abscisses,

$$Z \text{ radio} - \bar{Z} \text{ cluster} = \Delta Z = \text{the Einstein shift.}$$

These displacements are found to vary in a systematic monotone fashion increasing with the Hubble shift. (Slide 2)

$$z_R = z_{RH} + z_{RE}$$

$$\bar{z}_C = \bar{z}_{CH} + \bar{z}_{CE}$$

$$\bar{z}_R - \bar{z}_C = \Delta z = z_{RE} - \bar{z}_{CE}$$

$$\frac{G}{c^2}\left[\frac{M_R}{R_R} - \frac{M_C}{R_C}\right]$$

$\dot{R}_R$ and $\dot{R}_C$ may be the same but $M_R >> M_C$

But $\bar{Z}c \sim$ (distance). $H/C \sim To - T$

where T is the epoch at which the light left the source and To is the present epoch - i.e, the larger $\bar{Z}c$, the younger the object observed.

Also $$\Delta Z = \frac{GM}{\bar{c}^2 R}$$

If there is no mass change and no change in the values of the fundamental physical constants, G and c, then since ΔZ in decreasing with age, R must be increasing with age.

It will be recalled that Holmberg pointed out a systematic difference between the redshifts based on omission lines and those based on absorption lines for galaxies in the Virgo cluster, the emission objects being greater. This was explained by deVarcouleurs on the basis of the two Virgo clusters. The spiral (emission) cluster being further that the elliptical. Since luminosity criteria cannot be invoked to give independent estimates of the relative distances (there is no way of calibrating mean magnitudes of ellipticals except by assuming one Virgo cluster).

An additional property of the distribution

5

.0301
.0302
.0303
.0306
.0307
} identical to within observational error

.0568
.0570
}

.0820
.0825
}

.2201
.2206
}

A curious property is the large number of identities of redshifts in different parts of the sky — identical to within obs. error.

26.0
18.5
7.5

against which the redshift of the radio source could be checked for a displacement. Baum [3] has estimated a mean redshift for the cluster containing 3C295 from photometric measures of several galaxies in the cluster (not including the radio source). He finds z = 0.44 as against z = 0.461 for 3C295 measured spectroscopically. The displacement of 0.02 is consistent with the hypothesis but quantitative verification must await spectral redshifts.

Because cluster galaxy expansion is unknown, it is not possible to assign an absolute rate to the radio source expansion. If the cluster redshifts be taken as entirely cosmic in origin, then on the basis of the observed three-halves relation where the constant is approximately 1/1.8, it is deduced that in the redshift range 0.05 to .2,

$$\frac{\dot{R}_r}{R_r} = 2.2 \left(\frac{R_r}{R_s}\right)^{2/3} H$$

where R_s is the gravitational radius of the radio source and H is Hubble's parameter.

The expansion,of course, reduces the electron density and the radio power of the source. This is consistent with the explanation of the observed increases in number of radio sources to limiting flux densities as being due to a secular power decrease with time rather than a change in the number of sources per unit volume.

POSSIBLE REDSHIFT DISPLACEMENTS

CLUSTER-RADIO SOURCE REDSHIFT DISPLACEMENTS

1. CLUSTER DATA. Table 1. Fig. 1 distribution

   A 1020 Fig. 2. Cluster richness on counts

   Baum's 3 measures

2. RADIO SOURCE DATA Table 2. Fig. 3 distribution + LEQUEUX'S Mpg's

3. TABLES, $\frac{(1+z)^2-1}{(1+z)^2+1}$

Case for Bands - check 3 dimensional distributions

4. Statistics from random distributions Fig. 4. Distributions with $\frac{1}{\gamma^2}$ contraction.

   $\frac{1}{\gamma^2}$ contraction for ~~xxxxxx~~ uniform density

   TABLE 3 statistics

5. The displacement relation Fig. 5, $\log \Delta z$ vs. $\log z_{cl}$ etc.

   $\Delta z = 0.5\, z_{cl}^{3/2}$, errors, nearby and mixed systems

   Fig. 6 cl and RS together with cl corrected by displacement equation.

6. Possible Interpretations

   1. Holmberg Effect

      Em. vs. Abs., absorption blends

   2. Einstein Shift

      Instability (McVittie)

      Fit with astrophysical data

      Zwicky's masses

      Expansion cf. radii (angular), merge with Quasars

   3. Cross doppler, radio sources are "high velocity" objects

      radio and cluster relative rotation.

   4. Heirarchized Mach Principle, displacement due to RS + cl

      different ρ, different H, different q.

5. Equipartition

6. Time Dilation (see G. C. Chiu p. 9).

The "A-E" Effect

A systematic displacement of <u>bands</u> in the redshift distribution of radio sources with respect to the bands in the distribution of cluster redshifts is observed. Since, in general, the radio redshifts are derived from emission lines while the cluster redshifts are derived from absorption lines, the question arises whether this might not be some "absorption vs emission" effect.

Several possibilities may account for this A-E effect.

1. Absorption bleds vs emission sharpness.

   The blends would be systematically altered by the redshift.

2. A different Hubble parameter for radio objects than clusters.

3. A different q_o, acceleration parameter for clusters and radio source.

4. An Einstein shift

5. Equipartition

The A-E effect does not seem to be in evidence in the nearest portions of space. (Like the redshift being inoperative within the Local Group.) For values of $z \leq$, there seems to be no systematic A-E displacement. (This is approximately the region of $\sqrt[4]{2}$ law).

Reprinted from

THE ASTRONOMICAL JOURNAL VOLUME 70, NUMBER 2 MARCH 1965

Abstracts of Papers Presented at the 117th Meeting of the American Astronomical Society, held 28–31 December 1964 at Montreal, Canada

On Super-Organization Among Clusters of Galaxies. A. G. Wilson, *The Rand Corporation.*—The mean redshifts of clusters of galaxies do not appear to be distributed randomly. The small sample of available mean redshifts is consistent with the hypothesis that the clusters are located on a set of concentric shells which possess a definite relation between successive radii (Wilson, A. G., *Proc. Nat. Acad. Sci.,* **52,** 1964). If this distribution is real, the cosmological principle requires that apparent cluster distribution should be on concentric shells for all equivalent observers (i.e., observers located in or near a cluster). The actual spatial locations of clusters must then be at the intersections of the several sets of cluster-centered concentric shells. This requires structure in the *angular* distribution of cluster centers as seen by equivalent observers.

The investigation of regular structure in the distribution of clusters may be investigated further by combining the angular positions of clusters with the mean redshifts to generate additional "quasi-redshifts" by triangulation. If a linear redshift-distance relation and Euclidean space are assumed, the quasi-redshifts may be derived by the ordinary law of cosines. If these assumptions are valid and if the spatial distribution of clusters is regular the frequency distribution of quasi-redshifts should be a set of discrete peaks or resonances which represent the allowable separations between clusters.

The frequency histograms of the quasi-redshifts show a set of peaks distributed among a "noise" background. Statistical tests show that over fifteen of these resonances are not likely to be random fluctuations, (observed occurrence minus expected occurrence$>3\sigma$). It may be inferred that at least a subset of clusters manifests structured distribution. The noise may be due to breakdown of the Euclidean and linear-redshift approximations, to the coexistence of two or more independent organizations, and/or to actual random distributions.

It is further found that the ratios of the values at which some of the resonances occur are $3^{\frac{1}{2}}$, 2, $5^{\frac{1}{2}}$, $8^{\frac{1}{2}}$, $13^{\frac{1}{2}}$, suggesting the distance ratios which obtain for closely packed spheres.

The unlikelihood of the occurrence of these peaks and ratios in distances between clusters distributed in a random uniform manner suggests either that some form of super-organization exists among the clusters or that we are observing the vestiges of a structure whose angular and linear ratios have been preserved under a uniform and isotropic expansion from a time when the universe was in a highly compact stage. The latter hypothesis if physically consistent, would be corroborative of an oscillatory or other evolutionary model.

Alternatives to the vestige-hypothesis must account for an organization extended over 10^9 parsecs, the value bounding the separations of the clusters involved. It is difficult to explain such an extended organization without the introduction of physical communication processes not at present recognized.

Nov 13, 1964

ON SUPER-ORGANIZATIONS AMONG CLUSTERS OF GALAXIES

A. G. Wilson

The mean redshifts of clusters of galaxies do not appear to be distributed randomly, but rather show a tendency to be distributed in accordance with functionally related discrete values (Wilson, A. G., Proc. Nat. Acad. Sci. Vol. 52, 847, 1964). This may be interpreted as implying that clusters are located on a set of shells which possess a definite relation between successive radii. The cosmological principle requires that all equivalent observers (observers located at clusters) should view the clusters as similarly distributed. (For present purposes we may be associated with the Virgo Cluster.) Structured radial distribution of clusters observed by equivalent observers requires structured angular distribution of clusters observed by the equivalent observers. Hence if the regularity in radial distribution of clusters is real, angular structure in the distribution of clusters should also be in evidence.

The large numbers of clusters observed in all unobscured directions in the sky renders statistically meaningless any patterns selected ab initio on the basis of angular distribution criteria. This difficulty may be avoided by invoking an independent selectivity factor. A study was made of the clusters in Abell's catalogue (Abell, G. O., Ap. J. Suppl. 3, No. 31, 1958) selected on the basis of membership in the richest classes (4 and 5). Though widely separated, these clusters have angular positions consistent with structured rather than random distribution (the details to be published elsewhere). In addition, the same distribution properties observed for the richest clusters obtain in the subset of the rich nearby clusters. These non-random angular distribution patterns lend confirmation to the hypothesis of the existence of some sort of super-organization of which the clusters of galaxies are members.

In view of the same difficulties which arise in explaining super or second order clusters as dynamic systems (Zwicky, F. Pub. Ast. Soc. Pac. 69, 518, 1957), it is completely unsupportable to postulate the existence of a dynamic system with a diameter of the order of 10^9 parsecs, the value consistent with the distances and angular separation of the clusters involved. Consequently, if the apparent super-organizations to which these clusters belong are real, they must originate through physical communication processes other than those presently recognized.

~~Observational Evidence for~~ On the Possible Existence of a Super-Organization Among Rich Clusters of Galaxies

The discretization relation observed in the mean redshifts of clusters of galaxies (Wilson, 1) may be interpreted as due to ~~suggests~~ the existence of structure in the radial distribution of the clusters. If structure in the radial distribution of clusters exists for one observer, the cosmological principle requires that it exist for all equivalent observers. From the existence of radial structure observable by a set of observers, there follows the existence of ~~observable~~ angular structure observable for members of the set. It may thus be concluded that if the redshift discretization relation is real, structure in the angular distribution of clusters should also be observable. While it is not possible to predict the exact form of such angular structure from the redshift discretization relation, the presence or absence of angular structure may be taken as a supplementary test for the reality of redshift discretization, which at present is based on too small a sample for large statistical confidence. If both radial and angular structure in the distribution of clusters is found to exist this would in turn ~~simply~~ imply the existence of a super-organization or configuration governing cluster space-time distribution.

The question of the existence of super-organizations among clusters of galaxies has been an open one for several years. Controversy has focused on the existence of second order clusters ~~as special cases~~ rather than on more general cases of super-organizations. Although there is considerable evidence to support apparent super clustering (Abell 2, de Vaucouleurs, 3), Zwicky (4) has objected to the reality of such super-clusters as physical systems because the expected velocity dispersion among the ~~number~~ member clusters of a super cluster based on estimated masses and scales is much greater than is observed. Zwicky concludes that super clusters as physical systems ~~do~~ can not exist.

References:

(1) A. G. Wilson, Discretized Structure in the Distribution of Clusters of Galaxies -- Proc. Nat. Acad. Sci., Vol. 52, No. 3, pp. 847-854. Sept. 1964.

