Towards a More Well-Founded Cosmology

Hartmut Traunmüller
Stockholm University

Abstract: First, this paper broaches the epistemological status of scientific tenets and approaches: phenomenological (descriptive only), well-founded (solid first principles, conducive to deep understanding), provisional (can be falsified if universal and verified if existential), and imaginary (fictitious entities or processes, conducive to empirically unsupported beliefs). The ΛCDM "concordance model" involves such beliefs: the emanation of the universe out of a non-physical stage, cosmic inflation (invented ad hoc), Λ (fictitious energy), and exotic dark matter. Big Bang cosmology further faces conceptual and pragmatic problems in delimiting what expands from what does not. The problems dissolve after untying inertia from space. The cosmology that emerges appears immediately compatible with the considered observations and the 'perfect cosmological principle'. Waves and field perturbations that propagate at c expand exponentially with distance (a gravitational effect). The cosmic web of galaxies does not. Potential Φ varies as -H/(cz) instead of -1/r. Inertial forces arise from the gravitational action of the rest of the universe. Due to dilatation, they are reduced disproportionately at low accelerations. A cut-off value a0 = 0.168 cH is deduced. This explains the successful description of galaxy rotation curves by MoND. A fully elaborated physical theory is still pending. Wider implications are briefly discussed.

Keywords: foundations of science; cosmology: observations; cosmology: theory; galaxies: kinematics and dynamics; inertia; MoND

1. Introduction

Empirical science involves acquiring knowledge with an aim to organise, explain and understand phenomena. This knowledge is in part ‘existential’ and in part ‘universal’ (generic). The latter type of knowledge can be conceived of as a set of empirically testable universal claims that have not yet been convincingly falsified and so remain tenable. In addition to this, which is most prominent in Popper's [1] philosophy of science, science also makes existential claims, of which Pauli’s prediction of the existence of neutrinos [2] is an example. Existential claims can only be verified rather than falsified empirically. (Only the existence of something at a specific place and time can be falsified just as well as verified.) In this conception of science, it is in neither case necessary for the postulates and hypotheses, which give rise to the claims, to be understood. It suffices for them to be tenable given the empirical evidence. However, it can be argued that the ultimate aim of basic research is to extend the body of empirical knowledge that can be rationally explained ab initio, i.e., without reliance on any assumption that is not understood. These can never be explained within the theory based upon them. Assumptions or hypotheses that are not understood still have important functions in science at a less advanced stage of development, but they must be considered as tentative and provisional.
In order to really understand phenomena and the relations between these, we need theories that rest on a foundation of solid knowledge. This may involve other well-founded, more fundamental theories. Ultimately, well-founded theories are based solely on definitions and first principles of the kind that cannot easily be rejected using the Cartesian method of doubt. These are principles that are either accepted even outside the frame of the particular theory or indispensable for there to be a theory at all. In the present paper, the notion of ‘first principle’ is always to be understood in this narrow sense. An ‘axiom’ does not necessarily qualify as a first principle in this sense.

Indispensable axioms whose validity is independent of nature lie at the foundation of the formal sciences. These give us the rules of logic, algebra and geometry, which then can be taken as first principles in all sciences. It may not always be clear what can be taken as a first principle, but many theories build on a postulate that can easily be called into question. In such cases, it is clear beyond any doubt that the postulate does not qualify as a first principle in our sense. A theory that depends on it cannot be more than a speculative, conditional and provisional one, even if its predictions are compatible with all available empirical evidence, no matter how accurately. It will remain ‘just a theory’ even if ‘corroborated’ by evidence. While many theories are of this kind, there are also more well-founded ab initio approaches.

Physical ab initio approaches have been pursued in chemistry (e.g. ab initio quantum chemistry, ab initio molecular dynamics) as well as in physics. However, in physics, there is a strong tradition of attempting to reconcile empirical knowledge with a few traditional standard paradigms that may fall short of satisfying the mentioned criteria of well-foundedness. It is well known that inferior paradigms and standards tend to persist because of the legacy they have built up, like the QWERTY layout in typewriters [3]. Such “path dependence” is also prominent in the history, teaching and practice of science. This has, in effect, been noticed by Kuhn [4] in his study of scientific practice, but the undesirable “lock-in effects” of path dependence have not yet found the attention they require there. These have been mainly discussed in the field of economics, and the few papers on path dependence in epistemology also originated there [5, 6].

The history of science shows us that questionable assumptions on which previously established theories had been based tend to be retained not only as long as they remain compatible with the empirical evidence but as long as they can be made compatible with it by ad hoc means. Standard cosmology is a prominent case in point, and it had a precursor already in Newton’s questionable treatment of inertia as an effect of space (not of the matter in it), which Einstein retained and extended in General Relativity (GR).

In current standard cosmology, the Big Bang (BB) paradigm is taken for granted. Due to its free parameters and liberal allowance for evolution, it is flexible, but it happened that new or previously neglected evidence was found to be incompatible with it nevertheless. In such cases, a theory stands falsified until a convincing explanation of the discrepancy is presented. Although this is clear enough, it is not very rare in scientific practice that falsifications are brushed aside by advancing excuses in the form of ad hoc assumptions and constructs, also purely imaginary ones, which can only be believed in. Such adherence to traditional paradigms is characteristic of what Kuhn [4] called “normal science” as opposed to “revolutionary science”. It entails advantages for individuals and their collaboration, but approaches that depend on ‘credence’ in ad hoc fudge factors can, in the long run, hardly be claimed to remain within the bounds of ‘science’ at all. They are symptomatic of a degeneration of the science into a fossilized system of unquestioned doctrines.
We shall take a look at the epistemological status of the assumptions that have led to the standard model of BB cosmology, the $\Lambda$CDM concordance model, and contrast this model with the implications of alternatives in which ad hoc solutions are avoided and the most deeply rooted one of the questionable physical tenets, the association of inertia with space, is dropped, while conservation of energy is taken as a first principle and the “perfect cosmological principle” (PCP) as a generalising assumption. The latter implies that the universe is persistent instead of transient. It will be shown that the astronomical evidence that requires excuses in order to maintain the BB paradigm appears compatible with a persistent universe.

2. Tenets and Their Corollaries

Among scientific approaches to natural phenomena one can distinguish between phenomenological approaches, which are inductive and founded on observations, and theoretical approaches, which are deductive and founded on premises. There is often an interplay between these.

In purely phenomenological approaches, regularities among observations are searched and described without offering an explanation. They yield organised knowledge, empirical relationships, but only a superficial understanding. Exploratory data analysis is an archetypal method. Phenomenological models make use of formalisms and free parameters. A well-known example is present in Kepler’s laws of planetary motion.

