Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 4798–4805, June 1993
Colloquium Paper

This paper was presented at a colloquium entitled "Physical Cosmology," organized by a committee chaired by David N. Schramm, held March 27 and 28, 1992, at the National Academy of Sciences, Irvine, CA.

The redshift–distance relation

I. E. SEGAL

Department of Mathematics 2-244, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

ABSTRACT **Key predictions of the Hubble law are inconsistent with direct observations on equitable complete samples of extragalactic sources in the optical, infrared, and x-ray wave bands—e.g., the predicted dispersion in apparent magnitude is persistently greatly *in excess* of its observed value, precluding an explanation via hypothetical perturbations or irregularities. In contrast, the predictions of the Lundmark (homogeneous quadratic) law are consistent with the observations. The Lundmark law moreover predicts the deviations between Hubble law predictions and observation with statistical consistency, while the Hubble law provides no explanation for the close fit of the Lundmark law. The flux–redshift law $F \propto (1 + z)/z$ appears consistent with observations on equitable complete samples in the entire observed redshift range, when due account is taken of flux limits by an optimal statistical method. Under the theoretical assumption that space is a fixed sphere, as in the Einstein universe, this law implies the redshift–distance relation $z = \tan^2(r/2R)$, where R is the radius of the spherical space. This relation coincides with the prediction of chronometric cosmology, which estimates R as 160 ± 40 Mpc (1 parsec = 3.09×10^{16} m) from the proper motion to redshift relation of superluminal sources. Tangential aspects, including statistical methodology, fundamental physical theory, bright cluster galaxy samples, and proposed luminosity evolution, are briefly considered.**

The good news is that there is a simple redshift relation that appears consistent with observations in complete objectively defined samples in the infrared and x-ray wave bands as well as the optical. The bad news is that it doesn't at all resemble the Hubble law, which appears simply irreconcilable with these observations.

These facts come out in a systematic audit of redshift observations and theory, of the sort that every putative scientific theory should have periodically, as emphasized in the National Academy of Sciences booklet "Science and Creationism" (1). Outside of logic and mathematics, our basis for separating fact from fancy has to be probability, whose elementary laws are beyond dispute. A theory whose predictions are improbably deviant from direct observation is scientifically incorrect, while one whose predictions agree with direct observation within apparent statistical fluctuations is scientifically tenable and may be correct, although never in an ultimate sense. A "theory" that escapes deviations from direct observation by not making predictions just isn't a scientific theory, or in Pauli's terms, "isn't even good enough to be wrong."

Thus the burden of our audit has to be the objective, reproducible, maximally accurate determination of probability levels for the deviations of theoretical prediction from observation, in documented fair samples. Of course, the quintessential problem of observational astronomy, the inherent cutoff on flux, produces a kind of inherent bias, so that "fair" here doesn't mean that the sources are from the same population at all redshifts but that the criterion for inclusion in the sample is an objective and theory-independent one.

In this era of full disclosure, I would be remiss not to mention that I am the proposer of a non-Doppler theory of the redshift, called chronometric cosmology (CC), that predicts a different redshift–distance relation from those of Friedman–Lemaitre cosmology (FLC). This may affect our visceral responses, but we have absolutely no interest in being purveyors of an ultimately false theory of the nature of the redshift, to speak also for my principal collaborator, Jeffrey Nicoll, as well as a number of younger colleagues. We have gone to considerable lengths to devise and implement procedures that are of maximal statistical efficiency and physical conservatism; to document these procedures so that our results are fully reproducible from the observations in the designated sample; to base our analyses on well-documented and thoroughly studied flux-limited samples (in whose observation we had no hand, and which are used exactly as published, without any type of model-dependent correction); and to stick to the same basic procedure in all the analyses, in all redshift ranges and for all types of sources, to facilitate fair comparisons.

For conservatism on the physical side, we make no assumptions regarding the spatial distribution of the sources. These are uncertain, often questioned, and unnecessary for the estimation of the luminosity function (LF), when normalized to a total of unity, as suffices for cosmological testing purposes. For conservatism on the observational side we make no assumption as to the completeness of the sample in redshift (e.g., there may be gaps in the redshift distribution). All that is assumed is that at each redshift that is observed, *there is no selection on the basis of flux, down to the designated limit of the sample*. The conservatism of this assumption is further increased by analysis of successively brighter subsamples. Indeed, the flux limits may vary with the object, as when different parts of the sky are observed down to different limits.

Conservatism on the probabilistic side is attained by being nonparametric, computer-intensive, and using maximum-likelihood theory, all at the same time. At first glance, it may appear paradoxical to be both nonparametric and maximum-likelihood, because the latter theory has the goal of estimating a finite set of parameters, in terms of which the unknown

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: LF, luminosity function; CC, chronometric cosmology; FLC, Friedman–Lemaitre cosmology; LE, luminosity evolution; IRAS, infrared astronomical satellite; SIRAS, galaxy sample treated in ref. 14; CMB, cosmic microwave background; EU, Einstein universe; SR, special relativity; MP, Mach's principle; EEP, Einstein equivalence principle; BQS, quasar sample treated in ref. 7; SLF, sample LF estimated by ROBUST; QDOT, galaxy sample treated in refs. 17 and 18; EMSS, x-ray source sample treated in refs. 19, 20, and 42; AGN, active galactic nucleus; GR, General Relativity.

distribution function (here the LF, taken throughout in normalized form) is expressed. But, following a suggestion from Michael Woodroofe, we simply divided the absolute magnitude (or log flux) range subject to "Malmquist bias" into so many bins and used the values of the differential LF, assumed constant in each bin, as the parameters. *A priori* this appears as a rather horrible nonlinear problem, involving extensive number crunching with no assurance of convergence. Fortunately, we found a very simple closed-form solution, expressible simply as a rational function of the occupation numbers of the bins. Indeed, the solution was so natural that we guessed it to begin with and then determined, following a suggestion from Herman Chernoff, that it was indeed maximum-likelihood (2–3). We called our method ROBUST because it made no *a priori* assumption as to a parametric form for the LF.

Donald Lynden-Bell then called his heuristic infinite-number-of-parameters maximum-likelihood analysis (4) to our attention. While this work involved the assumption of spatial uniformity of the source distribution, as ours did not, the resulting LF estimate was nevertheless the limit of the ROBUST estimate as the bin size tends to zero.