(2) P. Maltby, J. A. Mathews, and A. T. Moffet, Brightness Distribution in Discrete Radio Sources IV, A Discussion of 24 Identified Sources. Ap. J., Vol. 137, p. 153, January 1963.

(3) M. Schmidt, Ap. J.

(4) F. Zwicky, Non-Uniformities in the Apparent Distribution of Clusters of Galaxies, Proc. Ast. Soc. of the Pac., Vol. 69, No. 411, p. 518-529, December 1957.

1) WILSON, A.G. Discretized Structure in the Distribution of Clusters of Galaxies. – Proc. Nat. Acad. Sci. vol 52, No. 3, pp 847-854 Sept. 1964

4) Zwicky, F. Non-Uniformities in the Apparent Distribution of Clusters of Galaxies – Proc. Ast. Soc. of Pac. vol. 69, No 411, p 518-529 Dec. 1957

2) Abell, G.O. Galaxies – Second-Order Clustering Ast. Journ. vol. 66. No 10. p. 607-613 Dec. 1961

3) de Vaucouleurs, G. Clusters and Superclusters – Ast. Journ. vol 66, No 10. p. 629-632, Dec. 1961.

Just, Kurt Ap. J. vol 129, p. 268

Maltby, Matthews, Moffet Ap. J. vol. 137, p 153, Jan 1963

D. KIRBY.
ENG. DIV.

P-1425

PROLOGUE TO A SYNTAX OF SPACE EXPLORATION*

A. G. Wilson

P-1425

July 15, 1958

*Paper read at the Second Lunar and Planetary Exploration Colloquium at The RAND Corporation on July 15, 1958.

The RAND Corporation
1700 MAIN ST. · SANTA MONICA · CALIFORNIA

During the pre-space era prior to October 4, 1957, an era which some of you present can probably vaguely recall, the activity of space enthusiasts had to be spent about half on research and half on salesmanship. These demi-researcher-demi-promoter types would like to ask themselves the questions we plan to discuss today: What do we want to learn in space and what is the best way to find out. But after realizing the futility of this approach they would end up by asking the really important question: What can we sell on space. In this connection the space promoters rivaled Madison Avenue in their contributions to the science of motivation research. Almost every type of probable and improbable gimmick was used as a lure to cast before the military and civilian fund dispensers in order to make a sale. Gimmicks included even such things as the fountain-of-youth packaged in a relativistic paradox and giant satellite mirrors for incinerating cities. (This latter was the space version of Madison Avenue's selling the sizzle instead of the steak.)

But with the sudden and overwhelming success of the oldest and best sell, viz., keeping-up-with-the-Joneses, which followed October 4, the purposes and cross-purposes of space flight became very confused in the minds of many people. Much of this was undoubtedly the result of some of these earlier sales programs; and it soon became apparent to many that what must be undertaken was a job of unselling on certain ideas; the type which were harmless so long as their execution was impossible, but very questionable if actually realizable. What was needed was a responsible, carefully planned, well-coordinated, and of necessity international, program for space exploration. And now, first steps are being taken in this direction. For example, the action taken recently by the National Academy of Sciences and the International Council of Scientific Unions in calling an international meeting, including the Russians, at The Hague, which resulted in the appointment of Professor M. Florkin of Liege as chairman of a committee to review the problem of the contamination of celestial bodies, a problem whose importance is only now beginning to be generally recognized.

It seemed appropriate to the present meeting that before discussing specific experiments and instrumentation for a scientific exploration of the moon that a few remarks be interjected concerning this problem of arriving at a unified scientifically responsible program, or syntax, for space exploration.

The first decision which must be made in defining a coordinated total space program concerns the relative role to be assigned to the rival motivations behind space exploration. Some of these contesting motivations are left over from the sale-of-space program, and some arise from the basic drives and interests of various groups; and all now compete for the position of charting the course in the exploration of space.

First, there are the motives arising from competition. Whenever individuals, corporations, or governments undertake to extend their domains of operation and influence, they are primarily motivated by objectives designed to affect their status within the structure of human relationships, competing for power, prestige, or wealth. In the light of past experience, it is certainly open to question whether any organization can successfully direct so vast an undertaking as the exploration of space when guided primarily by intra-human-affair criteria such as political or military. For regardless of what political and military advantages are apparently to be gained, these stand rapidly to become obsolete when considered against the background of the revolution in knowledge and perspective which will follow the successful penetration of space; and the space program will find itself tied to goals which are no longer valid.

Secondly, the old standby, the profit motive. An example of the profit motive in action in exploration is the case of a recent discovery in southern Arizona of a small crater about 20 meters in diameter believed to be of meteoritic origin. Experts were called in and plans laid for a thorough scientific exploration of the crater, measurements of size, depth, shape, mineralogical surveys, probing, sampling in situ, etc. Everything possible was to be learned about this most interesting little crater. But before the scientists could return and begin their systematic study, some fellow who had heard that meteorites had a good sale price took a bulldozer and completely obliterated the crater searching for specimens.

Thirdly, there are motivations arising merely from capability. A boy with a sling shot is motivated to break windows just because he has the means for doing it. This motivation has often played a role in exploration also. In the Southwest there are numerous archeological sites which, even though protected by the Antiquities Law, have been despoiled by pot hunters and much invaluable scientific data lost because irresponsible explorers happened to have spades and found an interesting place to dig.

Finally, there are the motivations arising from man's aspirations for growth, his spirit of adventure, and his innate curiosity concerning the universe. It is here that the scientific objectives of space flight find their first cause.

Each of these contending motivations has its spokesmen, and the decision of whether to be guided by political, military, scientific, or what have you considerations is now being worked out. It is probably the belief of all scientists and a great majority of non-scientists that scientific considerations should play the dominant role in guiding the space effort. But there is a curious thing. To spend effort making a case for placing scientific considerations in the dominant role in space exploration seems on the one hand to be entirely tautological - everybody believes in science just as everybody believes in Mother - save your breath, it is not necessary to sell science any more. But on the other hand in the growing space jungle of interservice rivalry, ARPA, NASA, private industry, etc., to say nothing of the Russians, to preach giving the pursuit of knowledge first weight in the face of what are termed more urgent and realistic considerations, is to be guilty of the most naive idealism. This sort of contradiction is not easy to explain. It reflects an irrational facet of human nature, a facet which has been with us a long time. We may all earnestly desire to bring about certain ends, then when they fail to materialize, we look about and wonder who has worked to defeat them - and discover ourselves. But be it tautology, naiveté, or tautological naiveté, the stakes in space will be high, and science will have to formulate its own plans for space exploration and will have to continue to sell them.

So, let us look at science's in-house problem. Here the importance of scientific values is postulated unequivocally, but there is still the problem of establishing some method of handling the conflicts which will arise from competing experimental projects, and to seek how to avoid the despoiling of one type of scientific data while searching for another. It becomes necessary to define a unified responsible approach to experimentation in space, enumerate the significant criteria, and decide which criteria are to take precedence in arriving at decisions with regard to purely scientific objectives.

Ideally, it would be desirable to have a general exploratory strategy from which the plans for exploring any specific unknown environment, such as a new planet, could be derived. Such a strategy or syntax would allow for all foreseen contingencies, minimize errors, and be the plan most adaptable for overcoming and exploiting unforeseen contingencies — in other words, it would consist of the pre-application of complete hindsight. This order for a definitive syntax of exploration of course will be capable of formulation only from the experience of exploration of large numbers of diverse type of celestial bodies; but nonetheless because of the existence of conflicting requirements today, we are called on here and now to formulate a preliminary syntax based on an experience in exploration limited to a series of fragmentary exploratory efforts on a single planet. We can only hope that principles drawn from this limited experience will help to lead to the criteria needed for the making of decisions in space exploration.

The general pattern which seems most often to have governed the exploration of, and the extension of human influence into, a new region can be outlined as follows:

1. The assembling of all pre-knowledge available concerning the region to be explored, which can be obtained by indirect methods. In the case of space, this pre-knowledge comes mostly from the researches of the astronomer and the astrophysicist.

2. After the assembling of the pre-knowledge, extensions of what is known must be made for specific exploration purposes. For example, the astronomer has not been concerned with landing a space ship on the Moon

and has not made all of the observations or performed any of the simulation experiments which would be useful for this purpose. This is an operation which is only now being proposed or conducted by non-astronomical research organizations.

3. From the available pre-knowledge and its augmentation follows the design and construction of the preliminary exploratory equipment to be used in direct exploration. This includes the designs of experiments to obtain the knowledge required for more sophisticated exploratory equipment and subsequent experiments.

4. The evaluation of the new knowledge acquired from the preliminary probes and the revision and modification of both hardware and plans for experiments on the basis of the new data.

5. The iterative repetition of Step 4, each repetition leading to an extension of knowledge, refinement of data, and the construction of new devices for promulgation of human influence and control over the alien environment.

These five steps are basic and automatic for any approach depending on successive approximations. And for a long time successive approximations in hardware and experimentation will be the only method available for exploration. Hence, it is found that criteria derived from scientific considerations are inextricably interlocked with criteria decreed by state-of-art.

Because of this interlocked aspect it is one of the first problems of a syntax to decide the relative roles of scientific criteria and feasibility criteria. It would be easy to permit state-of-art criteria to determine the entire manner in which space exploration unfolds and not bother about determining the proper sequence of events which would be dictated by scientific considerations. For example, hit the moon with an H-bomb because the state-of-art permits it, regardless of whether there is any reasoned scientific justification for it. But this irresponsible have-spade-will-dig philosophy of exploration has already proven ruinous to science on many occasions and it is an attitude to be on guard against.

It is clear that the freedom afforded by a very advanced state-of-art would allow close adherence to a syntax of completely independent scientific criteria. But in the present case where state-of-art and scientific exploration are proceeding more or less hand in hand, the latitude of departure from purely state-of-art criteria is limited. However, this does not vitiate the value of having a scientific syntax. For within the constraints decreed by state-of-art at any given time, there is always the freedom to do or not to do a given experiment and there is the freedom to do it now or wait until later. It is these freedoms which make meaningful the formulation of a syntax independent of feasibility considerations.

Further, planning is never constrained by feasibility. If it were, progress would halt. Planning permits extrapolation in state-of-art inputs, with the assumption of items not definitively known but reasonably anticipated. Only execution must be totally bounded by the state-of-art constraints. Hence, for the guidance of planning an independent scientific syntax must also be derived.

This question of scientific versus feasibility criteria is pointed out trenchantly in a lunar experiment proposed by Lederberg and Cowie.(1) I quote from their article in the June 27, 1958, issue of Science.

> Moon dust is cosmic material captured by the moon's gravitational field and presumably left undisturbed by atmospheric and biological

alteration. It should therefore contain a continuing record of cosmic history as informative with respect to the biochemical origins of life as the fossil-bearing sediments of the earth's crust have been in the study of its later evolution. (Ref. 1, p.1473)

Since the sending of rockets to crash on the Moon's surface is within the grasp of present technique, while the retrieval of samples is not, we are in the awkward situation of being able to spoil certain possibilities for scientific investigation for a considerable interval before we can constructively realize them. (Ref. 1, p.1473)

At the present pace of missile development we urgently need to give some thought to the conservative measures needed to protect future scientific objectives on the moon and the planets. (Ref. 1, p.1474)

Here the space planners are faced with a difficult decision, which may not prove to be critical in the case of the moon, but which illustrates a basic problem which may frequently come up in the exploration of space. Whether to perform experiments A and B as soon as is feasible and risk perhaps destroying the chance of ever performing experiments C and D, or postpone A and B until state-of-art will allow C and D to be successfully studied. On what basis should this decision be made?