Theoretical approaches offer, in addition, an explanation of observations. This can also be said of empirical approaches that invoke first principles. Among theoretical approaches, at least three epistemologically different types need to be distinguished. These are listed in Table 1 according to the nature of the tenets they profess.

1) First principles. In cases in which these are sufficient, they lead to well-founded theories and predictions and to explanations that can be understood ab initio. If this is the ultimate goal, first principles need to be respected also in advanced empirically founded approaches.

2) Tentative assumptions (sometimes called “postulates”) that appear reasonable but remain subject to doubt since they are not rooted outside the theory in question and can never be proven within it. These lead to provisional (conditional) theories and to explanations that hold to the extent to which the assumptions hold.

3) Assumptions that, in addition to not being rooted outside the theory in question, also lack independent empirical support. Any reasoning based on these remains within the domain of imagination. Such assumptions are ‘fictitious’\(^1\) and lead to empirically unsupported beliefs. Modern theoretical physics offers a range of “fairy tale physics” \cite{7} in which fictitious assumptions are either primary, as in string theory, or secondary, as in the “dark sector” of BB cosmology, discussed in Section 3.

\(^1\) By “fictitious”, we mean ‘merely existing in theory, not in reality’. In contrast, so called “fictitious forces” are never fictitious in this common sense, but rather in the opposite sense, which reflects a theory-centred world view that is characteristic of theoretical physics.
Table 1. Epistemologically different types of theoretical approaches and tenets, the confidence \( C \) these impart (a multiplicative variable), their type of adequacy and their epistemological yield.

<table>
<thead>
<tr>
<th>Tenets</th>
<th>Foundation</th>
<th>Confidence</th>
<th>Adequacy</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Well-founded</td>
<td>definitions, first principles</td>
<td>( C = 1 )</td>
<td>explanatory</td>
<td>deep understanding</td>
</tr>
<tr>
<td>2 Provisional</td>
<td>&quot; + tentative assumptions</td>
<td>( 0 &lt; C &lt; 1 )</td>
<td>tentatively explanatory</td>
<td>uncertain understanding</td>
</tr>
<tr>
<td>3 Fictitious</td>
<td>&quot; + fictitious assumptions</td>
<td>( C = 0 )</td>
<td>formal</td>
<td>empirically unsupported belief</td>
</tr>
</tbody>
</table>

The values listed in Table 1 under “Confidence” express the confidence we can have in the tenets and the explanations these suggest. They depend on how well the tenets are rooted in what is already understood. We can be fully confident if the tenets are well-founded (type 1). If they really are, our confidence remains undiminished even when confronted with discrepant empirical data. If, on the other extreme, an entity or process is fictitious within the frame of existing knowledge (type 3), it is not scientifically justified to rely on any explanation based on it: \( C = 0 \), exactly. This holds even if the approach leads to predictions that are compatible with the evidence. The provisional approaches (type 2) lie between the extremes 1 and 3 (\( 0 < C < 1 \)). In these cases, a numerical rating of confidence that would be generally valid is not obvious, except at the level of rank order. It is, e.g. justified to attach more confidence to a reasoning based on a generalising assumption that has not been falsified than to an alternative that can be said to involve the same assumption under a restrictive condition that needs to be specified. The latter is equivalent to having two assumptions instead of just one, and the higher confidence in an approach that needs fewer assumptions reflects the principle of parsimony (Ockham’s razor).

Valuable, even trustworthy predictions of entities that have never been observed are not precluded in this scheme. Such entities are not necessarily fictitious within the frame of existing knowledge. In order for us to be confident at \( C > 0 \) into their real existence, it is only required that \( C > 0 \) for each of the tenets on which the prediction is based.

It is well known that empirical falsifications are not firmly conclusive. They are only valid within the frame of the knowledge we have. A statement that stood falsified may even become tenable again in the light of new knowledge. Strictly speaking, universal statements can only be claimed “to be tenable” or “to stand falsified”, unless it follows from definitions and logic alone that they are “true” or “false”. The classification of a tenet as fictitious (\( C = 0 \)) might also change in the light of new knowledge, but as long as we lack this knowledge, our confidence in it must remain at zero.

When confronted with discrepant evidence, the descriptive adequacy of a theory can often be saved by introducing an ad hoc parameter. However, such a parameter has no explanatory power. Worse yet, it invites circular reasoning, and if it represents a fictitious entity, the approach turns into one of type 3 (\( C = 0 \)). This yields just an empirically unsupported belief, e.g. in dark energy. It promotes ‘credence’ – not ‘science’.

Some first principles with \( C = 1 \) can be derived logically on the basis of more widely valid principles and well-founded theories, but ultimately there remains a basic physical principle that

---

2 This is part of the Duhem-Quine problem.
can neither be verified logically nor empirically with full certainty, despite its wide range of empirically proven tenability. Its classification as a first principle is, instead, due to its being indispensable for there to be any ‘law of nature’ and any explanatory science at all. It says that the same physical laws are valid everywhere in space, direction and time. This universality principle expresses a precondition for physics.

There are several conservation laws that can be derived via Noether’s theorem from the homogeneity of spacetime that is implied in the universality principle: conservation of energy follows from the homogeneity of time, conservation of linear momentum from that of space, and conservation of angular momentum from the isotropy of space. While these may be shown to follow from one general symmetry principle, theories do not gain in confidence if the number of first principles they invoke (all with $C = 1$) is minimised. They gain in confidence if the number of tentative assumptions (all with $C < 1$) is minimised, provided that they do not involve any fictitious assumption (with $C = 0$).

While homogeneity and isotropy of spacetime belong to the set of first principles, this is not equally clear for the PCP. The PCP states that the distribution of matter in the universe is isotropic and homogeneous in space and in time. Bondi and Gold [8] considered it a first principle that gives us a reason for assuming that the same physical laws are valid everywhere, but it appears that the matter distribution in the universe is more in need of an explanation than the universality principle. The PCP is at least a clear case of a generalising assumption. The (imperfect) cosmological principle adopted in the BB paradigm is not. It exempts the temporal dimension of spacetime, which is possible in an absolute system of reference but not generally in Minkowski spacetime.

The foundational elements of theories can be listed in the following order:

• definitions and first principles
• generalising assumptions
• more specific testable assumptions
• assumptions involving fictitious entities or processes.

In order to obtain a radically more well-founded theory, it is necessary to reduce the number of required lines in this list from its end. Provided that no assumption with $C = 0$ is retained, theories gain already in confidence if the number of tentative assumptions they invoke is reduced.