Maximum-likelihood theory, which for practical purposes goes back to Fisher (1922), requires a finite number of parameters and in these terms defines the notion of efficiency (equivalent to that of maximal information content, roughly) and of sufficiency (extraction of *all* the information in the sample relevant to the parameters being estimated). ROBUST goes beyond generic maximum-likelihood estimation in providing such sufficient statistics, and its validity as actually programed has been tested by a vast number of simulations. This is exemplified concretely by the effectiveness of ROBUST in removing the flux cutoff bias for the Hubble law, or for FLC more generally; correlations between absolute magnitude and log$z$ are greatly reduced in absolute value, and the dispersion in the m given z relation, $m|z$ (i.e., the "conditional distribution of m for given z") increased only slightly and in fact reduced when the cutoff is especially severe, by straightforward corrections that derive from the ROBUST estimate of the LF.

Much of the recent astronomical literature uses a method basically due to Schmidt (5) for estimation of LFs. This method assumes spatial homogeneity or otherwise constrains the spatial distribution of the sources and thus is physically considerably less conservative than ROBUST. In addition, it requires completeness in redshift over the range in question and is thus observationally considerably more demanding. Moreover, the Schmidt estimate seems rarely to have been validated in concrete cases by direct *a posteriori* testing.

Perhaps the simplest and most direct test of an estimated LF is to draw random sources from it and place them at the observed redshifts, then reject them if the corresponding flux is below the sample limit, and continue until all observed redshifts are assigned an object from the estimated LF whose corresponding flux at the observed redshift is as bright as the sample limit. The resulting "random sample" should agree within statistical fluctuations with the observed sample, if the underlying cosmology is correct and the sample is fair, modulo the inherent flux cutoff or other objective constraints defining the sample. In particular, the application of the underlying LF estimation to the random sample (or the average LF obtained in a succession of random samples) should result in an LF that is statistically consistent with the original LF estimated from the directly observed sample data. This has apparently not been done with the Schmidt method, perhaps because it might seem (incorrectly) that the statistical agreement would be automatic, as would, for example, be the case when there is only one redshift, or no cutoff on flux. But this "Monte Carlo" type procedure has been performed many times for ROBUST estimates, both for observed and for artificial samples. On the basis of the latter trials, there is no question that the application of ROBUST to an artificial flux-limited sample constructed assuming any given cosmology and LF will reproduce that LF within statistical fluctuations, irrespective of the assumed given redshifts.

The results of parallel analyses of the low redshift to distance laws (e.g., for $z < 0.1$) predicted by FLC or CC (briefly, the Hubble law, or C1, and the Lundmark law, or C2, where Cp denotes the redshift–distance law of exponent p), on the bases of large, clearly equitable, and objectively defined flux-limited samples, have been qualitatively similar. (Historical note: the Lundmark law is named primarily in recognition of his methodological priority; the empirical law he published in 1925 was quadratic but distinct from C2.) The C1 predicted mean apparent magnitude at given redshifts is too bright on average, and quite conspicuously so at the lowest redshifts, which would be the least affected by assumed luminosity evolution (LE). Relatedly, the slope of the predicted magnitude–redshift relation consistently exceeds that observed. The C2 predicted mean apparent magnitudes are considerably closer to the observed values, and the deviations show no significant trend with redshift.

The correlation of absolute magnitude with (log) redshift is a highly relevant quantity because it speaks to the origin of the notion that redshift may be a function of distance, which arises from the observed correlation of apparent magnitude with redshift. One of the most crucial functions of a cosmology is to explain this correlation. In nonevolutionary cosmologies it is supposed to do so in terms of a redshift-independent population of sources, whose absolute magnitudes are uncorrelated with the redshifts in the population at large and only appear to be correlated in flux-limited samples because of the cutoff bias. But is the cutoff sufficient to explain the extremely strong negative correlations of C1 absolute magnitudes with redshifts in typical flux-limited samples?

The answer is definitely no: C1 almost invariably predicts that the observed correlation will be greater (in absolute value, less) than it is actually observed to be. One might hope to explain this by progressive incompleteness at lower flux levels, but this would suggest that subsamples with brighter flux limits should show improved agreement between the C1 predictions and the observed values, of the correlation of absolute magnitude with redshift, and this is not the case. In contrast, the C2 predictions for these statistics are in very good agreement with their directly observed values. But not only that, the C2 predictions for what the C1 correlation of absolute magnitude with redshift will be observed to be are extremely accurate—notwithstanding the inability of C1 itself to predict accurately its own statistic!

These results are indicative of serious flaws in C1, but one might hope that the deviations of its predictions could be explained by a conjunction of physical perturbations, local irregularities, and ancillary hypotheses. However, the deviations of the C1 predictions of the dispersion s_m in apparent magnitude from its directly observed values appear inherently irreconcilable. Large and otherwise appropriate samples will be discussed below in which the C1 predictions are of the order of 25% or more greater than the observed values; and this type of discrepancy between theory and prediction would tend to be enlarged rather than ameliorated by hypothetical perturbations and the like. In contrast, the C2 predictions for s_m are close to the observed values, typically within 2% (e.g., Figs. 1 and 4).

SOME EXAMPLES: OPTICAL SAMPLES

Before detailing these results for recent large complete samples that appear to be state-of-the-art, we exemplify the

4

statistical methodology just described by earlier analyses. Comparative tests of two or more cosmologies, in which each predicts the results of analyses predicated on the others, as well as on itself, may be made by testing the random samples constructed assuming the given cosmology in accordance with the precepts of the alternative cosmologies. Figures 2 and 3 of ref. 6 compare the FLC and CC LFs in this way, on the basis of the complete quasar sample of Schmidt and Green (ref. 7; hereafter, BQS). As not unexpected, from the general view that FLC is inconsistent with quasar observations, in its original nonevolutionary form, FLC cannot consistently predict its own LF, let alone that of CC (figure 3). However, CC appears quite self-consistent and predicts that analysis of the sample predicated on (nonevolutionary) FLC will yield, within apparent statistical fluctuations, just the ROBUST LF actually determined directly from the observations. This speaks to the question of whether the LE often postulated to reconcile FLC with quasar observations is physically real or merely exculpatory.

A sample complete to a given flux limit is of course complete to any brighter limit, and this provides a different type of *a posteriori* check on an estimated LF. The LFs derived from successively brighter subsamples should agree within statistical fluctuations, in their common absolute magnitude range, if the sample is fair and the cosmology is correct. As the apparent magnitude limit of the very low redshift galaxy sample of Visvanathan (9) is brightened from 12.4 to 12 to 11.6, figure 3 of ref. 8 shows the evolution of the C1 LFs and the coherence of those for C2.

The LF determines the observational implications of any given flux limit and, in particular, the mean apparent magnitude at the given redshift z. The observed $\langle m|z \rangle$ relation, when the redshift range is divided into a given number of bins, each containing the same number of objects (deleting objects at the large redshift end if necessary), is thereby predictable. The results for the sample (9) and the cosmologies Cp with $p = 1, 2, 3$ are shown in figure 3 of ref. 10, using 10 bins. The excessive brightness of the C1 prediction for the lowest redshift bin is conspicuous. A statistical summary of the prediction errors in the $\langle m|z \rangle$ relation in this and three other complete low-redshift galaxy samples is given in figure 1 of ref. 11. On the whole, C1 fits more poorly than Cp for any value of p in the range from $p = 1$ to $p = 3$, and C2 appears optimal; its prediction errors are of the order of half of those of C1.