Another factor on which the relationship between the scientific and feasibility criteria depends is whether the evolving state-of-art is determined directly by the scientific requirements or by some other requirements. For best scientific results hardware should be developed explicitly from the research requirements, rather than experiments being tailored to a hardware which evolves primarily according to extra-scientific considerations.

By now it should be apparent why a syntax is needed, why criteria must be decided upon and weighted, and what some of the basic difficulties in decision making will be. As one proceeds into the formulation of a syntax, the following questions will have to be analyzed:

- What do we wish to find out?
- What are the possible experiments for finding these things out?
- What is the best place to perform the experiments?

o What is the best time to perform the experiment with regard to the present and anticipated state-of-art?

o What is the bearing of the information received from the experiment on future experiments and on the subsequent development of the state-of-art?

What we wish to find out, what experiments we wish to perform, and what questions we wish to ask are, of course, a function of what we now know. Certainly we in 1958 can ask better questions about what to find out about the moon than Pythagoras or Archimedes could, had they been given the opportunity for space exploration. Nonetheless we are uncomfortably aware that we still may not be asking the right questions and that our plans for experimentation may result in unfortunate and unplanned consequences such as the premature contamination mentioned by Lederberg. This demands flexibility in the syntax.

In addition to setting up a systematic approach to the acquisition of new facts, the syntax of space exploration should seek the detection of presently unknown parameters. This must be mentioned because instrumentation designed to measure magnitudes of known parameters is not likely to detect new parameters. Let us take a very simple-minded example of this. Suppose we are familiar with all the properties of the circle but are not familiar with the third dimension and with bodies of circular cross-section like cones and cylinders. We are about to make our first excursion from the plane of the blackboard into three-dimensional space and are designing instruments to measure the radii of circles we have long observed in the plane of another blackboard. But in leaving the plane, for the first time the conical or cylindric natures of our circles are exposed to detection. Will our radius measuring devices discover these new attributes?

So in addition to instrumentation for measuring familiar aspects of phenomena, it is necessary to lay the groundwork for the search for new parameters whose detection may be possible for the first time because of being in space. Thus, research on how man thinks and how he develops awareness of hitherto unknown attributes by noting new differences and new similarities, also constitutes a basic part of the exploration of space and should find some place in the syntax.

Many possible criteria and other aspects of a scientific space syntax have undoubtedly occurred to each of you, and at some future date it is hoped this group may wish to formulate a syntax for space exploration. But before closing this prologue, I would like to point out two other types of mistakes which have been made in past explorations.

In an article in the May 31, 1958, issue of The Saturday Review[2] forty of history's most villainous characters were singled out for an all-earth, all-time Rogues' Gallery. The selections ranged from Herod to Himmler. But it was of special interest that in this list of forty who contributed so much to human suffering, it was seen fit to include a man whose name is unknown but who is nonetheless responsible for one of the major crimes of history: The unknown sailor of Columbus' crew who infected Europe with syphilis from the New World. To his case may be added the atrocity of Bishop Landa who burned the Mayan Codices in an arrogant attempt to eradicate the cultural contributions of a less advanced civilization.

It is not clear at the present time when, where, or if space explora- will find analogous opportunities to repeat these blunders. But it is hoped that man will have accumulated enough wisdom when these opportunities come that space explorers will add no new names to the list of all-time villains.

REFERENCES

1. Lederberg, J., and D. B. Cowie, "Moon Dust," Science, Vol. 127, No. 3313, June 27, 1958, pp. 1473-1475.

2. Ritner, Peter, "Rogue's Gallery," The Saturday Review, May 31, 1958.

THE SPACE ENVIRONMENT

by

A. G. Wilson

P-1427

February 24, 1958

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

LECTURES FROM "AN INTRODUCTION TO ASTRONAUTICS" PUBLISHED AS RAND PAPERS

Lecture	Title and Author	Paper
2	The Space Environment, by A. G. Wilson	P-1427
3	Trajectory Fundamentals, by S. Herrick	P-1303
4	Types of Flights, by R. W. Buchheim	P-1428
5	Propulsion Fundamentals, by B. Pinkel	P-1429
7	Materials for Space Flight, by G. A. Hoffman	P-1420
9*	Aerodynamics for Space Flight, by E. P. Williams and C. Gazley, Jr.	P-1256
10	The Penetration of Planetary Atmospheres, by	P-1322
13	Orientation and Control, by T. B. Garber	P-1430
14	Some Information Theory Considerations in Space Communications, by P. Swerling	P-1393
15	Communication in Space Operations, by C. M. Crain and R. T. Gabler	P-1394
16	Internal Environment of Manned Space Vehicles, by S. H. Dole	P-1309
17*	A Discussion of Energy Sources for Space Communications, by J. H. Huth	P-1318
19	Space Flight Ground Facility Requirements Problems--Launching Facilities, by J. J. L'Sullivan	P-1431
20	Scientific Exploration in the Fringe of Space, by W. W. Kellogg	P-1350
21	Physics of Solar-Terrestrial Space: Lunar Flight, by E. H. Vestine	P-1344
22	Interplanetary Exploration, by A. G. Wilson	P-1432
23	Some Elementary Facts of Relativity, by H. A. Lang	P-1433
24	Outer Space and International Law, by Leon Lipson	P-1434
25	Some Political Implications of the Space Age, by J. M. Goldsen and Leon Lipson	P-1435

*The paper is based on unclassified portions of the lecture.

THE SPACE ENVIRONMENT

by

A. G. Wilson

Present astronomical knowledge regarding the solar system is reviewed with emphasis on facts which may be of importance to the astronaut. The basic difference between space environment and terrestrial environment are set forth, and the material content of space which might offer a collision threat to a space vehicle is discussed.

Those who first venture into space will, unlike earlier navigators pushing across unexplored seas, find that much of the region to be traversed has already been charted and something of the character of both space itself and potential destinations in space is known. But there is always the difference between indirect knowledge and first-hand experience, the difference between reading about how to fly an airplane and actually piloting one, and this difference undoubtedly will show up trenchantly on the first flights into space. However it is useful to review briefly some of the present astronomical knowledge concerning the solar system with emphasis on facts which may be of importance to the astronaut.

Before entering into descriptive details, a few important basic differences between space environment and terrestrial environment should be mentioned and kept in mind in our discussions of space. First, the configurations of bodies in space are never static; relative distances are always changing. Second, the description of the solar system in terms of distances alone, is inadequate. The astronaut must think also in terms of all the orbital elements: the eccentricities, the inclinations, the nodes, the epochs, and the perihelions as well as the semi-major axes.

The third general difference is the relation between energy expended and distance traversed. In space this will be completely unlike anything in terrestrial experience. The laws of motion and energy are the subject of one of the lectures in this series.

Fourth is the matter of the scale of space. It is always most difficult to visualize the tremendous distances involved. The Moon, which is the closest sizable body, is over 30 earth diameters distant, while the sun is 12,000 earth diameters away. One of the most convenient methods for expressing the scale of space is by the time required for light to move from place to place. Light traveling with a velocity of 186,000 miles per second can circle the earth about seven times in one second. It traverses earth-moon space in 1.3 seconds; goes to the sun in 499 seconds (or 8.3 minutes); and goes from the sun to Venus in 6 minutes, to Mars in 12-1/2 minutes, to Jupiter in 43 minutes, and to Pluto in 5-1/2 hours - and, leaving the solar system, to the nearest star in +4.25 years.

A fifth difference which the astronaut must bear in mind is that space travel will be performed in vehicles which are intermediate in size between the small particles in free space and the massive planets. While the motions of the latter are influenced only by gravitational forces (Newtonian and relativistic), the small particles are, in addition, subject to magnetic, electrical, and radiation forces. It is to be expected that future space ships will, as intermediate-sized bodies, experience to some extent the effects of all of these forces.

One of the most important aspects of the space environment deals with the material content of space. A discussion of the properties and behavior of the smallest particles, atomic particles, free electrons, nuclei, and molecules will be the subject of a later lecture. Let us first consider somewhat larger bodies, ones in the range from cosmic dust to chunks of rock (i.e., say 20 microns to a few meters) commonly called meteoroids. These bodies are sufficiently numerous and of such size that encounters and collisions with a space vehicle become probable.

The definitive answer to the question of the likelihood that a space ship, satellite, or missile, or any body on trajectory outside the earth's atmosphere, will encounter a meteroid must await data actually obtained in space. The best answers that can be given today are derived from optical or radio observations of meteors made from the earth's surface.

But at this point a digression is needed to introduce an important unit of measure with which the astronaut should be acquainted: the astronomer's measure of the brightness of heavenly bodies by means of stellar magnitudes.

If L_1 and L_2 are the brightnesses, luminosities, or luminous energy radiated per unit time of two bodies, then the corresponding magnitudes m_1 and m_2 are defined by the relation $0.4(m_2 - m_1) = \log_{10} (L_1/L_2)$. Magnitude measure is thus a logarithmic measure of brightness ratio. A difference of one magnitude is equivalent to a brightness or luminosity ratio of $\sqrt[5]{100} = 2.512$; a difference of 5 magnitudes is equivalent to a ratio of 100 in brightness. Using this scale, the brightest stars in the sky have apparent brightnesses measured in magnitudes of about 0 or −1. The faintest stars which can be seen by the naked eye are about +6. The North Star is +2, Venus −4.3, the full moon −13, the sun −26. The faintest star detectable in the 200-inch telescope has a magnitude near +23 or about 10^{-8} as bright as the North Star.

Figure 2-1, based upon the observational and theoretical results of the Harvard Meteor Program, gives the mass and size of meteoric particles as functions of the visual magnitude. Figure 2-2 gives the number of such meteroids striking the earth per day, and the number striking a 3-meter sphere in the neighborhood of the earth per day. This last number is derived from relative sizes of the earth and sphere with the introduction of an earth shadow factor of one-half.

It is estimated by Whipple that a meteroid of magnitude 17, moving with a velocity of 18 km/sec, of which about two per day will strike a 3-meter sphere, will penetrate an aluminum skin of 0.01 cm, whereas a meteoroid of magnitude 5, one of which will strike the sphere every hundred years, would penetrate 4.5 cm of aluminum. A critical size would be one which penetrates 0.5 cm of aluminum. One this size will hit the sphere about every 50 days.

But the probability of striking meteroids depends upon where the vehicle is in space. Figure 2-2 applies to the immediate neighborhood of the earth. What about meteroid distribution at greater distances? Here good data are lacking. What is known, however, is that (a) the smallest dust particles (micrometeoroids) are concentrated in the ecliptic or plane of the earth's orbit, and (b) most meteoritic material is cometary refuse and is consequently largely distributed along the orbits of comets.