3. The Status of Standard Cosmology

In the BB paradigm, which in the late 1990s resulted in the ΛCDM “concordance model”, the universe is finite in age and has emanated under conditions to which physics does not apply. The initial event and the primordial state belong to the fictitious domain. Our confidence in any claims that crucially depend on such an event and state cannot be any larger than zero. This does not bring the confidence in the whole paradigm down to zero, since the event is not introduced as an initial postulate but emerges as a conclusion. Concordance cosmology may still describe reality in approximation if not projected too far into the past.

It is well known that GR allows for an expanding universe and for a contracting one but not for a stationary one, unless a cosmological constant ($\Lambda$) is introduced as a means of preventing the

---

3 The “PCP” needs the attribute “perfect” only because the term “cosmological principle” is in use for the imperfect cosmological principle that is respected also in GR-based cosmologies.
universe from collapsing, as in Einstein's [9] own model of an eternal universe. Einstein had introduced Λ reluctantly, since it did not reflect anything known from physics.

Prior to the advent of BB cosmology, most natural philosophers considered the universe as eternal, but since antiquity there had been a split opinion concerning its spatial extension. According to one, the world is spatially confined. The competing conception of an infinite universe that perpetually regenerates itself and that contains infinitely many similar "worlds", is also ancient. It was argued for by Epicurus, as communicated by Lucretius in De rerum natura.

The first physical model of an expanding universe was presented by Lemaître (1927) [10], who already knew that the redshift \( z = (\lambda_{ob} - \lambda_{em})/\lambda_{em} \) in the light from galaxies tends to increase with their luminosity distance. According to the most straightforward interpretation of this phenomenon as a Doppler shift, the galaxies are rushing away from each other. This interpretation had been adopted by Lemaître [10], but this was not yet a BB model. It assumed eternal expansion from an initial state, at \( t = -\infty \), such as described by Einstein's (1917) [9] model. In BB cosmology, Λ was skipped, but it was formally reintroduced in the ΛCDM model in order to make it compatible with the magnitude–redshift relation of distant supernovae. A non-zero Λ had already been required earlier in order to make the age of the universe indicated by the "Hubble constant" compatible with the estimated ages of the oldest star clusters.

The interpretation of the cosmic redshift as due to an expansion of the universe is compatible with the observed redshifts, but it predicts the angular sizes of distant objects whose proper size is the same to be larger than in a non-expanding universe.

In "tired light" models, which are in line with the Epicurean tradition, the universe does not expand. Instead, it is assumed that light loses energy due to interaction with ingredients of the intergalactic medium or for some other reason. In the tired light model that is most often considered, since Tolman (1930) [11], this loss causes a redshift in the light but no time dilation in its modulation. The tiredness is, thus, manifested in the carrier but not in the information conveyed by its modulation. This type of model stands falsified, since time dilation in accordance with the redshift has been observed in the light curves of distant supernovae [12–18] and in their spectroscopic aging rates [19]. However, it has never been shown that "entirely tired light", with a time dilation consistent with its redshift would be impossible.

Bondi and Gold [8] assumed their PCP to hold. Since they also considered the cosmic redshift as indicative of an expanding universe, they were led to the Steady State theory, in which creation is an ongoing process by which the density in an expanding space is kept constant. This sets the Steady State theory apart from Epicurean cosmology, in which the PCP is also implied, while creation out of nothing is disallowed. Unlike the BB paradigm, which does not adhere to the PCP, and which allows models of 'our universe' (among other 'universes') to be adapted to new observations that falsify previous versions, the Steady State theory made more definite predictions. It lost adherence after the discovery of the cosmic microwave background radiation (CMB), for which it provided no convincing explanation. It can be questioned whether it ever deserved confidence, since the perpetual creation it postulates has remained as fictitious as creation in BB cosmology.

The BB paradigm also fails to provide explanations for several kinds of observational facts. In order to retain it when faced with unexpected observations, it was necessary, in the process of time, to introduce more and more free variables and fudge factors. Some of these arise directly as rational conclusions that can be drawn if the paradigm is accepted a priori. The most important were, in
temporal order, 1) dark matter, 2) cosmic inflation, 3) dark energy, and 4) a particular size evolution of galaxies.

**Dark matter** was suggested by the observed cohesion of galaxy clusters [20, 21] and rotation curves of individual galaxies [22]. These would require much more than the visible matter to be present in order to be compatible with Classical Mechanics (CM) and GR. Initially, the hypothesis that unseen matter in form of gas, dust and substellar objects is responsible for the discrepancy was reasonable (C >0). This matter would need to be present in haloes around galaxies and additional amounts in galaxy clusters. Since the discovery of the discrepancy, the presence of large amounts of gas in galaxies has in fact been verified, but it does not have the required mass and distribution. Neutrinos may also be considered, but the number that would be required by far exceeds the number that can be expected to have been created in a BB universe. Dark matter in form of hypothetical weakly interacting massive particles (WIMPs) is more problematic. Since attempts to verify the existence of WIMPs experimentally have so far failed, it is not justified to attach a non-zero confidence to them. They remain of type 3 in Table 1. As long as the required amount of dark matter is neither predicted on independent grounds nor empirically confirmed to be present, its supposed presence represents an excuse with C = 0. This means in fact that, *at the present state of our knowledge*, GR (as well as CM) stands falsified already at the scale of galaxies. Therefore, we cannot be confident in models of the whole universe based on GR or CM. In a similar vein (based on less numerous cases of clear evidence) CM stood falsified when faced with the anomalous perihelion advance of Mercury and the search for the supposedly responsible planet Vulcan had failed. This problem was solved by GR. The present one is still awaiting its solution – which will be attempted in Section 5.3.

The galaxy rotation curves suggest that the ratio of inertial to gravitational forces is reduced for low accelerations, with a transition value of $a_0 \approx 1.1 \times 10^{-10} \text{ m/s}^2$. This is the essence of Milgrom’s Modified Newtonian Dynamics (MoND) [23, 24], which allows accounting for the rotation curves of all kinds of galaxies in terms of a single function. MoND also provides an explanation of the Tully-Fisher relation, which describes the otherwise unexplained close relation between luminosity and rotation velocity of galaxies. While MoND describes regularities that remain unpredicted by the dark matter hypothesis, it represents a phenomenological approach comparable to Kepler’s approach to planetary motion. It has been shown to be successful for a wide range of different galaxies. This includes galaxies with very low mass, in which the discrepancy with CM and GR is substantially larger than in the galaxies considered when MoND was originally proposed [25, 26]. The fact that MoND describes the rotation curves of galaxies successfully in terms of a function that is at variance with CM and GR suggests that something is wrong with these theories, although MoND still requires substantial amounts of dark matter to be present in galaxy clusters [27].