COMPLETE SAMPLES IN THE INFRARED AND X-RAY BANDS

All of the foregoing complete samples are optical, for which there are relatively large absorption and aperture effects. At the low redshifts of complete galaxy samples, these appear however to be inconsequential as regards comparative cosmological testing, on the basis of studies using any of various estimates of absorption, or restriction to polar regions, and the isophotal magnitudes or diameters involved in the above cited work. The only apparent hope for the validation of C1 lies in the possibility that the deviations of its predictions from observation reflect a subtle combination of local perturbations that are absent at larger redshifts or in less ambiguous wave bands.

IRAS (infrared astronomical satellite) galaxy samples provide flux-complete subsamples for which redshifts have been increasingly observed. The sample of Soifer *et al.* (12) in the redshift ranges between 500 and 5000 km/s, inclusive of 202 galaxies, has values of s_m and of the correlations ρ_p of absolute magnitude (where we convert infrared fluxes using a Pogson scale to facilitate comparison with optical observations) with log redshift that impugn C1 and are consistent with C2. For example, in 100 random samples the predicted values of s_m all exceed the observed value, when C1 is assumed as the basis for the construction of the samples, but the predicted values are consistent with the observed value when the construction of the random samples is based on the assumption of C2. This provides some independent confirmation for the results of studies of complete optical samples, but it is possible that a significant motion of the Galaxy materially affects the results and conceivable that it effectively mimics the configuration predicted by C2. Statistically precise studies (13) have earlier shown that the comparative fits of C1 and C2 are insensitive to assumed motions for the whole sky sample (8), but this may not apply to a sample of limited sky coverage such as ref. 12.

Fortunately, the deeper sample of Strauss *et al.* (14) provides a basis for avoiding sensitivity to the motion of the Galaxy. We have made parallel statistical analyses of this sample testing C1 and C2, on the same basis as our earlier analyses, in the redshift range $500 < cz \leq 30{,}000$ km/s that is usually considered pre-evolutionary, inclusive of 2551 galaxies. Denoting the ROBUST LF determined from a given sample on the assumption of Cp as SLFp, the naive (uncorrected) sample statistics such as the correlation ρ_p of absolute magnitude with log redshift and the standard deviation σ_p of the residuals in the magnitude–redshift relation may be corrected for the apparent magnitude cutoff straightforwardly, as detailed in the references for optical samples, from the SLFp. In the sample of Strauss *et al.* (hereafter SIRAS), such correction increases the naive ρ_1 from the value -0.91 to 0.02 and <u>reduces</u> the naive σ_1 from 1.73 to 1.56 mag, thus confirming the effectiveness of ROBUST. In the case of C2, a naive ρ_2 of -0.67 is increased to one of 0.05, and the naive σ_2 is increased from 0.90 to 1.08 mag. (The decrease in σ_1 and increase in σ_2 reflect the comparative strength of the flux cutoff as viewed from the respective perspectives of C1 and C2.) On this basis, there is not much to choose between C1 and C2, apart from the lower dispersions for C2, which are not conclusive. However, the C1 prediction errors for $\langle m|z \rangle$ are of the order of twice those for C2, and the C1 prediction is again conspicuously bright at the lowest redshifts. There is also a pronounced trend with redshift in the C1 prediction errors, corresponding to an excessive slope for the regression of the predicted $\langle m|z \rangle$ on log$z$, and no such trend for C2.

In order to establish an overall probabilistic significance level for the deviations of the Cp predictions from observation, it is necessary to focus on individual collective statistics, of which the dispersion s_m in apparent magnitude appears as one of the cosmologically most sensitive. The Cp predictions for s_m in 100 random samples constructed assuming each cosmology were computed. The directly observed value of 0.77 mag is invariably exceeded by the C1 predictions, which have an average value of 0.98 ± 0.019, an excess of 10σ, while the C2 prediction is 0.77 ± 0.013. In view of the origin of the concept of functional dependence of redshift on distance in the correlation of apparent magnitude with redshift, the presumptively correspondingly vanishing (apart from the bias produced by the flux cutoff) correlation of absolute magnitude with redshift is a crucial statistic. Can the large absolute values of these correlations actually be explained by the flux limit? In fact, the C1 predictions for ρ_1 in the random samples are invariably too large, having an average value of -0.85 ± 0.0066, a deviation from observation of almost 10σ. The average C2 prediction is for the directly observed value, with a standard deviation of 0.0034. As earlier noted, the respective cross predictions of C1 and C2 are relevant: the C1 prediction for ρ_2 is deviant by 0.15 ± 0.016, while the C2 prediction for ρ_1 is the observed value $\pm$ 0.012. The slope of the linear regression of apparent magnitude on log$z$ has been emphasized traditionally. The random sample fluctuations in the predictions for this statistic are larger than those for s_m and ρ, but the C1 prediction is again deviant, by about 5σ, while the C2 prediction is accurate. For a detailed analysis of this sample, see ref. 15.

In the foregoing analysis, the observations were used as given, without correction for possible motion of the Galaxy (or Local Group). To check for sensitivity to such motion, the same analysis was made with redshifts corrected by two possible motions. There is no direct evidence that either motion is applicable to the SIRAS sample, but they serve to indicate the insensitivity of the foregoing results to conceivable motions. One motion (called CMB) is associated with the cosmic microwave background and is assumed to be 622 km/s towards galactic $b = 30$, $l = 277$ (16). The other (called OPT) was estimated (13) from a whole-sky optical sample to be a best-fitting motion assuming C1. The application of these motions changes the prediction errors in the statistics just described by at most 0.02 for s_m and ρ and has only a marginal effect on significance levels—e.g., with the CMB motion the deviation in the C1 prediction for s_m is reduced from 10σ to 9.3σ. The conspicuously excessively bright prediction of C1 for $\langle m|z\rangle$ in the lowest redshift bin is unaffected, as are the other qualitative features of $\langle m|z\rangle$. This is significant in relation to proposals for LE as an ancillary perturbation to C1; if linear to first order in z on the magnitude scale, as appears natural in the context of C1, the LE can correct the excessive brightness of predictions at low redshifts only if extremely rapid and then disturbs the fit at higher redshifts.

As earlier, to test for possible incompleteness in flux, the subsamples brighter by 0.5 and 1 mag than the reported sample limit were also tested in the same way, and with the same possible motions. Since the sample sizes go down by approximate factors of 2 for each 0.5 mag brightening, the significance levels of the C1 deviations are reduced, but never go below 3σ, and are generally considerably higher, at the same time that the qualitative features of the C1 predictions and their comparison with the C2 predictions are quite unchanged.