Let us first review some of the evidence for the ecliptic concentration of cosmic dust. After evening twilight, especially near the 21st of March in northern latitudes, a faint tapered band of light can be seen extending up from the horizon centered along the ecliptic. This band of light, which can be photoelectrically traced through the complete night sky, is called

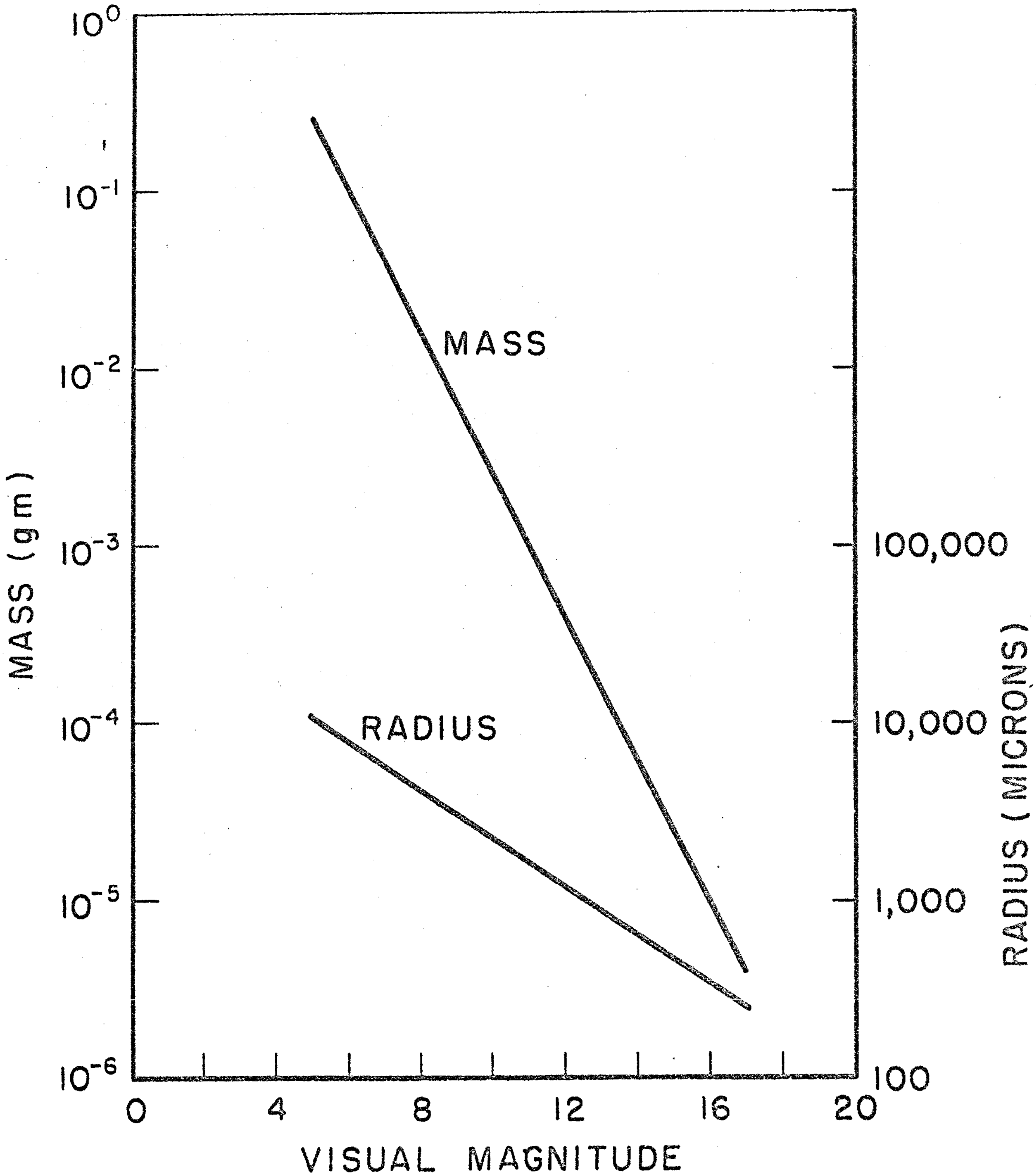


FIG. 2-1

METEOR BRIGHTNESS VS SIZE

P-1427
2-24-58
6

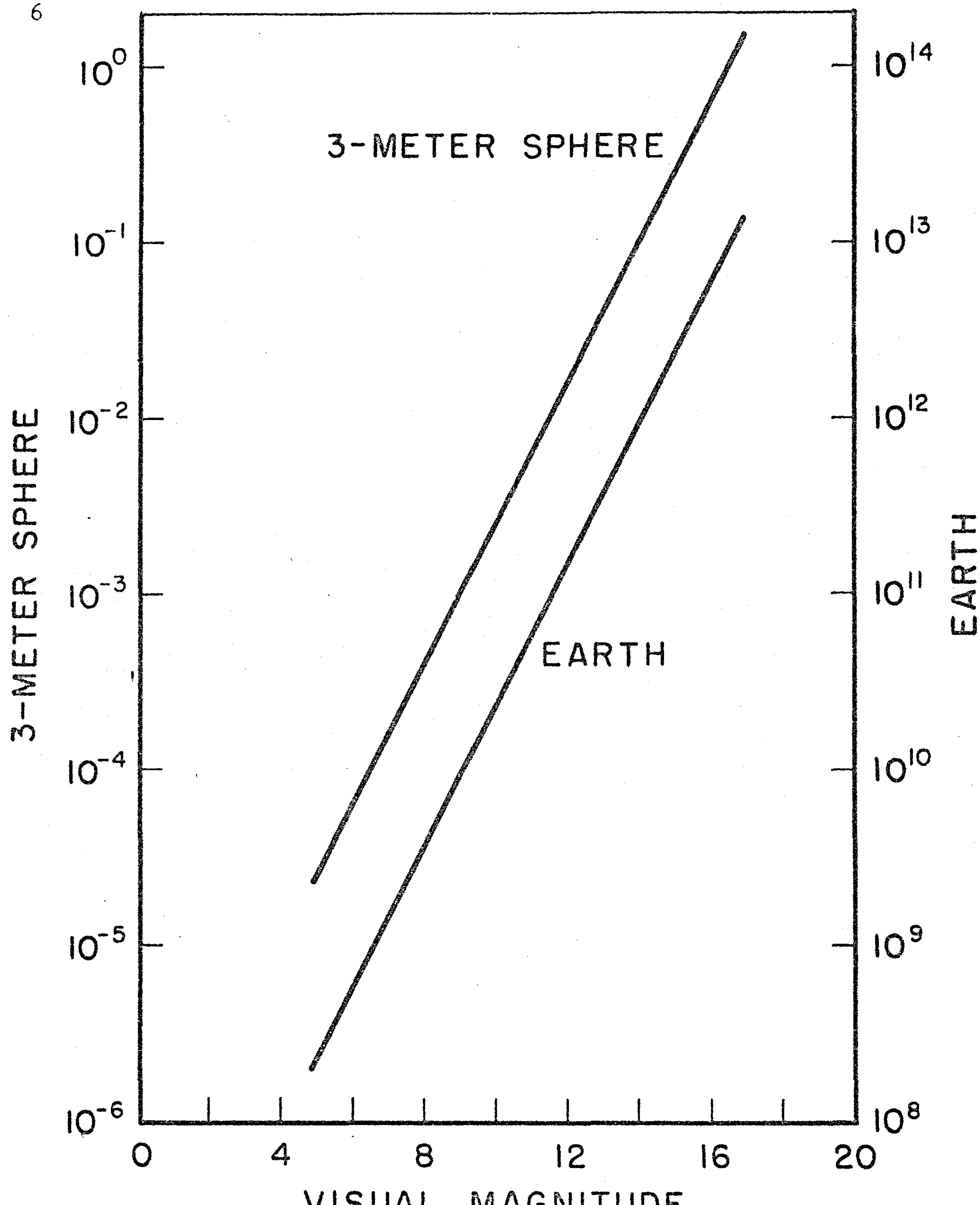


FIG. 2-2
METEOR IMPACTS
PER DAY

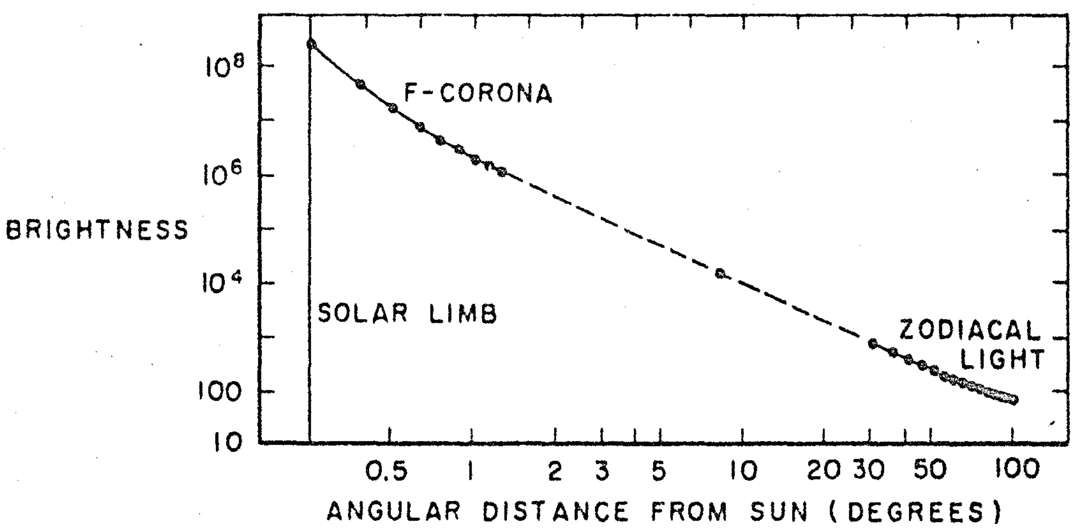
the zodiacal light. Spectroscopic observations of this light show a continuous spectrum like that of scattered sunlight. The color of the zodiacal light is nearly the same as that of the sun but shows approximately 20 per cent polarization. These observational facts suggest that the zodiacal light is caused for the most part by sunlight scattered from small dust or meteoroidal particles at least 20 microns in diameter. Since light scattered by free electrons is strongly polarized, it is probable that free electrons represent a fraction of the particles present. This is also substantiated by the fact that the total light present seems to vary with solar activity, being least when ionizing radiations from the sun are at a minimum. However, since scattering by gas atoms and molecules alters the color of the light, it must be concluded that the zodiacal particles (except for the free electrons) are much larger than molecules.

It has been suggested that the zodiacal light is an extension of the Fraunhofer or outer solar corona. This idea is reinforced by the fact that the corona has a color and continuous spectrum agreeing with the zodiacal light. But most interesting is the comparison of the brightnesses, as shown in Fig. 2-3.

This lenticular layer of small meteoroidal particles must extend from the sun well beyond the orbit of the earth, being concentrated toward the ecliptic or fundamental plane of the solar system. Further, this dust cloud is probably continuously being resupplied by cometary wastage and possibly by material from asteroid collisions. It is at the same time being drained off by the action of the Poynting-Robertson effect, which causes the particles to spiral in toward the sun. It has been estimated that in 60,000,000 years all particles smaller than 1 mm within the diameter

FIG. 2-3

CHANGE OF BRIGHTNESS OF THE SUN'S OUTER CORONA AND THE ZODIACAL LIGHT WITH DISTANCE FROM THE SUN



of the orbit of Mars would reach the sun due to this effect.

One of the most interesting features of the zodiacal light is its broadening and intensification in the night sky exactly opposite the position of the sun. This phenomenon, called the counterglow or gegenschein, consists of a very faint patch of light extending about 10 degrees along the ecliptic and for 6 degrees at right angles. One explanation of the counterglow is that in this area the particles are seen in full-phase illumination and are consequently much brighter. Another explanation is that radiation pressure pushes a tail of atmospheric particles from the earth out into space opposite the sun. But most probably, the counterglow is due to the concentration of interplanetary particles at one of the earth-sun Lagrangian or libration points. These points are points of metastable equilibrium which occur in the three-body problem. There are five such points; three lie on the line joining the earth and sun, and two form equilateral triangles with the earth and sun in the plane of the ecliptic. Small bodies may become trapped in these "gravitational sinks" until a perturbing force allows them to escape. Such may be the explanation of the counterglow.

The major concentration of the smallest meteoric material (producing no visual effects when striking the earth) is in the ecliptic, but other concentrations are intimately associated with comets and other bodies. The visible meteors, or shooting stars, are of two types — those associated with showers and those which are sporadic. The shower meteors are of cometary origin; the sporadics are probably traceable to asteroids.

Next let us review a few facts concerning comets and meteor showers.