Among deductive approaches to MoND, two alternatives can be distinguished: 1) modified gravitation (increased where $g < a_0$, and 2) modified inertia (decreased where $g < a_0$). The theories proposed so far [28, 29] are of type 1). They involve, in addition to the Newtonian gravitational force, which varies $\propto r^2$, an otherwise unknown force that varies $\propto r^1$. In a different approach [30], a Newtonian force combines with a Yukawa type of force instead. So far, no deductive approach to the dynamics of galaxies provides a deep understanding. Keeping GR and introducing a new, additional force whose existence has not been verified does not bring about any higher confidence than C = 0. Our confidence in MoND as such is not so low, but it involves uncertainty about the value of $a_0$ and the interpolating function between the regimes $a < a_0$ and $a > a_0$. 

Cosmic inflation [31] is a purely theoretical ad hoc construct. It serves the explicit purpose of reconciling the fact that the universe appears flat, clumpy and yet homogeneous on the largest scale with the BB paradigm, in which such a universe would be an extremely unlikely outcome. It increases the likeliness of such an outcome by assuming physics itself to have been expeditiously different when the universe had not yet reached an age of $10^{-32}$ s. The whole approach and even its logical conclusiveness have long been under debate even amongst those who proposed it [32, 33]. It has not been shown that cosmic inflation is anything else than a fictitious process, and these deserve no more than zero confidence. Amongst astrophysicists, cosmic inflation is not accepted by all, but this leaves the problem it is meant to solve unsolved, which does not lend any markedly higher confidence to the paradigm.

Dark energy is an unpredicted fictitious form of energy with anti-gravitational properties. It is an embodiment of the cosmological constant $\Lambda$, which Einstein [9] introduced as a fudge factor (in form of an integration constant) when he still believed that the universe ought to be static. This $\Lambda$ was reintroduced in order to make the observed magnitude–redshift relation of distant type Ia supernovae compatible with the BB paradigm [34, 35]. In the alternative quintessence cosmology, $\Lambda$ is treated as a parameter that is allowed to vary over time [36]. This is, then, a fictitious variation of a fictitious entity.

Assuming the existence of non-baryonic dark matter and dark energy has sometimes (e.g. [37]) been compared to Pauli’s hesitant prediction of the neutrino, whose existence was verified only 25 years later [2]. These cases had in common that the existence of an entity that had not been known previously was suggested. However, the foundations on which these suggestions rested were epistemologically very different. The nuclear mechanism known as $\beta$-decay appeared to violate a first principle: conservation of energy. Given that this is a principle of the kind in which we can be confident even when faced with evidence that appears to contradict it, the existence of a new particle, which was later named the neutrino, was the simplest conclusion that could be drawn. This was not a fictitious assumption but a well-founded prediction. In contrast, the magnitude–redshift relation of a type of supernovae appeared to violate just the BB paradigm, in which it was not justified to be confident, and which rests on a theory (GR) that in fact stood falsified already in view of the dynamics of galaxies. The BB paradigm stands falsified also in view of this magnitude–redshift relation. $\Lambda$ (dark energy) remains a fictitious excuse that lends no confidence to any reasoning about reality that is based on it.

In models in which the PCP holds, the factor by which waves are stretched per unit of distance $D$ is necessarily constant and everywhere the same. If no other mechanism contributes to the redshift $z$, we have

$$1 + z = \exp\left(\frac{H}{c} D\right),$$

where the Hubble parameter $H$ is a true constant (units $s^{-1}$, km $s^{-1}$ Mpc$^{-1}$). In BB models, the relation is more complicated.

Inverting Eq. (1), $D$ can be calculated as

$$D(z) = \frac{c}{H} \ln(1 + z).$$

Since these expressions differ from those in the usual BB models, they lead to a different predicted relation between redshift and other observables, such as the apparent magnitude of type Ia supernovae.

In a static and flat geometry, the intensity (W m\(^{-2}\)) of light received from a source, flux \(F\) ("apparent luminosity"), varies as \(F \propto D^{-2}\). \(F\) is proportional to absolute luminosity \(L\), defined as the total power radiated by the object. If both the energy of each photon and the number of photons arriving per time unit are reduced by factors of \((1+z)^{-1}\) and if no additional factors are involved, this gives us for an object that radiates isotropically

\[
F = \frac{L}{4\pi D^2(1+z)^2}.
\] (3)

While Eq. (3) has been claimed to be valid in tired light models as well [38], the number of photons arriving per time unit was not reduced in the casual analysis [11] on which the "Tolman test" is based. In this case we get a factor of \((1+z)\) in the denominator of Eq. (3) instead of \((1+z)^2\). Recently, an analysis of redshift and magnitude data from 892 type 1a supernovae, which are the best "standard candles" we have, has shown that the two \(D\)'s that can be calculated on the basis of redshift (Eq. 1) and flux (Eq. 3) are proportional to each other [38], so that astronomical magnitude \(m\) satisfies the relation

\[
m = 5 \log[(1+z) \ln(1+z)] + \text{const}.
\] (4)

This had already been observed previously [39, 40][4] in a smaller set of data that was available then. The conclusion that the redshift factor \((1+z)\) increases exponentially with distance (Eq. 1) was also arrived at in an investigation [41] in which the same tendency was shown to be present also in data from gamma ray bursts. The redshifts of these exceed those of the observed supernovae substantially, but it is questionable to what extent they are due to the distance of the sources. They had been taken into consideration for testing models that involve a \(\Lambda\) parameter [42].

It is clear a priori that a good fit can be obtained in standard cosmology if \(\Lambda\) is allowed to vary as a function of time. However, \(\Lambda\) just describes the error of the \(\Lambda\)-free model, and such fudge factors lack explanatory power. Further, it does not threaten the empirical validity of the simple relation (4), which follows form Eqs. (1), (3) and the definition of \(m\). In a \(\Lambda\) CDM model, the corresponding relation, \(m(z; \Omega_M, \Omega_\Lambda)\), is more complicated and less elegant since it requires numerical integration. If neither \(\Omega_M\) nor \(\Omega_\Lambda\) was fictitious, an alternative that conforms to Eq. (4) directly would still be preferred because of the principle of parsimony.