Since it has been fairly widely assumed that the Hubble law is applicable to the low-redshift regime (and the average redshift for the SIRAS sample is <4500 km/s), the same analysis that was just reported was applied to the deeper IRAS QDOT sample of Rowan-Robinson *et al.* (17). This consists of a random ("sparse") subsample of the IRAS sample down to its 60-micron flux limit and includes more than 2000 galaxies. With this flux limit and in the redshift range $500 < cz \leq 30{,}000$, comprising 1991 galaxies, the overall deviations of the C1 predictions from observation and the comparative fit of the C2 predictions are quite similar to those for the SIRAS sample. Fig. 1 shows the Cp predictions for s_m in 100 random samples constructed assuming each cosmology. Fig. 2 shows the same for the predictions of the correlations of absolute magnitude with log$z$, including cross correlations. Fig. 3 shows the same for the slope of the linear regression of apparent magnitude on log$z$. The results are quite similar to those for the SIRAS sample.

Saunders *et al.* (18) have argued that evolution is required to fit the Hubble law to the QDOT sample, but it is not clear what falsifiable (and hence scientific) content would remain in the Hubble law if there is no specification of the redshift range in which it applies in nonevolutionary form and/or of nontrivial constraints on the form of the evolution. Obviously, LE equal to the simple difference between the C2 and C1 regression functions for apparent magnitude on redshift for a fixed luminosity class will reconcile C1 to the observed $m|z$ relation, but to postulate such LE, or an effective approximation thereto, would appear scientifically redundant and render the present epoch special, considerably diminishing the theoretical attractiveness of the Doppler explanation for the redshift.

However, the SIRAS and QDOT samples are at fairly low average redshifts, and one might entertain the hope that they are affected by an unprecedented type of irregularity that is absent at higher redshifts and spuriously renders their predictions for the observed dispersion in apparent magnitude highly excessive. Fortunately, the EMSS-AGN sample of Gioia *et al.* (19) and of Stocke *et al.* (42) provides an observational basis for dealing with this possibility and for further comparative testing of C1 and C2. This sample is complete in x-ray flux and extends over a large redshift range; at the same time its wave band minimizes absorption and aperture ambiguities that may afflict optical samples at larger redshifts. At the limiting sensitivity for AGNs given in ref. 19, the sample includes 128 AGN up to $z = 0.15$, 224 to $z = 0.3$, and 277 to $z = 0.5$. The EMSS sample as a whole is a composite of subsamples with varying limiting fluxes, as in the case of the BQS, but the limiting fluxes applicable to the individual objects have not been available in this case. We tentatively adopt the reported limiting sensitivity as an overall effective, if approximate, complete flux limit for the AGN at the lower redshifts, subject to *a posteriori* checks, such as the analysis of successively brighter subsamples, of increasing completeness.

The results of parallel analyses of C1 and C2 in all redshift ranges up to $z = 0.3$, beyond which the FLC and CC predictions begin to depart significantly from those of simple power laws, are similar. For brevity, we give primarily the results for the sample up to $z = 0.3$, in which range C1 has traditionally been established from observations on bright

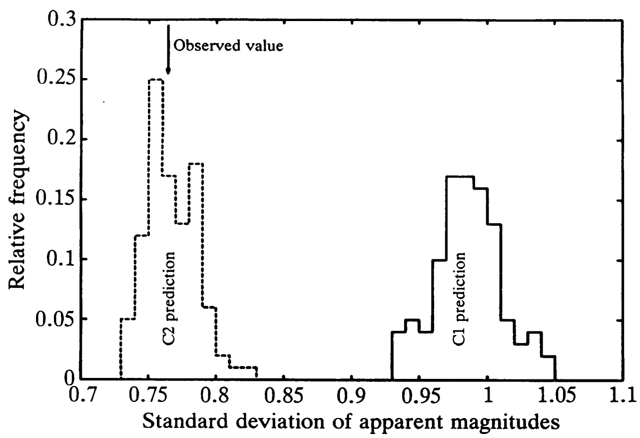

FIG. 1. Distribution of the Cp predictions of the dispersion in apparent magnitude, assuming Cp ($p = 1, 2$), for the QDOT sample in the range $500 \leq cz \leq 30{,}000$ km·s^{-1} (1991 galaxies), without any assumed motion.

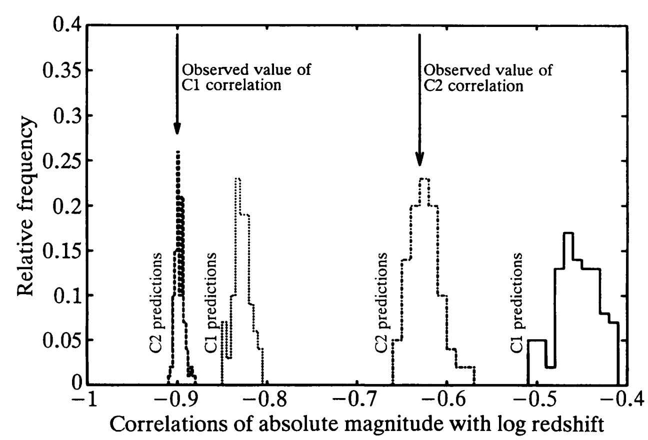

FIG. 2. The same as Fig. 1 for the correlations of Cp absolute magnitude with log redshift, as predicted by Cp, and also as predicted by C($3 - p$).

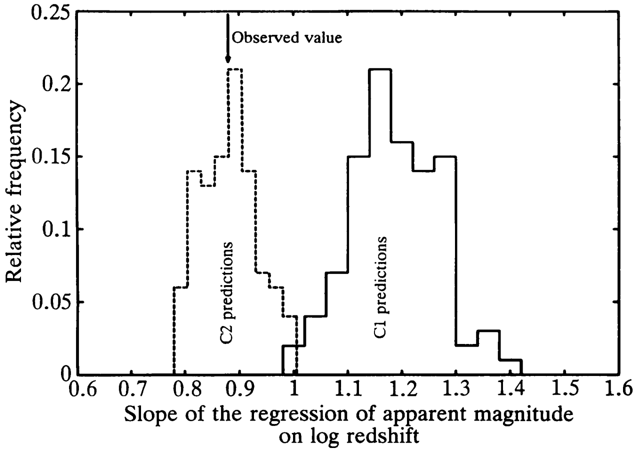

FIG. 3. The same as Fig. 1 for the regression of apparent magnitude on log redshift.