Cometary orbits are of two general types, parabolic and periodic. No comet has yet been observed with a definitively hyperbolic orbit. This means that most comets, at the present time, must be a part of the solar system and not visitors from interstellar space. However, at some past time they may have been captured by planetary perturbation after orbiting in from outer space. Comets on so-called parabolic orbits probably really travel in exceedingly elongated ellipses with periods running to centuries. (A comet with a major axis equal to the distance to the nearest star would have a period of 100 million years.) These comets which travel in parabolic orbits are usually highly inclined to the plane of the ecliptic, half even being in retrograde orbits. The mean perihelion distances of parabolic comets is about the same as the earth's distance from the sun.

The periodic comets have orbits more like those of the planets. Only one periodic comet (Halley's) travels in a retrograde manner, the inclinations of most being less than 45 degrees. The eccentricities are commonly in the neighborhood of 0.5 (the most eccentric planetary orbits are less than 0.25). A few periodic comets have their perihelion within the earth's orbit, while most aphelions are near the orbits of one or the other of the major planets. Some median values for periodic comets are: period, 7 years; semi-major axis, 3.6 astronomical units (a.u.); perihelion distance, 1.3 a.u. Thus meteoric material driven from comets by solar radiation or thermal processes may be found almost anywhere in the inner parts of the solar system, with greatest concentrations not far from the earth's mean distance from the sun.

Some of the most interesting facts about comets are associated with their changes in appearance as they come near the sun. A comet usually consists of a bright nucleus surrounded by a fainter luminous envelope. While the typical head, nucleus plus coma, is of the order of the size of

Jupiter (100,000 miles), the nucleus is only a few kilometers in diameter, reaching a size as large as 1000 km only in rare cases. As the comet approaches the sun a tail is usually forced out by the radiation pressure of sunlight. The tail sometimes extends millions of miles at maximum size. The tail of the great comet of 1843 stretched twice the distance from the earth to the sun.

No accurate masses of comets have been determined since they are not massive enough to exert any measurable perturbative forces on other bodies. But is is estimated that typical masses are of the order of 10^{12} tons (earth ≈ approximately 10^{21} tons), and the densities are such that in a thousand cubic miles of a comet's tail there is less matter than in a cubic inch of air.

In 1949 Whipple hypothesized a comet-model which satisfactorily explains a great many observed facts about comet. Whipple holds that a comet's nucleus is a cosmic iceberg, a porous mass of solidified gases or ice plus some solid particles. The substances present are largely water ice, ammonia, and methane with some carbon dioxide and cyanagen. There is also the possibility of the presence of free radicals. As the comet approaches the sun, these gases evaporate, sending out jets which form the coma and tail. Some of the energy of the jets may come from the free radicals.

The structure of the tail is determined by several forces: orbital momenta of the particles, ejection velocities on evaporation, radiation pressure, gravitational, and possibly other fields. Oftimes the tail shows a large amount of structural detail due to a sort of "mass spectrograph effect" separating particles of different size.

But what is of special interest is that on each trip near to the sun, the comet is partially disintegrated by the action of these forces and leaves a "wake" of small solid particles and ices. So the regions of space where an astronaut is likely to find higher than average densities of meteoric material are along the orbit of comets, either "live" comets or old disintegrated comets.

Whenever the earth passes through one of these cometary wakes a meteor shower results. Hundreds of shooting stars are observed to emerge from a small area of the sky called the radiant, the direction being determined by the orbit of the comet wake in space. As the wakes or meteor streams become older, the particles are spread out thinner and both the size of the radiant and the period of time over which the meteors are observed increases. In general these small solid particles or bits of ice, a few microns in size, which cause meteor showers will not cause penetrative disasters to a space vehicle, but they may in time cause considerable skin attrition. It is the sporadic meteoroids that are likely to cause trouble in space flight. These bodies are most probably fragments of asteroids which have resulted from collisions. Like comets, none seems to have a definitively hyperbolic orbit. However, these sporadic meteoroids may be quite sizable, form fireballs, and frequently strike the earth. (No meteor observed during a shower has been known to strike the earth.) They range from a few grams up to thousands of tons like the large meteorites (or even small asteroids) which cause craters like the Barringer Meteor Crater in Arizona.

Let us now review briefly a few facts concerning the minor planets or asteroids themselves. Since the discovery of the first asteroid on January 1, 1801, the orbits of more than 1500 of these bodies have been determined. However, their total number must run into the hundreds of thousands, it having been estimated that there are 80,000 brighter than the 19th magnitude alone. Only those which have been observed sufficiently for an orbit to be computed are catalogued and assigned numbers and names. Most of the asteroids follow orbits which lie between the orbits of Mars and Jupiter, occupying a place in the solar system where Bode's Law predicted a major planet which does not exist. But some asteroids depart considerably from the mean orbits. At one extreme there is Hidalgo (944), having the largest asteroidal orbit known. It has a perihelion at 2.0 a.u., not far beyond the orbit of Mars, and an aphelion of 9.6 a.u., nearly at the solar distance of Saturn. At the other extreme is Icarus, which has a perihelion of 0.19 a.u., taking it well within the orbit of Mercury (0.39 a.u.), and an aphelion of 1.98 a.u., beyond Mars.

A study of the distribution of the orbits of asteroids shows that many are grouped in families. But perturbations over millions of years obscure the original picture and possible clues to the origins of the asteroids. One family of asteroids is of special interest. It occupies the equilateral Lagrangian points in Jupiter's orbit (Fig. 2-4). These asteroids – known as the Trojans – number about 12, some leading Jupiter, some following. Searches have been made for possible Trojan-type asteroids associated with the equilateral Lagrangian points in the orbits of other planets, but none has been found.

The distribution of the periods of the asteroids shows a series of gaps (Fig. 2-5). These are the effects of the perturbations of Jupiter. The period of Jupiter is 11.9 years. It is found that there are no asteroids with periods of 5.95, 4.76, and 3.97 years, i.e., exactly 1/2, 2/5, and 1/3 of Jupiter's period. There are also depressions in the distribution

FIG. 2-4

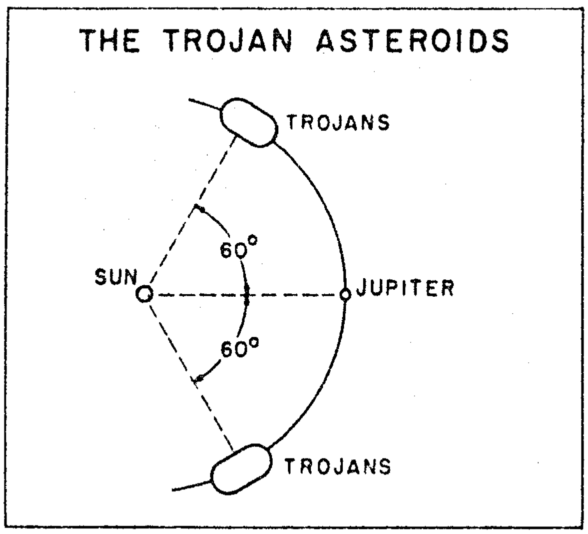
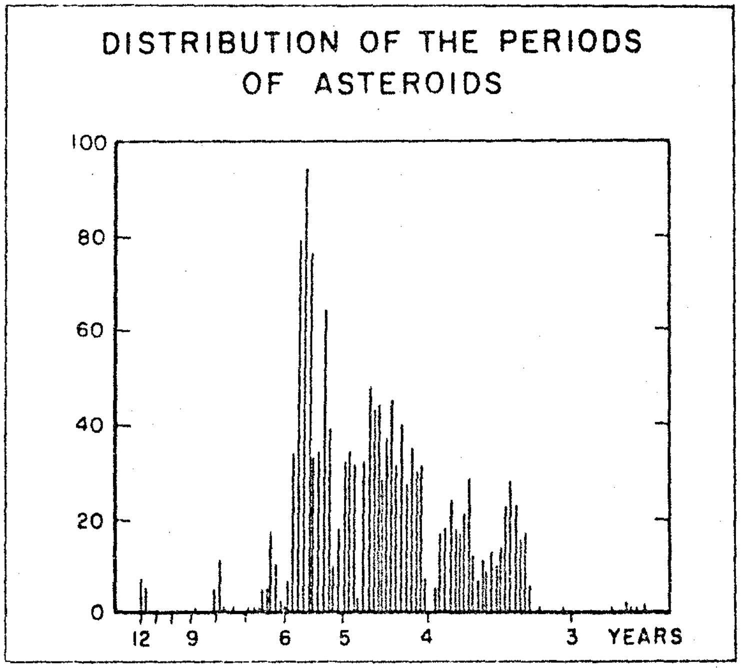


FIG. 2-5


DISTRIBUTION OF THE PERIODS
OF ASTEROIDS
100
80
60
40
20
0
12
9
6
5
4
3
YEARS

curve for orbits with periods which 1/2, 1/5, 3/5, and 3/7 of Jupiter's period. Orbits whose periods are exact fractions of Jupiter's period are called resonant orbits. The effect of perturbations on these resonant orbits is to render them unstable and force the asteroids into other orbits, a fact which might be of interest to astronauts; similar effects would operate on earth satellites whose periods were exact fractions of the lunar period. Thus if a satellite were placed on an orbit with a period of say exactly 1/4 a month, it would soon move into some other orbit.

In recent years high-powered, wide-field photographic telescopes have recorded thousands of faint new asteroids, some of them on orbits which bring them close to the earth. In 1932 an asteroid, later named Apollo, passed within 3,000,000 kilometers of the earth. In 1936 an asteroid was discovered which passed at only 1,000,000 km, and in 1937 an asteroid swept within 800,000 km or roughly twice the moon's distance. Orbits are now known for at least 10 such objects which come within the earth's orbit. Undoubtedly there are scores more, and over hundreds of thousands of years collisions with the earth must occur.

What are the sizes of the asteroids? The largest asteroid (and the first discovered) is Ceres with a diameter of 730 km (Mercury is 5000 km, Mars 3476 km). The sizes range on down to a few kilometers. Assuming that the ratio of reflecting power to size is the same for small asteroids as for large ones, we have

Absolute magnitude.........	5.0	10.0	15.0	20.0
Diameter (km)..............	270	27	2.7	0.27

Since the number of bodies increases by a factor of 2.7 with each magnitude, there are probably 100,000 asteroids with diameters in excess of 250 meters. It is estimated that all the asteroids together would make up a spherical body about 1000 km in diameter with a mass less than one thousandth the earth's mass.

While large planets are stable and pursue orbits relatively unperturbed for millions of years, the asteroids suffer frequently from perturbations, and collisions occur from time to time. It is felt that all of the asteroids may have originated from the collision of two planets between the orbits of Jupiter and Mars, one planet being larger than Mars, but smaller than the earth, the other being about the size of Ceres. Subsequent collisions between the original fragments have resulted in further fragmentation and the process continues, wearing the pieces down to meteoroidal chips.

The remarks of this description of the matter to be encountered in space have been restricted to the smaller bodies of the solar system — meteors, comets, and asteroids. Later in the series some special facts which are of interest to astronauts concerning the sun, the moon, and the planets will be discussed.

But in concluding this brief survey, it might be well to say something of the larger setting in which the solar system itself figures. The nearest star outside our solar system is in the star system of αCentauri, a first-magnitude star in the southern sky about four light-years away. αCentauri is a double star whose two components orbit about one another. The star "Proxima," also associated with this system, is actually at the present time the star closest to our solar system; "Proxima" itself may be also be double. It is not known whether this star system has any planets, but observations of some other near-by stars, e.g., 61 Cygni, indicate, from wobbles in their motion, the presence of orbiting dark bodies with masses comparable to Jupiter's. So there is indirect evidence for the existence of other planetary systems. Within 20 light-years of the sun there are only about 100 stars and possibly one or two planetary systems. Kuiper estimates on the basis of the ratio of the masses of components of double stars that 12 per cent of all stars have planetary systems. When we realize that there are some 2×10^{11} stars in our galaxy, this leaves 20 billion with planetary systems. It seems reasonable to assume that out of this number there must be some with earth-like planets, and probably on some of these life similar to our own.