**Size evolution of galaxies:** If the universe expands in proportion to a scale factor \(a\) so that \(a(t) = (1+z)^{-1}\), while gravitationally bound objects, such as galaxies, do not expand, the angular size \(\delta\) of these will be, in small angle approximation, \(\delta = (1+z)d/D\), where \(d\) can be the major axis diameter of a galaxy and \(D\) its comoving distance. This angle is enlarged by the redshift factor over that in a flat and static universe, where \(\delta = d/D\). While tired light models predict \(\delta \propto \ln(1+z)^{-1}\), all models in which the universe expands but not the galaxies, predict the relation to flatten substantially with increasing \(z\) and \(\delta\) to slowly increase again at large values of \(z\). (With exponential expansion, the prediction is \(\delta \propto (1+z)\ln(1+z)^{-3}\) with a minimum for \(\delta\) at \((1+z) = \text{e}\). In the past, when angular sizes were still considered to make a crucial test of the paradigm possible [43] (pp. 23–25), several investigations

---

4 I was not aware of [39-42] when submitting [38].
returned instead an approximate empirical relation of $\delta \propto z^{-1}$ [44, 45]. Meanwhile, measurements of the angular sizes of galaxies have progressed in scope and reliability without leading to a substantially different result [46]. Allowing a reasonable margin for uncertainties, the observations are, instead, immediately compatible with what would be expected in a universe in which the PCP holds. If the BB paradigm is taken for granted nevertheless, this suggests that galaxies grow in size as $d \propto a(t)$ [47, 48, 49] or slightly more, as $d \propto (1+z)^{-1.2}$ [50]. It is justified to attach some confidence ($0< C < 1$) to this suggestion, since galaxies are expected to evolve in some way within the BB paradigm. This evolution was formerly thought to affect mainly the luminosity rather than the size of galaxies. There is now a hierarchical theory of galaxy formation with too many free parameters [51], according to which galaxies grow by mergers of smaller pieces, dominated by dark matter. This allows modelling the empirical data, but it is an addition to rather than a prediction of the ΛCDM model. A very extensive investigation [52] of 4993 Lyman break galaxies ($4 < z < 10$) reported, within standard cosmology, a growth by $(1+z)^{-0.95}$ for the mean, $(1+z)^{-1.10}$ for the median and $(1+z)^{-1.26}$ for the mode, the discrepancy in the exponents being unexpected.

More recently, the dark sector has been enriched by dark flow. This is an observed large-scale bulk flow of galaxy clusters that appears to be in conflict with concordance cosmology. It has been tentatively ascribed to influences from pre-inflationary inhomogeneities [53, 54]. Even a dark force, a fifth force that affects only the fictitious kind of dark matter, has been contemplated [55]. This leads deep into fairy tale physics.

Even if none of the mentioned excuses would be required, BB cosmology would still be problematic. The energy that radiation loses due to the cosmic redshift, a fraction of $1-(1+z)^{-1}$, disappears without being transformed into any other form. This violates the most basic first principle of physics.5 ΛCDM concordance cosmology stands out as exceedingly speculative. Scientists who are not bound to the path that has led to it can easily see 1) that CM and GR stand falsified at the scale of galaxies and subsequently also at any larger scale, and 2) that the cosmic redshift probably has been misunderstood and so given rise to fudge factors in addition to dark matter. Although even the adherents of concordance cosmology do not usually claim that they understand the universe, it is clear that by pursuing an approach that deserves no confidence, they have driven their disciples and audience into a veritable Dark Age. When a theory persists in standing falsified, it is likely that a wrong choice has been made at a branching of the path that has led to it. In such cases, one should preferably search for the right path, instead of proceeding on the once chosen path and dreaming up an imaginary environment (WIMPs, dark energy, etc.) in which this would be the right path. Let us now find out where this wrong choice appears to have been made.

4. Inertia and Gravitation

CM rests on Galilei’s principle of inertia, according to which physical objects remain in their state of motion as long as no external force impinges on them. Newton took this as his first law of motion. In CM, rotation and motion in general are considered in an ‘absolute space’ and in ‘absolute time’. Most other scholars, such as Descartes, Leibniz, Berkeley and later, Mach, were of the opinion that motion can only be specified with respect to actual objects in space. Newton’s interpretation of

5 If analysed in comoving coordinates, there is no loss of energy, but in these coordinates there is no redshift either.
inertia as an effect of space as such appears to have been the wrong choice that has led us into the dark. Since for a long time it met no empirical counter-evidence, it established itself as the accepted standard, but the empirical fact that the inertia of bodies is proportional to their gravitational attraction remained unexplained and intriguing.

On his path to GR, it was Einstein’s explicit objective to devise a theory in accord with Mach’s view. However, he actually continued on Newton’s path and consolidated it when he conceived of motion in a gravitational field as inertial motion in a curved spacetime. In GR, gravitational and inertial mass are by axiom taken to be the same, and space has an even greater role than in CM, although it is now no longer absolute. In CM, space acts on matter, and this goes without a reaction. In GR, there is a reaction: matter curves space, but this is linked to gravitation only, while inertia remains as exceptional as in CM. Further, the axiomatic linkage between inertia and gravitation makes it impossible to explain the relation between the two within the frame of GR. If a theory of gravitation and inertia is to be well-founded, it must offer such an explanation, which neither CM nor GR does.

On the basis of Mach’s principle, inertial forces can be expected to decrease when the universe expands and its matter density decreases, and to increase in the opposite case. Further, inertial forces will be increased in the vicinity of a gravitating body, since all distant bodies appear blue-shifted and closer there.

An attempt to develop a theory in which inertia is induced by the relative acceleration of the masses of the universe is due to Sciama [56]. This induction is analogous to electrostatic induction in response to relative acceleration of charges. Sciama described a vector theory as a step towards a tensor theory compatible with GR, which he did not put into question. While he investigated some of the cosmological consequences of his theory, he did not devise a cosmology from scratch on the altered premise. He assumed the universe to be expanding, in accordance with the prior interpretation of the cosmic redshift. In attempting to account for inertia on the basis of the density of the universe inferred from astronomical observations, he obtained a missing mass problem of the same magnitude as in GR-based cosmological models, but he expected large amounts of uncondensed and yet unobserved matter to be present between galaxies.

In order to be tenable and useful at the scale of galaxies and above, a theory of gravitation and inertia must also explain the dynamics of these. As a phenomenological model, MoND does this only at the most superficial level. It allows predicting the rotation curves of different galaxies and so can be said to explain the relations between them, but it does it in the absence of an understanding of the underlying physics. The physically founded theories proposed so far [28, 29, 30] attempt to improve this, but since they introduce an ad hoc force (C = 0) of some kind, they actually lack epistemological value. They are also instances of proceeding on the established path and “dreaming up” something that would make it right. The alternative would be to reconsider the path that was vaguely suggested by Mach and not fully appreciated by Sciama, and see where it leads. This will be approached in Section 5.3.