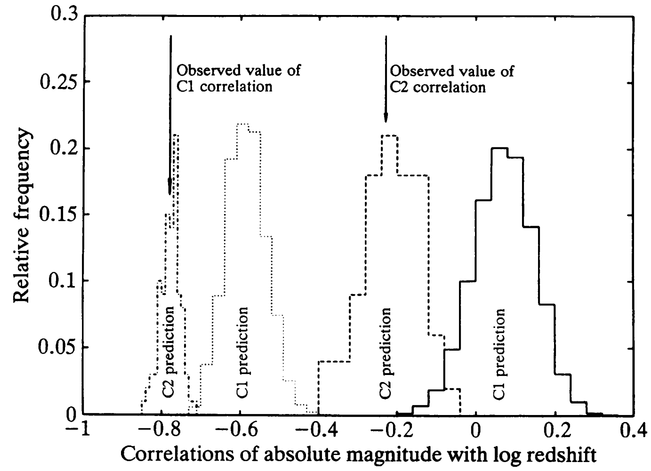

FIG. 5. The same as Fig. 2 for the sample of Fig. 4.

cluster galaxies (discussed later). In all analyses, as earlier, the absolute magnitude range subject to cutoff varying with redshift (Malmquist bias) has been divided into 10 bins and the LF assumed constant within bins of the resulting size. Figs. 4 and 5 are the parallels for this sample to Figs. 1 and 2 for the SIRAS sample and appear qualitatively identical. As an *a posteriori* check, one may compare the SLFs for C1 and C2 estimated directly from the observations with the average predicted LF in 100 random samples. C2 appears self-consistent, but C1 is deviant. A further check shows that while the SLF2s in the redshift ranges $0 < z \leq 0.15$ and $0.15 < z \leq 0.3$ are consistent with the overlapping portions of SLF2 in the redshift range $0 < z \leq 0.3$, the C1 LFs differ, in keeping with the concept of LE. However, extremely rapid evolution is required to achieve self-consistency for C1. For example, with the weak evolution estimated in ref. 20, it remains the case that in 10,000 random samples, the C1 prediction for the dispersion in apparent magnitude invariably exceeds that observed, and similarly for the observed and predicted correlation of absolute magnitude with redshift.

These results have been checked by analysis of a variety of subsamples, both with brighter flux limits and in other redshift ranges. However, C1 is consistently deviant in qualitatively the identical fashion to the samples discussed earlier, with or without the LE proposed in ref. 20, in the redshift ranges up to 0.1, 0.15, 0.2, or 0.3, as well as the intermediate ranges $0.1 < z \leq 0.2$ or $0.1 < z \leq 0.3$, or with

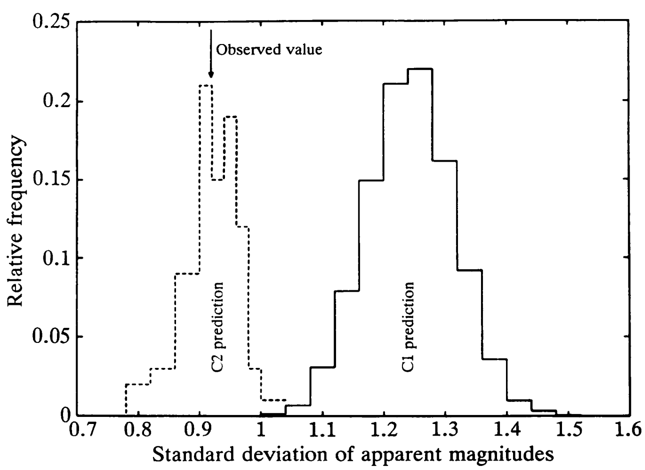

FIG. 4. The same as Fig. 1 for the EMSS-AGN sample in the range $z \leq 0.3$ (224 sources).

brighter flux limits. For example, with a limit 1.5 mag brighter than the reported sensitivity limit for AGN, in the well-observed range $z \leq 0.3$, the predictions of C1 for the dispersion in apparent magnitude in the corresponding subsample of 90 AGN, in 100 random samples, invariably exceed the observed dispersion. At the same time, ROBUST correction reduces the absolute value of the correlation with redshift from 0.91 to 0 and considerably <u>reduces</u> the C1 dispersion in absolute magnitude, so the problem is not with the effectiveness of ROBUST. There is also the same inability as in the IRAS samples for C1 to predict the observed correlation of the C1 absolute magnitudes with redshift, although this correlation is accurately predicted by C2 (as well as the correlation of the C2 absolute magnitudes with redshift). Very bright optical subsamples of the EMSS-AGN appear likely to be substantially complete optically to apparent magnitudes of 17 or 16.5 and show qualitatively identical statistical effects.

These results are very much of a piece with our earlier optical studies, but the dispersion in apparent magnitude, which was not treated in the optical studies, adds a new dimension. *A priori*, one would expect the prediction by a correct theory for this dispersion to be, if at all deviant from direct observation, somewhat less than the observed dispersion. Presently unknown or unobserved factors—e.g., possible discordant redshifts as proposed by Arp (21)—and of course peculiar velocities, could <u>add</u> to the dispersion. But there is no known mechanism in general statistics or specific mechanism in astrophysics that could <u>diminish</u> the dispersion by the order of the more than 25% deviation of the C1 prediction from observation. C1 thus appears irreconcilably inconsistent with substantial and model-independent observations in a variety of wave bands and redshift ranges, from that just beyond the range in which it originated to that of the bright cluster galaxies that have been claimed to establish it definitively. One might suspect that an unknown and unprecedented factor was involved if all power laws were deviant in the same way, but the C2 predictions are quite accurate, although they involve no adjustable cosmological parameters or evolution.

Of course, no theory can fit discordant observations, of which there are a plenitude in samples observed or corrected on a model-dependent or subjective basis. Moreover, basically equitable and objective samples can be subjected to ad hoc selection procedure or special tests designed for the occasion that may appear to invalidate any given theory, since observed samples always occupy a set in event space of probability zero, where continuous variables are involved. A variety of such scientifically inappropriate criticisms of CC have been published, but space permits consideration of only the major one—that the C2 predictions do not fit the cluster galaxy observations. These observations have been described here by van den Bergh, without however noting that those that most closely fit C1 consist of samples *selected in whole or in part assuming the Hubble law*—e.g., from the Abell catalog (22), which states (p. 214) *"in determining whether a cluster meets this criterion, it was assumed that the red shifts of clusters are proportional to their distances"* (our emphases).