However, communication with such planetary systems is not conceivable with our present state of knowledge. When we recall that our galaxy is some 100,000 light years across, the sun being an insignificant star some 30,000 light years from the galactic center, circling in an orbit of its own every 200,000,000 years as the galaxy rotates, we realize that trying to grasp the nature and scale of the universe beyond the solar system is futile. Nor is interstellar space and the galaxy the end. Beyond are the millions of other galaxies all apparently rushing from one another at fantastic speeds.

Today no one of acute awareness beholding the dawn of the first space age can fail to feel that man is about to enter upon his greatest adventure. Where it will end we do not know, but whenever man's environment has been altered he has discovered hitherto unsuspected ruts in his thinking. And certainly the differences between man's terrestrial environment and the environment he will encounter in space will furnish a greater alteration than any to which he has ever been subjected. It would be surprising indeed if, with the penetration of space, many generically new concepts do not make their impact on both scientific and social thought.

From what we already know of the nature of space, our ideas have been repeatedly forced from their restrictive channels. The Copernican revolution, the discovery of the multiplicity of worlds, and, recently, the observational evidence supporting the possibility of extra-terrestrial life have all caused a profound revision in our way of looking at ourselves. Perhaps following the first flights into space, the proposal by both Khruschev and Western spokesmen to substitute the goal of the mastery of space for wars of mutual destruction can be so implemented as to serve as an adequate challenge for man's basic aggressiveness and as a yardstick of the relative effectiveness of competitive social and economic systems.

GALACTIC SCALE DISCRETIZATION II: OBSERVATIONS

RM-3771-RC

RSR PROJECT 7086

ALBERT WILSON

Draft No. 2

I. Introduction

Assuming the Einstein Field Equations

$$B_{AB} - 1/2\ Bh_{AB} = \kappa T_{AB} \qquad (A, B = 0, 1, 2, 3)$$

(B_{AB} is the Ricci tensor, h_{AB} is the ~~first~~ ~~fundamental form~~ metric tensor, $B = B_{AB} h^{AB}$ is the scalar curvature, κ is the relativistic gravitational constant, and T_{AB} is the momentum-energy tensor); and employing an averaging operator which leaves all tensor structure invariant, Edelen ~~(1963a, b)~~ [1], [2] has derived consequences of the General Theory of Relativity which ~~permit the formulation of~~ allow observational tests of the theory in its general form.

In particular, under assumptions that (a) the boundary of a galaxy may be represented by a geometrically stable time like hyper surface Σ, imbedded in a four dimensional Einstein space, (b) the time sections of Σ depart from oblate spheroidal symmetry by only small time independent deviations and (c) there exists a jump type discontinuity in at least one component of the momentum - energy tensor across the hyper-surface (but with no further assumptions concerning the momentum-energy tensor), Edelen ~~(1963a, b)~~ [1], [2] has shown that the ellipsoid defined by the constant time section of Σ has a semi-major axis which is proportional to a function $f(n,m,\epsilon,\xi)$ where m and n are integers, with $n > 0$ and $0 \leq m < n$; ϵ is the eccentricity of the ellipsoid; and ξ is a physical parameter which represents the jump in energy density at the surface of the galaxy. For ϵ equal to zero, the function f assumes the particularly simple form

$$r = \xi^{-\frac{1}{2}} [n(n+1)]^{\frac{1}{2}} \qquad (1.1)$$

where r is the radius of the galaxy. Equation (1.1) is exact for $\epsilon = 0$ for all values of m and n. However, it is also a good approximation for $\epsilon \leqq 0.1$ for all m and n (see table Ia, b) since for these ranges the derivative $df/d\epsilon$ is approximately equal to zero while at $\epsilon = 0$, $df/d\epsilon \equiv 0$.

According to present classification systems, the only morphological types of galaxies having $\varepsilon \doteq 0$ are the early ellipticals. Equation (1.1) thus may be interpreted to predict for E0 and near E0 galaxies, diameter sizes which are discretized in proportion to the eigen sequence $\sqrt{2}, \sqrt{6}, \sqrt{12}, \sqrt{20}, \ldots$, whenever ~~whever~~ ξ is constant. This interpretation of equation (1.1) will be termed the Edelen Discretization Hypothesis.

In the present paper we shall attempt to construct an observational test of this discretization hypothesis using as few additional assumptions as possible, and apply the test to various sets of published and unpublished diameter measurements.

II. An Observational Test of the Discretization Hypothesis

The ~~Our~~ first task is to determine under what conditions the discretization prediction as it is formulated in [1], [2], or in the introductory section, is testable. In order to do this it iwll be necessary to investigate the following epistemological or technical questions:

1. Is the quantity r of equation (1.1) observationally interpretable.
2. Under what restrictions on ~~What are the effects of the values of~~ the parameter ξ ~~on the~~ will ~~testability of~~ the hypothesis be testable.
3. What are the effects of unknown axis orientations on the selectability of a suitable test sample of low true eccentricity ellipticals.
4. How can observed angular diameters be converted into linear diameters which will be meaningful in equation (1.1).

Each of these questions will be discussed in turn.

1. The r of Eq. (1.1) is defined as the radius or semi-major axis of an ellipsoid whose boundary is a surface across which there is an assumed jump in some component of the momentum-energy tensor. While r is thus well defined in terms of the mathematical model, ~~when~~ the question arises, does an r defined in this way exist in the real world, and if so it can be identified with the properties of some observable quantity.

In observations to date no jump discontinuity ~~seems~~ appears to exist in any observable parameter of early elliptical galaxies. It might seem ~~in one sense~~ that this lack of observation raises an argument against the validity of the discretization prediction. But this is not so, not only ~~Certainly,~~ in that not all components of the momentum energy are observable, but in that even if they were and if no jump discontinuity were still observed, this observed absence ~~lack of observation~~ would not invalidate the arguments, because ~~since~~ the assumption ~~condition~~ of the existence of the jump discontinuty in the mathematical model is only a sufficient and ~~but~~ not a necessary condition for discretization. Hence, even if discretization exists and is observed in the real world, there is no implication ~~necessity~~ that a jump discontinutly exist in real galaxies. ~~But~~ ~~However~~

¶ Thes lack of observation, (or even ~~the~~ non existence ~~reality~~) of a jump ~~dixxxxxx~~ discontinuity, while not invalidating the theoretical argument, non-the-less does

preclude the possibility of a direct observational means of defining a diameter which would be known to correspond to the theoretical r of Eq. (1.1).

no ¶ In the absence of the availability of this direct definition, we may, nonetheless, proceed to construct an observational test if we arewilling to introduce a new assumption, namely that one or more classes of operationally defined diameters (such as isophotal, micrometric, effective, etc.) are propoetional to r. But it must be realized in introducing such an assumption that a failure to confirm Eq. (1.1) with such classes of operational diameters does not disprove the discretization hypothesis - the defect may lie in the proportionality assumption. ¶ This assumption imposes a strong additional constraint on the structure of early elliptical galaxies and constitutes a basic departure from the original discretization hypothesis. The essense of the assumption is that discretization , if it exists, should be manifest over a range of sizes rather than only at some unique size as the Edelen discretization hypothesis predicts. Such a strongly modifying assumption while undesirable is unavoidable because in the absence of observable jump discontinuities or some other structural feature leading to a unique definition of diameter, our test must be based on diameters defined by arbitrary operations on the luminosity distribution of the photographic image of the galaxy. It is the arbitrary element in the definition of operational ~~xxxxx~~ diameters (e.g. the intensity level selected for the isophotal diameter or the percent luminosity to be included within the effective diameter etc.) which ~~xxxxxxx~~ introduces this necessity of proportionality over a range.

In the event of confirmation of equation (1.1) by a set of operationally defined diameters it is probable that some ranges does exist and could be manifested if adjacent similarly defined sets of diameters were tested, and not probable that one has fortuitously found a unique set of diameters for which discretization holds. Nontheless, in the event of confirmation of equation (1.1) by a certain set of ~~xxxxxxxxxxxx~~ operationally defined diamaters,* ~~for~~ if the proportionality assumption were found to be invalid for adjacent sets of diameters some quite different conclusions would have to be drawn; and it is well to hold that without investigation of this possiblity no confirmation of the discretization hypothesis should ~~xx~~ be held as definitive.

---

*It is important that the proportionality assumption be idependently tested and a range exhibited.

In the case of operational diameters defined by isophotes at a fixed intensity, proportionality is well approximated in the outer parts of galactic images. Hubble (3) found that the luminosity profile of elliptical galaxies are well described, except in the innermost parts of a galaxy by a relation of the form

$$\log I = \log I_o - 2 \log (R/A + 1)$$

where I is the intensity, I_o the central intensity, R is the radial distance from the center, and A a scale parameter constant for each profile. When the galaxy scale parameter A has been determined the log profiles constitute a family of parallel lines defined by the single parameter I_o. It follows that at a fixed intensity, for galaxy i and galaxy j

$$\log (R_i/A_i + 1) - \log (R_j/A_j + 1) = \text{constant}.$$

and this constant is independent of the intensity picked. Therefore, in the outer parts of the profile, for $R/A >> 1$, R_i is proportional to R_j over a range of intensities.

2. A second features of Eq. (1.1) bearing on the feasibility of an observational test is the value or set of values assumed by the parameter ξ which represents the jump in energydensity seen by an observer moving along a trajectory of the irrotational isometry which generates Σ (Edelen, (2)). The theory makes no assertions concerning the constancy or range of variability of ξ. The two extreme possibilities are that ξ may be an absolute constant, or that it may be different for each galaxy. In the former case, equation (1.1) may be readily tested and unequivocally; in the latter, any discretiation determined by the $\sqrt{n(n+1)}$ factor would be completely masked by the variations in ξ, and confirmation would probably be impossible.

In the absence of any a priori knowledge concerning ξ, we can only note that in addition to ξ being an absolute constant, two other possibilities render Eq. (1.1) subject to test. The first possiblity is that ξ may assume only a small number of distinct values, where by small is meant a number such that the sample of diameters which exhibit the $\sqrt{n(n+1)}$ discretization for each separate value of ξ is large enough that Eq. (1.1) may be considered independantly confirmed for each ξ with statistical confidence. The second possibility is the identification of in galaxies having the same ξ of some observable related to ξ which can be independently measured.

In view of these considerations, it appears that an observational test of Eq. (1.1) is feasible under certain conditions which can only be known a posteriori. That is to say, the feasibility of a test will be established along with confirmation of Eq. (1.1) in the event of positive results. Whereas in the case of negative results, it will not be possible to say whether the discretization hypothesis is wrong, the proportionality assumption is wrong, or that ξ varies in such a way as to conceal any discretization. Thus, the discretization hypothesis in its present formulation is ~~that~~ such that it lends itself to observational confirmation but not to observational refutation.

~~Several additional complications are involved in a test of the discretization hypothesis.~~ Equation (1.1) applies to EO galaxies and, to within ~~specifiable~~ deviations, to ellipticals of small true eccentricity. However, samples of elliptical galaxies to be tested must be selected on the basis of their apparent eccentricities. Hubble (Ap. J. Vol. 64, 1926, pp. 321-369) has shown that in any sample of elliptical galaxies with random orientation of axes of symmetry that only 55% of the apparent EO's are true EO's and the remainder are ellipticals with larger true eccentricities so oriented as to give an apparent ellipticity from 0 to 0.05 (corresponding to a range in eccentricity from 0 to 0.3). It is thus probable that any sample of low apparent eccentricity galaxies will be "contaminated" with galaxies of larger true eccentricity for which equation (1.1) does not hold. It is therefore necessary to investigate the probable degrees of contamination in our samples in order to determine whether or not discretization tests ~~will~~ will be statistically vitiated by this effect.