5. In Search of a Tenable Cosmology

5.1. Primary Considerations

If cosmology is to be an empirical science, it is a minimum requirement that its tenets impart more than zero confidence. The theory must not assume any unpredicted fictitious entities or
processes to be in effect. Ideally, it should be based on definitions and first principles alone, but at the present stage of our knowledge it may be necessary to accept, in addition, the PCP (a generalising assumption) in order to progress:

The universe is homogeneous and isotropic in time as well as in space.

Such a universe is persistent instead of transient. Its statistical properties do not change as a function of time, space, and direction. The PCP should, however, only be assumed to hold within volumes that are sufficiently large – that of a Hubble sphere or larger.

Further, gravitation is to be treated in such a way that it allows inertia to emerge as an effect of the gravitation of cosmic masses (Mach’s principle). This, as well as the PCP, is at variance with GR, which is actually promising, considering that GR is incompatible with quantum mechanics and fails to account for the rotation curves of galaxies.

The PCP puts a narrow constraint on the redshift–distance relation. Abstracting influences of nearby masses away, the function must be self-similar and the same everywhere in spacetime: in a flat geometry, this can only be a constant exponential function \((1+z) \propto \exp(D)\), so that Eqs. (1) and (2), which, in addition, only contain the constants \(H\) and \(c\), must hold. Further, if extinction of light is negligible or compensated for, Eqs. (3) and (4) must also hold. This is in excellent agreement with the empirical flux–redshift (and magnitude–redshift) relation of supernovae SN1a [38], which corroborates the tenability of the PCP.

In order to be tenable, a cosmological model must satisfy all the conditions that are listed in Table 2.

<table>
<thead>
<tr>
<th>Phenomenon</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time dilation factor</td>
<td>(1+z)</td>
</tr>
<tr>
<td>Flux–redshift relation and magnitude–redshift relation</td>
<td>(F \propto [(1+z) \ln(1+z)]^{-2}), (m = 5 \log[(1+z) \ln(1+z)] + \text{const}) (5)</td>
</tr>
<tr>
<td>Angular diameter of galaxies</td>
<td>(\delta \propto (1+z)^\alpha \ln(1+z)^{-1}), with (\alpha = 0) (6)</td>
</tr>
<tr>
<td>Cut-off acceleration (a_0) of galaxies (in MoND)</td>
<td>(0.13 &lt; a_0 c^{-1} H^{-1} &lt; 0.22) (7)</td>
</tr>
</tbody>
</table>

Unfudged BB models fail for all phenomena in Table 2, except time dilation. Cosmologies with exponential expansion, such as the “Scale Expanding Cosmos theory” (SEC) [57] and the de Sitter model [58] satisfy also the flux–redshift relation, but the latter model has the blatantly fictitious property of containing no mass. The observed angular sizes of galaxies are at variance with the distance-duality relation. This relation [59] is said to hold if photons travel along null geodesics in a Riemannian geometry, and their number is conserved, but it becomes practically inapplicable if galaxies evolve in luminosity or size with \(a(t)\). If they do not, luminosity distance \(D_L\) is related to angular distance \(D_o\) as

\[D_L = \frac{D_o}{1+z} \propto a(t) \exp(Ht)\]

This model is said to evolve as \(a(t) \propto \exp(Ht)\), but in de Sitter’s original conception, there was no real expansion: “the frequency of light-vibrations diminishes with increasing distance from the origin of co-ordinates. The lines in the spectra of very distant stars or nebulae must therefore be systematically displaced towards the red, giving rise to a spurious positive radial velocity”.

---

6 This model is said to evolve as \(a(t) \propto \exp(Ht)\), but in de Sitter’s original conception, there was no real expansion: “the frequency of light-vibrations diminishes with increasing distance from the origin of co-ordinates. The lines in the spectra of very distant stars or nebulae must therefore be systematically displaced towards the red, giving rise to a spurious positive radial velocity”.

\[ D_L = D_a(1+z)^2. \] (8)

This would require \( \alpha = 1 \) in Eq. (6). A violation of the distance-duality relation has been observed between \( D_L(z) \) of supernovae Ia and \( D_a(z) \) of radio galaxies, compact radio sources and X-ray clusters [60]. Analyses of data from galaxy clusters have more recently been reported to be compatible with the relation (8), the elliptical model fitting better than the spherical [61], or vice versa [62]. However, these papers are concerned with the relation between \( D_a(1+z)^2 \) and \( D_L \) in Eq. (8) and do not report an estimate of the crucial exponent \( \alpha \) in Eq. (6). The result of the investigation by Shibuya et al. [52] is similar to those of previous investigations, which all suggest the exponent \( \alpha \) of the first term in Eq. (6) to be closer to 0 than to 1, but the difference between 0.05 for the mean, –0.10 for the median and –0.26 for the mode is in need of an explanation.

The universality principle supports the assumption that Planck’s radiation law, i.e., the Stefan-Boltzmann law and Wien’s displacement law, should be valid for sources at any distance. This allows for the surface brightness \( SB \) of a redshifted blackbody a maximum of \( SB \propto (1+z)^{-4} \). In the absence of extinction and redshift, \( SB \) does not change with distance. If there is a redshift but no extinction, so that flux accounts for a reduction by \((1+z)^{-2}\), the solid angle the object subtends must increase by a factor of \((1+z)^2\). This condition is satisfied if the distance-duality relation holds, but this is not immediately evident in other cases. In the case of galaxies, there can hardly ever arise a conflict, since their \( SB \) is much lower than that of a black body, and in the case of stars in distant galaxies, the problem remains an academic one, since these are bound to remain point-like sources for which \( SB \) cannot be measured.

The cut-off acceleration \( a_0 \) is the most challenging one of the explananda in Table 2. Its empirical boundaries correspond to \( cH = 6.1 \pm 1.5 \, a_0 \). This should preferably not be a free parameter, which it is in MoND, but emerge from well-founded cosmological considerations.

5.2. The Delimitation Problem

Within GR based cosmology, galaxies are thought of as essentially remaining at rest in an expanding space that brings light waves to expand with it. Standards of comparison, such as sources of radiation and the metre, which is defined in terms of light waves, are, however, tacitly exempted. They are treated as if they did not participate in the expansion of the space they occupy, or as if this space did not participate in the otherwise general expansion. In the absence of such an exemption, there would be no observable cosmic redshift. The FLRW metric and various GR-based alternatives\(^7\) describe the relation between spacetime in presence of gravitation and expansion and spacetime in absence of these factors, e.g. in the equation for the line element. They do not leave anything more extended than a point unaffected. In order to be consistent in exempting anything, such as standards of comparison, it is necessary to know the delimitation criterion, which the metric does not provide.