Such samples can hardly be used properly to test theories with different redshift–distance laws, notwithstanding claims that *de facto* if not *prima facie*, the selection criteria were model-independent. (Additional selection problems, such as the attempted deletion of cD galaxies, are no less questionable from the standpoint of statistical equitability.) They can however be used for consistency checks on the Hubble law. The sample of Hoessel *et al.* (23), which was described by van den Bergh (24) as a complete sample drawn from the Abell catalog that is an "almost ideal database for the study of the velocity–distance relation," has a closeness of fit to the Hubble law, as measured by rms scatter, that is "remarkable." However, a statistical analysis of the same type as earlier indicates that C1 is self-inconsistent, as regards its predictions for s_m and ρ_1. More specifically, in 100 random samples constructed assuming C1, from the SLF1 for the sample assuming its completeness to its faintest apparent magnitude, the dispersion in apparent magnitude of the 116 galaxies of the sample invariably exceeded the observed dispersion. The observed and C1 predicted values for s_m were 0.96 and 1.05 ± 0.03 mag, a deviation of 3σ. Similarly, the observed and predicted values for the slope of the magnitude to log$z$ regression were 4.61 and 5.01 ± 0.16.

In addition to being subjectively selected, the bright cluster galaxy samples typically made extensive corrections that are theory-dependent. As noted by Humason *et al.* (25), inappropriate corrections for aperture, which are substantial and depend on q_0, which is unknown, may lead to a trend with redshift in the observations that tends to obfuscate cosmological implications. It is interesting to note that the sample of Gunn and Oke (26), which attempts a more objective treatment of aperture and other corrections than earlier studies, and is largely independent of the Abell catalog, is fit by C2 slightly more closely than by C1 (in terms of rms scatter), notwithstanding that the apertures involved are based on C1.

Thus bright cluster galaxies, as actually observed and reported, at best provide a weak consistency check for C1 and do not in the least rule out C2. The knowledgeable strictures of Zwicky (27) concerning cluster galaxy samples, "Hubble's method is subject to many objections . . . great caution is indicated," seem too rarely to have been heeded. In any event, bright cluster galaxies form a small fraction of the general galaxy population, whose selection model-independent magnitude observation is complex and difficult, if possible at all, and now appears unnecessary for cosmological testing purposes.

THE LARGE-REDSHIFT REGIME

Our charge was to treat the redshift–distance relation, but at larger redshifts the concept of distance becomes too ambiguous to be appropriate for direct observational analysis. The issue may be moot at redshifts up to about 0.3, within which both FLC with empirical values for its parameters and CC are well approximated by power laws, and space appears close to its local euclidean tangential approximation. But a more positivistic approach is essential at large redshifts, and before entering into the theoretical aspects required to specify the distance concept, the directly observed flux–redshift relation will be examined.

FLC now postulates various forms of evolution, especially LE at larger redshifts, and only CC makes a specific prediction, to the effect that the flux F varies as $(1 + z)/z$, apart from the cutoff bias, the usual corrections for nonstandard spectra, etc. This relation is a directly observable one that is testable in flux-limited samples by the same statistical methods as earlier.

Two of the largest and most fully documented samples that extend to large redshifts are the BQS (up to $z = 2.15$) and the EMSS-AGN (up to 2.87). The former was treated in ref. 28 by the methods earlier indicated, and it was found that the CC predicted flux–redshift relation is consistent with these observations. The latter has recently been studied in its full redshift range and, under the same completeness assumption as earlier, is also consistent with CC. Moreover, as in the case of the BQS also, CC implies that the predictions of nonevolutionary FLC will deviate from direct observation in the way that is observed. As in the low-redshift regime, this indicates the scientific redundancy of the evolutionary hypotheses that have been proposed for the exculpation of FLC, but which appear incapable (at present, at least) of direct observational substantiation.

Although the term "distance" has to be a theoretical one in the extreme distance regime, a unique redshift–distance relation can be derived from the flux–redshift law, if one makes the simplest assumptions about the structure of space; and this relation is moreover confirmed by further analysis of observed data. Specifically, assuming that space is either spherical or euclidean, then the variation of F with $(1 + z)/z$ together with the assumed absence of singularities in finite parts of space first rules out the euclidean possibility and then implies that if r is the distance in radians on spherical space, then $z = \tan^2(r/2)$. The argument here is based on the known variation of flux inversely with the surface area at a given distance. The r resulting quasi-empirical law is identical to the theoretical redshift–distance law in CC (31, 43).

The determination in laboratory units of the length of one radian in spherical space S^3 of three dimensions—in effect, the cosmic distance scale—can be made from the directly observed proper motion to redshift relation for superluminal sources. The proper motion, like the apparent diameter, varies inversely with the square root of the redshift in CC. From this follows (29) a statistically consistent estimate of the cosmic distance scale (or "radius of the universe") R that is an explicit function of the observed proper motions and redshifts. Inserting the data reported in ref. 30 into this analytic expression, R is estimated as 160 ± 40 Mpc (1 parsec = 3.09×10^{16} m). Because of projection and statistical effects (31), this value is quite consistent with estimated ages for the oldest discrete sources, notwithstanding that it appears an order of magnitude smaller. To the extent, however, that there may be a significant cutoff in ref. 30 on the observation of smaller superluminal motions, the estimation procedure involved in ref. 29 will benefit by refinement along the lines of ROBUST.

The final redshift–distance relation that results:

$$z = \tan^2(r/2R); \quad R \approx 160 \pm 40 \text{ Mpc},$$

serves to explain, without ancillary hypotheses, the nonparticipation of the Local Group in the "general expansion" noted by Hubble. At the same time, it explains the tendency of estimates of the Hubble parameter H to increase with redshift and in some degree to reconcile disparate estimates of H, for which it gives the value 100 km/s/Mpc at $z = 0.01$. An incidental result is the reconciliation of the extant observations on superluminal sources with the expected isotropy of the angles between the directions of the proper motion and the line of sight; this in fact provides the basis for a nontrivial *a posteriori* check on the estimate.

4

THEORETICAL CONSIDERATIONS

The scientific evaluation of a proposed redshift–distance relation necessarily involves theory as well as the statistics of direct observations, as already noted. Moreover, the questions naturally arise of whether CC applies to extragalactic astronomy but not on the microscopic scale, and relatedly, of whether it is only an effective theory and not a fundamental one. Finally, the origin and historical development of FLC and CC are relevant to the evaluation of their proposed redshift–distance laws.

CC has been claimed by some to be deviant from known principles of local physics and therefore *a priori* dubious. But it is very much in line with contemporary theoretical physical development, from Maxwell, Mach, and Einstein to particle physics, and it is rather FLC that is fundamentally deviant. First, FLC introduces the radical concept of an intrinsically Expanding Universe, for which there is no precedent, and a nontrivial degree of philosophical skepticism. In contrast, in CC space is fixed, as in traditional physics. Second, while a suitable *local* Doppler theory may be compatible with the law of the conservation of energy, a *cosmic* Doppler theory abandons this law. There is no generic global conserved energy in GR (e.g., ref. 32), and it is clear from the explicit time dependence of space in FLC that there can be none in these cosmologies, implying even the absence of local energy conservation. In contrast, in CC the energy is conserved, as in traditional and modern microphysics. The redshift arises because some of the energy becomes diffuse (quasi-gravitational) and only a local component is directly observed by existing telescopes.