First, let us specify more precisely the degree of fit of Eq. (1.1) in terms of the <u>true</u> eccentricity. The deviations in diameter from the values given in Eq. (1.1) as functions of ϵ may be computed from the theoretical eigen functions for m = 0, (m = 0 ~~the values of m which~~ will be assumed to correspond to the elliptical galaxies). In Table Ia the percent increase in diameter over the $\epsilon = 0$ diameter of Eq. (1.1) is tabulated against the value of ϵ and the branch number n (as determined ~~derived~~ from the table of the eigen functions ~~given~~ in reference [2]). The deviation ~~derivation~~ is seen to be about 1% for all branches out to an eccentricity of 0.2; less than 7% out to $\epsilon = 0.5$, etc. Table Ib gives the estimated maximum ~~estimated~~ per cent deviation for large n ~~longer with~~ corresponding to various true eccentricities. Thus if the true eccentricity is always less than 0.42 the per cent deviation will be ~~#1~~ five per cent or less for all n.

It is thus seen that the testable sample for eq. (1.1) consists not just of true EO's but of all ellipticals whose true eccentricity is less than or equal to a value corresponding to some specified deviation, as given in Table I. Without here stating the value of the statistically acceptable precision~~s~~ for the test, the appropriate value of ~~it is evident that the~~ deviation selected ~~used~~ should be ~~taken~~ the same as the precision of measurement. For example, if the diameters are known to within 5% then ~~4%,~~ eq. (1.1) may be assumed to hold with the same percent deviation, i.e. to the same precision out to true eccentricities of 0.42 ~~0.38~~, etc.

With the test sample defined in terms of the true eccentricity and degree of precision, it is now required to define the sample in terms of the apparent eccentricity. The apparent and true eccentricities are related by ~~the equation~~

$$e_a = e_t \sin \varphi$$

where φ is the angle between the axis of the galaxy and the line of sight. Let us assume for illustrative purposes that the acceptable precision is 5%. Then any galaxy with an apparent eccentricity of less than 5 per cent will be admittable to the test sample. In order for e_a to be ≤ 0.05

$$\sin \varphi \text{ must be } \leq \frac{0.05}{e_t}$$

The values of $\sin \varphi$ and φ corresponding to different e_t's are given in the table below

Table II

~~Table IIa~~

Ranges of $\sin \varphi$ and φ for $e_a \leq 0.05$

e_t	$\sin \varphi$	φ
0.0	0 to 1.000	0 to 90°
0.1	0 to 0.500	0 to 30°
0.2	0 to 0.250	0 to 14.5°
0.3	0 to 0.167	0 to 9.6°
0.4	0 to 0.125	0 to 7.2°
0.42	0 to 0.119	0 to 6.8°
0.5	0 to 0.100	0 to 5.7°
0.6	0 to 0.083	0 to 4.8°
0.7	0 to 0.071	0 to 4.1°
0.8	0 to 0.063	0 to 3.6°
0.9	0 to 0.056	0 to 3.2°

The dotted band dividing the table into two parts corresponds to $e_t = 0.42$, which from Table Ib is the limiting value of e_t for a galaxy to be testable to within a precision of 5%. All galaxies above the dotted band are testable regardless of orientation, although those having $e_a \gtrsim 0.05$ will be rejected from the sample.

The galaxies below the dotted band will not be testable since their major axes deviate by more than 5% from the values of equation (1.1). Those below the dotted band whose ϕ's are greater than those given in the right column will have values of $e_a \gtrsim 0.05$ and hence will be detected and rejected from the sample. The remaining galaxies, those below the dotted line with ϕ's less than given in the right hand column will neither be testable nor detectable as interlopers and will constitute the contaminating portion. The contaminating set is thus a subset of the set of galaxies below the dotted band for which $\phi \lesssim \phi_o = 6°8$.

If we assume with Hubble that the distribution in e_t is uniform and that the axes are randomly oriented, then the probability of an axis falling in this contamination cone will be

$$1 - \cos \varphi_c .$$

Thus for a precision of 5%, the per centage of undetectable contaminating galaxies will be less than

$$1 - \cos 6.8^{\circ} = 0.7\%$$

Table III gives the per cent contamination for various precisions

Table III

% Precision	1	2	3	4	5	10
φ_c	3.0°	4.25°	5.75°	6°	6.8°	10°
Bound on %Contamination	0.1	0.3	0.5	0.6	0.7	1.5

The small per centage of contamination given by Table III, based on random orientation of axes, allow the discretization hypothesis to be restated in the form of a prediction that Eq. (1.1) is valid for apparent EO's and low apparent eccentricity samples of ellipticals with exceptions determinable from the specified precisions.

4) The fourth problem in the construction of an observational test of the discretization hypothesis, the conversion of observed angular diameters to linear diameters, involves the cosmological questions which arise when objects located at different distances whose observed radiation originated at different times are to be compared. Since it is desirable to hold the number of additional assumptions entering into the test hypothesis to a minimum, it will be best initially to test samples of galaxies for which cosmological effects are negligible. These samples will in general be either galaxies with small redshifts or galaxies located in the same volume of space, such as members of the same cluster. If then it develops that discretization effects do exist and are readily observed in these samples it is evident that we will have at our disposal a yardstick for calibration of linear sizes of galaxies and hence possess a tool for cosmological exploration when samples at different distances are compared.

First let us consider the samples which in internal comparison should be relatively free of cosmological effects. The first sample consists of the large bright presumably nearby galaxies. The problem of distribution of these galaxies in distance can be met by assuming that the redshift is essentially an indicator of distance and that the peculiar component of the redshift due to dynamical effects other than the general expansion is negligible. This evidently is not a bad assumption. The size of the peculiar component has been estimated by de Vaucouleurs (5) to have a maximum value of 60 km/sec. More recently Neyman and Scott (6) have estimated the mean maximum as 10 km/sec. If these bounds are correct then with redshifts in the range ____ to ____, errors from this source should be less than ____ per cent. On the other hand, Zwicky (7) from arguments based on maximum observed negative velocities estimates the peculiar component to be as large as ______ km/sec. If the estimates quoted from the first two papers are to be used as mean values, the Zwicky result must still be assumed to be a possible value and caution exercised in assuming the peculiar component to be negligible especially in cases of closely interacting objects where large dynamical velocities are clearly present.

Under these assumptions the conversion of angular diameters to linear diameters is straight forward. Assuming Hubble velocity-distance relation in the form

$$H\Delta = cz$$

where H is Hubble's constant, Δ is the distance, c the velocity of light and $z = \delta\lambda/\lambda$ is the spectral shift; the linear diameter S should be given to good approximation by

$$\log S = \log \theta + \log cz - \log H$$

where θ is the angular diameter. Any discretization properties of S may thus be investigated for an E0 galaxy for which θ and z are known.

However, Discretization, if it exists, should be measured in galaxies as they are at a given cosmic time and not as they appear to a particular observer. Hence, in order to eliminate preferences given to the observer's particular location in space, it will be necessary to correct the observed quantities for light travel time. The Δ, S, and θ of the above equations accordingly must represent ~~not the observed values~~, the distance, the linear, and the angular diameters reduced to a common epoch, and not the observed values.

On the basis of purely kinematical considerations an isochronous representation for angular diameters may be derived directly as follows: ~~Equation (1.2) is derived from purely kinematical considerations and does not include higher order redshift or evolutionary effects. It should, however, be applicable locally.~~ If a signal leaves a galaxy at time t' when the galaxy is at distance Δ', then the distance Δ, of the galaxy at time t_o, when the signal is observed will be.

$$\Delta = \Delta' t + cz\ (t_o - t) = \Delta'(1+z)$$

or in terms of the linear and angular diameters,

$$\frac{S}{\theta} = \frac{S'}{\theta'}\ (1+z)$$

where S' and θ' are respectively the <u>observed</u> linear and angular diameters corresponding to epoch t'; and S and θ are the linear diameters reduced to their values at epoch t_o.

If the linear diameter has remained unchanged during the light travel time $(t_o - t')$, then

$$\theta = \theta'(1+z)^{-1},$$

which is the kinematic correction to angular diameters applicable in all models. (Hoyle [8])
The above relation may then be rewritten in terms of observed angular diameters θ' as

$$\log S' = \log \theta' + \log z - \log (1+z) + \log (c/H)$$

or

(1.2) $$\log S' = \log \theta' - \log u + \log (c/H)$$

The quantity $u = 1+z/z$ will be designated the "synoptic redshift." Equation (1.2) is derived from local kinematic considerations and does not include higher order redshift or evolutionary effects. It should, however, be applicable locally.

~~However,~~ For more distance galaxies, various cosmological models introduce modifications involving different powers of (1+z) into the diameter distance relation. To be as inclusive as possible the tests for discretization among non local samples should therefore be based on generalizations of the form

(1.3) $$\log S = \text{constant} + \log \theta + \log z + \sigma \log (1+z)$$

metric $\sigma = -1$
isophotal $\sigma = +1$

where the parameter σ should be empirically determined and compared ~~further~~ with values predicted by various cosmological models.

The basic test equation of the discretization hypothesis may now be written by combining the logarithmic form of Eq. (1.1) with either Eq. (1.2) which is valid locally or with the more general form of Eq. (1.3) when a cosmological parameter σ must be derived. The local test equation:

$$\text{(1.4)} \qquad \log \theta' - \log u = \text{constant} + \tfrac{1}{2} \log n(n+1) - \tfrac{1}{2} \log \xi$$

where θ' is the observed angular diameter $u = (1+z)/z$, n is a positive integer and ξ is a physical parameter of unknown distribution as discussed in Section 2.

The general test equation:

$$\text{(1.5)} \qquad \log \theta' + \log z + \sigma \log(1+z) = \text{constant} + \tfrac{1}{2} \log n(n+1) - \tfrac{1}{2} \log \xi$$

where σ is to be determined.

¶ In the special case of the second type of test sample, namely galaxies belonging to the same cluster, the test equation takes a simple form.

~~One approach to this question is to select a test sample of apparent low eccentricity ellipticals, located in clusters~~. no ¶ ~~For~~ If the cluster ~~membership~~ membership of each member of the sample can be established by redshifts or some other suitable criterion, then the spread in distance among members of the sample should be small and linear diameters should be directly replaceable with angular diameters. An estimate of the size of error introduced by this approximation ~~of the size of error introduced by this approximation~~ may be made by assuming that the member galaxies of a cluster are distributed in depth along the line of sight by an amount equal to their linear distribution at right angles to the line of sight. If Δ is the distance to the cluster and Θ is its angular extension (measured in radians in the plane of the sky), then the relative distribution of depth, $\frac{\Theta \Delta}{\Delta} = \Theta$.

For example, if the Coma Cluster has a diameter of 2°, its member galaxies will have a 3% fluctuation in relative distance. Whenever the relative distances are of the order of or less than the acceptable error in the diameter measurements, as will be true for the Coma Cluster and beyond, angular diameters uncorrected for distance may be tested for discretization effects directly. Thus for purposes of tests in clusters the test equation may be written in the form

$$\log \theta' = \text{constant} - \tfrac{1}{2} \log \xi + \tfrac{1}{2} \log n(n+1) \qquad (1.6)$$

where θ' is the observed angular diameter.