In the praxis of GR and the FLRW metric, the proper spacetime defined by local standards of comparison is not identified with the spacetime that is valid in the local environment, but with the spacetime that is valid in the absence of gravitation and expansion. This choice is conceptually in error but appears empirically right, since it makes frequency shifts observable. The error is avoided in the common pseudo-Newtonian approach, in which “expansion” is an attribute of structures and distances rather than of “space” in a substantivalist sense. It assumes the expansion, described by the

---

\(^7\) This includes the so called Swiss cheese models.
FLRW metric, to take effect only to the extent to which it is not prevented by forces. In this matter, gravitation is not treated differently from other forces. Thus, it is, essentially, assumed that

(1) waves and incoherent objects expand – coherent objects do not.

In such an approach, it has been calculated [63] that expansion sets in at a radius of 10 Mpc for a structure like the Virgo cluster, which comprises more than 1000 galaxies within a radius of less than 3 Mpc. If objects are coherent up to this size, and with a substantial margin, this delimitation leaves only the voids between galaxy clusters to expand. However, it is questionable whether even these can expand. The universe looks like a reticulated foam or a three-dimensional web of galaxy clusters, which are connected to their neighbours via filaments whose matter density suffices for coherence along their axes, otherwise the filaments would not persist. Persistent filaments might prevent even the voids from expanding. It is, at least, clear that in a young BB universe, the gravitationally bound regions would overlap with each other, which would have prevented the expansion of the universe. Therefore, this BB approach fails pragmatically in delimiting what expands from what does not. Statement (1) is incompatible with an expanding material universe but it remains compatible with a non-expanding one, in which only waves expand. As mentioned under “Size evolution of galaxies” in Section 3, the observed angular sizes of galaxies, Eq. (6), are also hard to reconcile with an expanding material universe.

It is also possible to consider the expanding space of (1) as static while coherent objects shrink and processes speed up [38, 64]. In short,

(1b) coherent objects contract

If the expansion in (1) is exponential, the contraction in (1b) is exponential in reciprocal proportion. Version (1b) describes the same situation as version (1) in a frame of reference that is co-expanding with the waves. Any observable effects are the same. If the cosmic web is coherent, there is no observable expansion or contraction of objects of any size. The entities that contract in the contraction model (1b) include any real standards of comparison. The material universe remains, therefore, metrically static and so compatible with the PCP. The metric space being the same, this brings us back to the expansion model (1).

Within the frame of GR, on which the FLRW metric is based, statement (1) appears incongruous since GR does not draw a distinction between coherent and incoherent objects. Instead, it draws a distinction between non-gravitational forces and gravitation, which it links with space. This link is broken in (1) by exempting gravitationally bound objects from the expansion. GR rather suggests a delimitation between a space in which radiation propagates, and which is also the space of gravitation, and the space of non-gravitational forces, which can be equated with that of CM. A corresponding assumption would be that

(2) anything under free gravitation expands – objects under control of other forces do not.

This alternative predicts the universe, the cosmic web, galaxies and planetary systems all to expand, which is incompatible with the PCP but not necessarily with the observations. The angular size discrepancy may disappear if galaxies participate in the expansion. Further, if planetary systems

\footnote{In [38], I described such a model (not fully understood) as making different predictions from expansion models. However, if the cosmic web does not expand, the predictions based on (1) and (1b) will agree.}
expand, this would be reflected in an increase of the Astronomical Unit (AU). With AU = 149.6 10^9 m, and H = 60 km s^{-1} Mpc^{-1}, there would be a secular increase by 17.8 m for expansion by (1+z). A secular increase of the AU by 15±4 m has actually been reported to be present in empirical data [65]. Essentially the same explanation might also account for the increasing eccentricity of the lunar orbit [66]. However, Pitjeva & Pitjev [67] reported a non-significant increase of the AU by only 1.2±3.2 m per century (at the 3 σ level).

As for the rotation curves of galaxies, criterion (2) amplifies the discrepancy with the astronomical observations. Aside from this trouble, it does very well if it is true that planetary systems and galaxies participate in the general expansion, or perhaps just in the expansion supposedly caused by dark energy (roughly 50%) [68]. The efficacy of criterion (2) can be falsified by demonstrating that the AU does not increase correspondingly. The reports of its increase [65, 66] or absence of significant increase [67] were based on results that are highly sensitive to small errors of various kinds.

An error that might feign or hide a change in the AU appears to be the cause of the “Pioneer anomaly” [69, 70], an unexplained acceleration of about 8.7 10^{-10} m/s^2 directed towards the Sun, observed in the trajectories of space probes. While the anomaly was explained away as a thermal effect [71], it rather reflects an erroneous modification in the acquisition or processing of data that was introduced in 1990. This is evident from an exercise in which publicly accessible data were analysed in order to verify the presence of the anomaly [72]. The figures in that paper show that there was no anomalous acceleration before a certain date, when it suddenly appeared and remained in the data from both Pioneer 10 and 11 (launched 13 months later). This goes unmentioned in the cited papers [69–72]. Since the mistake, perhaps GR related, turns up in the tracking of at least two space probes and remains there over the years, it may be present also in data that have been used in investigations of the constancy of the AU [65, 66, 67]. This needs to be cleared up in order to judge whether alternative (2) is tenable or stands falsified if the PCP is allowed to be violated.

If the PCP holds and the cosmic web is coherent, statement (1) implies that the material universe is static on all scales, while waves are stretched. If the factor by which waves are stretched per unit of distance is constant and everywhere the same, the redshift factor (1+z) increases exponentially with distance D, as described by Eq. (1). If the number of periods between a source of radiation and the observer is conserved, which is the case in “entirely tired light” models, the expanded distance $D_{\text{exp}}$ can be calculated by integration as

$$D_{\text{exp}} = \exp\left(\frac{H}{c} D\right) - 1. \quad (9)$$

Under this condition, $D_{\text{exp}}$ is simply proportional to z,

$$D_{\text{exp}} = \frac{c}{H} z. \quad (10)$$

9 In [70] one can read: “One can demonstrate that beyond 15 AU the difference between the predictions of Newton and Einstein are negligible”. This is said without telling that it holds only for observations made from a still larger distance – not for those made by us from Earth. This evokes a suspicion that the Pioneer anomaly may have arisen from a similar inadvertence.
This is illustrated in Figure 1. $D_{\text{exp}}$ is a distance that is valid for signals that propagate at $c$. Since equations (9) and (10) hold irrespective of frequency, down to zero, they also hold for the effective lengths of lines of force. It is noteworthy that the unexpanded distance $2D$ from a source to a mirror and back can be measured directly by counting the periods of a stable monochromatic signal that can be sent towards the mirror until the first period of the reflected signal returns. The front of the signal moves at $c$. The expansion occurs from there in the direction opposite to the propagation. Not only waves but any field modulations that propagate or are maintained at $c$ are dilated in this way, whereby slopes and gradients are reduced.