4

CC arose as a further implementation of the "ansatz" that Minkowski proposed, to explicate the essential idea of special relativity (SR)*. In the terminology, for example, of Faddeev and of Lichnerowicz, SR is a *deformation* of Newtonian physics, to which it tends as $c \to \infty$. Quantum theory can be understood in these terms as a similar deformation of classical physics, as $h \to 0$. But physics requires three fundamental units—e.g., the idea of a fundamental length was proposed by Heisenberg in 1946 (33). CC connects Minkowski's ansatz with Heisenberg's proposal by a deformation involving the cosmic distance scale R, rather than the microscopic length contemplated by Heisenberg, thereby providing the last fundamental unit required by physics. Perhaps unexpectedly and paradoxically, this length is an invariant under symmetry transformations, and its introduction effectively removes prototypical ultraviolet divergences in nonlinear quantum field theory (34). The elimination of such divergences was the main objective that Heisenberg sought by the introduction of a microscopic length.

The intractability of the ultraviolet divergences in quantum field theory led Schwinger (35) to conclude that "a convergent (quantum field) theory cannot be formulated within the framework of present space–time concepts." Basically, CC modifies these concepts by postulating that the proper space–time arena for fundamental physics is a variant of Minkowski space that may be called the Einstein–Maxwell cosmos. It is conformally equivalent to the Einstein universe (EU) at the same time that it is the maximal space–time to which solutions of Maxwell's equations (or those of Yang–Mills, etc.) canonically extend.

In the language of quantum field theory, it is a bare space–time, or in classical terms, an empty or reference space–time, which is conformally invariant. In mathematical terms, it is the homogeneous space consisting of the quotient of the (conformal) group $SU(2,2)$, in its simply connected form G, modulo its maximal parabolic subgroup (which includes the Poincare group and scaling). It possesses an essentially unique locally Minkowskian causal structure that is invariant under G. There is no invariant metric under G, but there is an invariant length—namely, the distance between antipodal points—when represented as the EU, $R^1 \times S^3$. Again in quantum field theoretic terms, this bare space–time is clothed by the energetic contents of the universe, by a natural interpretation of Mach's principle in this context. More specifically, these contents determine a time × space frame in which the total energy is minimal; this clothed frame is precisely identifiable with the EU. Thus the large-scale contents of the universe break the symmetry down from the conformal group G to the Einstein isometry group K (which in mathematical terms is the maximal essentially compact subgroup of G, unique within conjugacy). There is an eight-parameter family of such splittings into time × space, or equivalently of Ks, and minimization of the energy picks out the splitting that may be described as the standard of rest, or inertial frame, of the universe, according to CC.

At first glance, this may appear to depart from a concept central to GR, that space–time exists only as a manifestation of the interactions of its energetic contents. In fact, however, it merely adapts this *classical* format to the standpoint introduced into quantum field theory by the advent of renormalization theory and lays preliminary ground for the quantization of gravitation. The Einstein equation survives and achieves a direct physical interpretation as the simplest causal equation for the propagation of local versions of the EU, which arise from local perturbations of the "bare" (locally Minkowskian) causal structure by local matter, apart from a locally unobservably small $1/R$ term representing the intrinsic curvature of S^3.

The absence of an invariant metric under G thus appears not at all as a blemish, but as a physically essential feature, which provides a basis for resolving more fully the classic dispute as the nature of space and time, between Newton on the one hand and Huygens and Leibniz on the other (36, 44). The position of the latter was updated by Mach and Einstein and seems unquestionably correct from a fundamental standpoint. However, the Newtonian position is essential for practical purposes and is in effect updated by modern quantum field theory. In this, it seems necessary for experimental purposes to represent the observed physical particles, which are continually interacting with the contents of the universe, by "free" particles that partake of the fundamentally mythical quality of Newton's absolute space and time. These free particles are represented as inhabiting a Poincare group invariant space–time, that of Minkowski, whose structure is however unaffected by the particles. The G-invariant space–time of CC is merely a canonical variant of Minkowski space—namely, the universal cover (i.e., simply connected form) of its conformal compactification.

Minkowski's explication of SR was that in essence it replaced the generators of transformation to moving axes in Newtonian theory, $x_0\partial_j$, by the corresponding Lorentz boost, obtained by the addition of $c^{-2}x_j\partial_0$, which evidently tends to zero as $c \to \infty$. CC makes an analogous proposal for the generator of time evolution. In essence, it replaces the Newtonian generator ∂_0 by adding to it $R^{-2}[(1/2)x_0(x_0\partial_0 + x_j\partial_j) - (1/4)x^2\partial_0]$, which looks complicated but simply serves to produce the generator ∂_t of time evolution in the EU, t being the Einstein time. The two operators ∂_t and ∂_0 do not commute, so that the corresponding energies cannot be simultaneously conserved. CC ascribes the redshift to the difference between these energies. A careful study of Maxwell's equations (37) shows that while the Minkowski energy (i.e., that corresponding to ∂_0) of a photon is determined by the local structure of its wave function, principally the infinitesimal oscillation frequency, the Einstein energy (i.e., that corresponding to ∂_t) includes an additional positive term

*Minkowski, H., Address to the 80th Assembly of German Natural Scientists and Physicists, 1908, Cologne, Germany.

that depends on the global structure of the wave function, principally the number of oscillations in addition to their infinitesimal frequency. This term is proportional to the space curvature R^{-1} and is unobservably small for a freshly emitted localized photon, but in the course of cosmic time it increases nontrivially, while the more localized energy corresponding to ∂_0 that is observed by present photon detectors decreases. The result is that a cosmically aged photon appears redshifted, but its total Einstein energy is conserved.

The explicit time dependence of the FLC models break Lorentz invariance, as well as conservation of energy–momentum. The loss of such theoretically fundamental as well as experimentally well-established symmetries must be weighed in the balance against the apparent simplicity of a Doppler explanation for the redshift. Hubble and Tolman (38), among others, expressed doubts about the Doppler theory; they suggested that the redshift would be more intelligible as a space curvature effect. As just described, the chronometric redshift is consistent with this suggestion. Equivalently, every free photon wave function in Minkowski space extends canonically to one that is defined throughout the EU. Its Einstein energy is then given by the integral over the Einstein space S^3 of the square of the field strength in curvilinear coordinates and exceeds the corresponding expression in rectilinear coordinates by an amount that varies as R^{-1} (45).