REFERENCES

1. Edelen, D.G.B., Astron. J., 68, 535 (1963).

2. Edelen, D.G.B., Galactic Scale Discertization - I Theory, The RAND Corporation, Rm-3941-RC (in press).

3. Hubble, E. P., Distribution of Luminosity in Elliptical Nebulae, Ap. J., V. 71, 1930, p. 231-276.

4. Hubble, E. P., Extragalactic Nebulae, Ap. J., Vol. 64, 1926, pp. 331-369.

5. de Vaucouleurs, G. A.J., Vol. 63, 253, 1958.

6. Neyman and Scott A.J., Vol. 66, 148, 1961.

7. Zwicky, F.

8. Hoyle, F. Cosmological Tests of Gravitational Theories Proceedings of the International School of Physics, "Enrico Fermi" XX, 1961, pp. 141-173

III. FITS OF OBSERVED ANGULAR DIAMETERS TO THEORETICAL DISCRETIZATION SEQUENCES

The difficulties in obtaining a uniform set of angular diameters of elliptical galaxies have been discussed by several authors, Hubble (   ), de Vaucouleur (   ). These difficulties in defining and measuring diameters have been such that it has not been possible to obtain sets of measurements of sufficient accuracy and consistency to be useable for the law of redshifts or other cosmological investigations. Similarly, most published diameters are also unsuitable for discretization tests. For example, the Shapley Ames Catalogue (   ) which would provide a large test sample of early ellipticals, turns out to be unsuitable for discretization tests because of the large per cent of diameter measurements which suffer round-off effects.

As discussed in the first part of Section II, several sets of similarly defined diameters (such as diameters defined by different isophotes intensity levels) should be tested in order to establish the existence of a proportionality range. It might seem that different sets of published diameters which include the same galaxies might be substituted for this purpose. But unless the operational definitions used are identical this is not necessarily so. It will, non-the-less be important to test data which may be too limited for proportionality tests. This is because, while positive results may not be regarded as definitively establishing discretization, negative results demand rejection of the test as formulated in Section II.

Accordingly in this section the angular diameter measurements from three published sets of data are compared with the predicted discretization sequences and a measure of the degree of fit between the observations and theory established for each sample. The statistical significance of the fits will be discussed in Section IV.

The first test sample consists of the bright EO galaxies from A. and G. de Vaucouleurs' Reference Catalogue of Bright Galaxies (Univ. of Texas Press, 1964). The data from the de Vaucouleurs Catalogue consists of all galaxies classified by the authors as EO for which redshifts and diameters are available. The first column in Table III gives the NGC numbers with the parentheses indicating galaxies whose classification as EO's may be uncertain but whose major and minor axes are given as equal. The second column gives the common logarithm of the apparent major axis or diameter in tenths of a minute of arc. The third column gives the velocity corrected for galactic rotation. The fourth and fifth columns are respectively the common logarithms of the velocities and the synoptic velocities, $u = (c + V_c)/V_c$. The sixth column gives the B-V colors reduced to the diameter D and corrected for galactic absorption and redshift. The seventh column gives the spectral class given by Humason, Mayall, and Sandage, (A. J.......) and the last column, again from the Reference Catalogue, gives the blue magnitude to diameter D. Details of the magnitude and color systems and the corrections are given in the Catalogue.

The angular diameters are derived statistically from published values of micrometric diameter by several observers, The Reference Catalogue being based on 30 different sources. (See de Vaucouleurs, A.J....). The redshifts are either Humason-Mayall-Sandage values or recent determinations by A. and G. de Vaucouleurs.

The tabular values are carried to a larger number of places than the precision in the individual measurements warrants. This is done throughout in order to minimize round-off effects (a discretization must not be that established by rounding off).

The errors in log D are not known, but if they may be inferred from the internal errors of the various measurements used in arriving at log D's,

TABLE III

EO BRIGHT GALAXIES

NGC	log D	V_c	log V_c	log u	Color	Spectra	m_{pg} Magnitude
83	0.94	6745	3.8290	1.658	0.91	G3	14.21
382	0.81	5354	3.7287	1.756		G5	
596	1.28	2097	3.3216	2.158	0.78	G3	12.31
741	1.15	5636	3.7510	1.734	0.90	G5	13.12
(751)	0.88	5291	3.7235	1.761	0.75	G2	14.11
1374	1.01	1136	3.0554	2.423			
1379	1.05	1303	3.1149	2.364			
1399	1.26	1311	3.1176	2.363	0.73	G4	11.19
1407	1.16	1707	3.2322	2.246	0.95	G3	11.43
1889	0.65	2334	3.3681	2.112	0.77	G2	14.29
2673	0.89	3669	3.5646	1.918	[0.65]	G3	[14.38]
(2694)	0.90	5165	3.7131	1.771	0.89	G4	15.27
3348	1.02	3010	3.4786	2.003		G5	12.45
4283	0.94	1078	3.0326	2.446	0.86	G8	13.42
4339	1.15	1173	3.0693	2.410	0.83	G3	12.83
4458	1.05	311	2.4928	2.985	0.80	G7	13.32
4552	1.30	195	2.2900	3.187	0.91	G7	11.30
4636	1.48	629	2.7987	2.679	0.85	G2*	11.01
4782	0.96	3858	3.5864	1.896			
4783	0.97	4527	3.6558	1.828			
(4827)	1.04	7657	3.8841	1.604	0.81		14.35
4915	0.96	3034	3.4820	1.999	0.79	G5	13.30
(4926)	1.00	7673	3.8850	1.603	0.96		14.40
5061	1.18	1888	3.2760	2.204			
5173	1.00	2508	3.3993	2.081		G4*	14.12
5812	0.96	2040	3.3096	2.170	0.88	G7	12.90
5846	1.32	1784	3.2514	2.228	0.89	G0*	11.76
5898	0.88	2231	3.3485	2.132	0.74	G2	12.91
5930	1.09	2868	3.4576	2.024			
(5953)	0.99	2188	3.3401	2.140			
7507	1.05	1684	3.2263	2.253	0.81	G5	11.87

( ) means uncertainty in type

D = major axis in [illegible]

V_c = velocity corrected for galactic rotation

Color = NC(0) = B-V for D = D(0) [illegible]

Spectra from HMS

Magnitude ~ B(0) blue [illegible] diameter D(0)

D(0) = log D, [illegible]

a value $\leq \pm$ 0.01 (i.e. 3%) may be taken. This error will be assumed throughout the range of log D for the galaxies reported in Table III. The error in the redshift is more or less independent of the size of the displacement resulting in a large relative error for nearby objects and a small relative error for distance objects. A constant error of $\pm$ 70 km/sec, (which is a mean value for nearby galaxies reported in the HMS Catalogue) is assumed for $\delta . V_C$. The errors in log V_C and log u range from almost 0.2 for the nearest galaxies in Table III to about 0.004 for the most distant. The redshift and diameter errors are equal at about 3000 km/sec.

The functions $A = \log D + \log V_C$ and $B = 2 + \log D - \log u$ are tabulated in Table IV together with errors $\delta . = \delta A = \delta B$ estimated from the considerations of the previous paragraph. The δ's are not ~~intended more for relative weights of the A's and B's than~~ to be taken as a quantified ~~direct~~ measure of the internal error but are measure relative uncertainties. 3 a selectivity effect 3. δ is correlated with n

According to Eq. (1.4), there should exist a positive integer n, such that for ~~each~~ each B_n from Table IV

$$B_n - \tfrac{1}{2} \log n(n+1) = \text{constant} - \tfrac{1}{2} \log \xi$$

to within some specified deviation (such as the δ's of column 4). Further, in order for the test to be interpretable, ξ must assume only a limited number of values.

In order to determine whether such integers exist, the unknown constants may be eliminated by ~~example~~ comparing the differences between B's with quantities of the form $\tfrac{1}{2} \log i(i+1) - \tfrac{1}{2} \log i(j+1)$ with $i \geq j$. To do this a "<u>B</u>" <u>difference table</u> is compared with a <u>$\tfrac{1}{2}$ log n(n+1) difference table</u>. The important ~~maximum~~ differences are those between the lowest values of B which presumably should correspond to the differences between the lowest values of $\tfrac{1}{2} \log n(n+1)$. Fits for larger values of n have little significance unless they are part of a sequence which also fits to the lowest values of n. ~~Table Va is a difference table constructed from the smallest values of B. The values of B~~

TABLE IV

NGC	A	B	δ
4458	3.543	0.065	0.08
4552	3.590	0.113	0.1
4283	3.973	0.494	0.03
1889	4.018	0.538	0.02
1374	4.065	0.587	0.03
1379	4.165	0.686	0.02
4339	4.219	0.740	0.03
5898	4.229	0.748	0.02
5812	4.270	0.790	0.02
7507	4.276	0.797	0.02
4636	4.279	0.801	0.04
5953	4.330	0.850	0.02
1399	4.378	0.897	0.02
1407	4.392	0.914	0.02
5173	4.399	0.919	0.02
4915	4.442	0.961	0.01
2673	4.455	0.972	0.01
5061	4.456	0.976	0.02
3348	4.499	1.017	0.01
382	4.539	1.054	0.01
4782	4.546	1.064	0.01
5930	4.548	1.066	0.01
5846	4.571	1.092	0.02
751	4.604	1.119	0.01
596	4.602	1.122	0.02
2694	4.613	1.129	0.01
4783	4.626	1.142	0.01
83	4.769	1.282	0.01
4926	4.885	1.397	0.01
741	4.901	1.416	0.01
4827	4.924	1.436	0.01

$A = \log D + \log V_c$

$B = 2 + \log D - \log u$

TABLE V

[illegible]	B	n	P_n	$B-P_n$	δ	m	Q_m	$B-Q_m$	δ
[illegible]	[illegible]					1	0.081	[illegible]	0.08
4552	113					1	0.081	+0.032	0.10
4283	494					3	0.470	+0.024	0.03
1889	53	2	0.538	0.000	0.04				
1374	[illegible]					4	0.581	-0.006	0.03
1379	0.586	3	0.689	-0.003	0.02				
[illegible]	[illegible]					6	0.742	-0.002	0.03
	0.748					6	0.742	+0.006	0.02
5812	0.790	4	0.800	-0.010	0.02				
7507	0.797	4	0.800	-0.003	0.02	(7)	0.804	-0.007	0.02
4636	0.801	4	0.800	+0.001	0.04	(7)	0.804	-0.003	0.04
5953	0.850					8	0.859	-0.009	0.02
[illegible]	[illegible]	5	0.888	+0.009	0.02	(9)	0.908	-0.011	0.02
1407	0.914					9	0.908	+0.006	0.02
5173	0.919					9	0.908	+0.011	0.02
4015	0.961	6	0.961	0.000	0.01				
2673	0.972	6	0.961	0.011	0.01				
5061	0.976	(6)	0.961	0.015	0.02	11	0.991	-0.015	0.02
3348	1.017	7	1.023	-0.006	0.01	(12)	1.024	-0.007	0.01
382	1.054					13	1.060	-0.006	0.01
4782	1.064					13	1.060	+0.004	0.01
5930	1.066					13	1.060	+0.006	0.01
5846	1.092					14	1.091	+0.001	0.02
751	1.119	(9)	1.126	-0.007	0.01	15	1.120	-0.001	0.01
596	1.122	9	1.126	+0.003	0.02				
2694	1.129	9	1.126	+0.003	0.01				
4783	1.142					16	1.147	-0.005	0.01
83	1.282	13	1.279	+0.003	0.01				
4926	1.397	17	1.392	+0.005	0.01				
741	1.416	18	1.416	0.000	0.01				
4827	1.436	19	1.439	-0.003	0.01				

$$B = 2 + \log D - \log u$$

$$P_n = 0.149 + 1/2 \log n(n+1)$$

$$Q_m = -0.070 + 1/2 \log m(m+1)$$