More generally, the CC framework applies with significant effect not only to the redshift but also to the diffuse background radiation (41, 46), to microphysics (34, 39, 40), and to gravity (36, 44).

The cosmological statistical studies described above would not have been possible without the collaboration of J. F. Nicoll. Extension of chronometric theory to the microphysical contest was very greatly facilitated by collaboration with the late S. M. Paneitz. I thank also D. Johnson, W. Segal, P. Wu, and Z. Zhou for their collaboration in the cosmological studies. For collaboration in the development of relevant group and quantum field theoretic aspects, I thank H. Jakobsen, B. Orsted, B. Speh, M. Vergne, D. A. Vogan, Jr., and Z. Zhou. Thanks are also due B. T. Soifer, J. T. Stocke, M. A. Strauss, and A. E. Wehrle (IPAC) for e-mail communication of data treated above.

1. Commission on Life Sciences (1984) *Science and Creationism* (Natl. Acad. Sciences, Washington, DC).
2. Nicoll, J. F. & Segal, I. E. (1980) *Astron. Astrophys.* **82**, L3–L6.
3. Nicoll, J. F. & Segal, I. E. (1983) *Astron. Astrophys.* **118**, 180–188.
4. Lynden-Bell, D. (1971) *Mon. Not. R. Astron. Soc.* **155**, 95.
5. Schmidt, M. (1968) *Astrophys. J.* **151**, 393.
6. Nicoll, J. F. & Segal, I. E. (1985) *Astron. Astrophys.* **144**, L23–L26.
7. Schmidt, M. & Green, R. F. (1983) *Astrophys J.* **269**, 352.
8. Nicoll, J. F., Johnson, D., Segal, I. E. & Segal, W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6275–6279.
9. Visvanathan, N. (1979) *Astrophys. J.* **228**, 81–94.
10. Segal, I. E. (1989) *Mon. Not. R. Astron. Soc.* **237**, 17–37.
11. Segal, I. E. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7129–7131.
12. Soifer, B. T., Sanders, D. B., Neugebauer, G., Danielson, G. E., Lonsdale, C. J., Madore, B. F. & Persson, S. E. (1987) *Astrophys. J.* **320**, 238–257.
13. Nicoll, J. F. & Segal, I. E. (1982) *Astron. Astrophys.* **115**, 398–403.
14. Strauss, M. A., Davis, M., Yahil, A. & Huchra, J. P. (1990) *Astrophys. J.* **361**, 49–62.
15. Segal, I. E., Nicoll, J. F., Wu, P. & Zhou, Z. (1993) *Astrophys. J.*, in press.
16. Smoot, G. F., Bennett, C. L., Kogut, A., Aymon, J., Backus, C., De Amici, G., Galuk, K., Jackson, P. D., Keegstra, P. & Rokke, L. (1991) *Astrophys. J.* **371**, L1-L5.
17. Rowan-Robinson, M., Saunders, W., Lawrence, A. & Leach, K. (1991) *Mon. Not. R. Astron. Soc.* **253**, 485.
18. Saunders, W., Rowan-Robinson, M., Lawrence, A., Efstathiou, G., Kaiser, N., Ellis, R. S. & Frenck, C. S. (1990) *Mon. Not. R. Astron. Soc.* **242**, 318–337.
19. Gioia, I. M., Maccacaro, T., Schild, R. E., Wolter, A., Stocke, J. T., Morris, S. L. & Henry, J. P. (1990) *Astrophys. J. Suppl.* **72**, 567–619.
20. Maccacaro, T., Della Ceca, R., Goiia, I. M., Morris, S. L., Stocke, J. T. & Wolter, A. (1991) *Astrophys. J.* **374**, 117–133.
21. Arp, H. (1987) *Quasars, Redshifts, and Controversies* (Interstellar Media, Berkeley, CA).
22. Abell, G. O. (1958) *Astrophys. J. Suppl.* **3**, 211–288.
23. Hoessel, J. G., Gunn, J. E. & Thuan, T. X. (1980) *Astrophys. J.* 486–492.
24. van den Bergh, S. (1992) *Proc. Natl. Acad. Sci. USA* **90**, 4793–4797.
25. Humason, M. L., Mayall, N. U. & Sandage, A. (1956) *Astron. J.* **61**, 97.
26. Gunn, J. E. & Oke, J. B. (1975) *Astrophys. J.* **195**, 255.
27. Zwicky, F. (1959) *Clusters of Galaxies: Handbuch der Physik LIII* (Springer, Berlin).
28. Segal, I. E. & Nicoll, J. F. (1986) *Astrophys. J.* **300**, 224–241.
29. Segal, I. E. (1990) *Mon. Not. R. Astron. Soc.* **242**, 423–427.
30. Cohen, M. H., Barthel, P. D., Pearson, T. J. & Zensus, J. A. (1988) *Astrophys. J.* **329**, 1.
31. Segal, I. E. (1976) *Mathematical Cosmology and Extragalactic Astronomy* (Academic, New York).
32. Dirac, P. A. M. (1959) *Phys. Rev. Lett.* **2**, 368–371.
33. Heisenberg, W. (1946) *Z. Naturforsch.* **1**, 608.
34. Segal, I. E. & Zhou, Z. (1992) *Ann. Phys.* **218**, 279–292.
35. Schwinger, J. (1958) *Selected Papers on Quantum Electroddynamics* (Dover, New York).
36. Segal, I. E. (1987) in *Infinite-Dimensional Lie Algebras and Quantum Field Theory*, eds. Doebner, H. D., Hennig, J. D. & Palev, T. D. (World Sci., London), pp. 78–96.
37. Segal, I. E. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 669–673.
38. Hubble, E. P. & Tolman, R. C. (1935) *Astrophys. J.* **82**, 302–337.
39. Segal, I. E. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 994–998.
40. Pedersen, J., Segal, I. E. & Zhou, Z. (1992) *Nucl. Phys. B.* **376**, 129–142.
41. Segal, I. E. (1983) *Phys. Rev. D Part. Fields* **28**, 2393–2401.
42. Stocke, J. T., Morris, S. L., Gioia, I. M., Maccacaro, T., Schild, R., Wolter, A., Fleming, T. A. & Henry, J. P. (1991) *Astrophys. J. Suppl.* **76**, 813–874.
43. Segal, I. E. (1972) *Astron. Astrophys.* **18**, 143–148.
44. Segal, I. E. (1986) in *Quantum Theory and the Structures of Time and Space*, eds. Castell, L. & Weizsäcker, C. F. v. (Hanser, Munich), pp. 120–137.
45. Segal, I. E. (1984) in *J. C. Maxwell*, ed. M. S. Berger, (North-Holland, Amsterdam), pp. 245–262.
46. Segal, I. E. (1987) *Astrophys. J.* **313**, 543–546.

4