![Figure 1](image)

**Figure 1.** An unexpanded wave train (below) from source S to observer O at distance $D = 0.5$ Hubble length units ($cH^{-1}$) and its expanded equivalent (above). The chosen $D$ gives a redshift $z = 0.649$ and an expanded distance $D_{\text{exp}} = 0.649 cH^{-1} = 1.297 D$.

Statement (1) is not explicit about the reason for the dilatation, but one can easily see a reason for waves to expand if one considers $c$ as the velocity of escape from the universe. Under this premise, anything that moves at $c$ will have to overcome a non-zero gradient: it is pulled back by the gravitational attraction of the universe which, in Section 5.3, will be shown to be finite. This gradient will be disproportionately smaller at $v < c$, when the distant masses no longer pull in the same direction. Such cases remain outside the frame of the present paper, which, in addition to objects moving at $v \ll c$, is only concerned with waves and signals that propagate at exactly $c$. For these, Eqs. (9) and (10), illustrated in Figure 1, are valid immediately.

The two delimitation criteria are contrasted in Table 3. For both, a contraction model has been entered in addition to the equivalent expansion model. One is free to choose one of the two ways of regarding the situation. There is no such freedom if fudge factors are introduced in the way this is done in $\Lambda$CDM cosmology, and the two ways result in equally simple descriptions only if expansions/contractions are exponential functions.

Criterion 1 reflects the practice in BB cosmology, which, as we have seen, appears to be incompatible with an expanding material universe. It is, however, compatible with a cosmology in which the PCP holds and which, unlike CM and GR, offers an explanation for inertia, as detailed in Section 5.3. It appears also to be compatible with the empirical data unless the AU actually increases as it would according to criterion 2. This needs yet to be checked.
Table 3. Delimitation criteria of cosmological models in which expansion is exponential as in Eqs. (1) and (2), flux $F \propto [(1+z) \ln(1+z)]^{-2}$ (5). The entities that contract in contraction models are those that remain unaffected in expansion models.

<table>
<thead>
<tr>
<th>Delimitation criterion</th>
<th>Expansion models</th>
<th>Contraction models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 Coherence</strong></td>
<td>Waves and incoherent objects (radiation, slopes and gradients propagating or maintained at $c$; universe if incoherent)</td>
<td>Coherent objects (atoms, rocks, planets, stars, galaxies, galaxy clusters; cosmic web if coherent)</td>
</tr>
<tr>
<td><strong>2 Gravitation</strong></td>
<td>Anything under free gravitation (waves, universe, cosmic web, galaxies, planetary systems)</td>
<td>Objects under control of non-gravitational force (atoms, rocks, planets, stars)</td>
</tr>
</tbody>
</table>

Criterion 2 is in the gist of GR, in which gravitation is special by being linked to space. It implies an expanding universe in which also planetary systems expand, but the empirical flux–redshift relation (5) suggests the expansion to be exponential, as it is in SEC [57], in which the delimitation criterion has not been made explicit either. In the de Sitter universe [58], any measurements are imaginary since it contains no matter, but this is not much different in FLRW models, which, strictly speaking, do not either contain any of the aggregations of matter, such as atoms, instruments, planets, stars and galaxies, but only an abstract fluid. We shall restrict the further discussion to models in which criterion 1 applies.

Even if the contraction model (1b) is mathematically equivalent to the expansion model (1), there is a heuristic difference between the two. Model 1 suggests the mentioned explanation of the redshift. This is hidden in model 1b, which discloses no direct explanation for the contraction of coherent objects either. On the other hand, 1b makes it obvious that there must be time dilation in proportion to the redshift factor. This is hidden in model 1, which might, mistakenly, be thought to lack overall time dilation.

5.3. Gravitational Potential, Gradients and Inertia

In GR, a static gravitational force and a force due to uniform acceleration of a body have been made equivalent, in accordance with Einstein’s equivalence principle [73], by treating also gravitation as an action of space. One can also reason like this: If the force that acts on a body at rest on Earth is given by a gradient in the gravitational potential field of the Earth, the force that acts similarly on a body that is accelerated must then be given by a gradient in a field that is present in the comoving frame of the accelerated body. In this frame, the rest of the universe is seen as accelerating in the opposite direction and must so give rise to a force in this direction. This force must be counterbalanced in order to accelerate the body. This alternative concretisation of Einstein’s equivalence principle implements Mach’s principle immediately.

It is well known that under ordinary conditions, the inertial force is $F = ma$, so that the effect of the universe must be very close to unity under these conditions. The dynamics of galaxies could then possibly be explained if it could be shown that $F$ is reduced disproportionately for accelerations that are not much larger or even smaller than Milgrom’s $a_0$. 
In the classical theory of fields, the scalar gravitational potential $\Phi$ due to all the masses of the universe can be calculated for any point in spacetime by summing up the contributions from all masses $m$ at their distance $r$ from the point,

$$\Phi = -G \sum m \frac{m}{r}. \quad (11)$$

In a homogeneous non-expanding universe, this $\Phi$ comes out as $\infty$, which leads to absurdities. However, in the suggested cosmology, the effective $r$ in Eq. (11) is not the static distance $D$ but the expanded distance $D_{\text{exp}}$ of Eqs. (9) and (10). With this distance, we get

$$\Phi = -G \sum \frac{mH}{cz}. \quad (12)$$

In a universe in which matter is homogeneously and isotropically distributed in static space, the potential can be calculated by integrating the contributions from shells of thickness $dr$ at distance $r = 0$ to $\infty$:

$$\Phi = -4\pi G \rho \frac{H}{c^2} \int_0^\infty \frac{r^2}{z} dr. \quad (13)$$

The contributions to $\Phi$ by shells up to $r = 8ch^{-1}$ in Eq. (13) are shown in Figure 2 (continuous line). The integrated contributions (from $r = 0$ to $\infty$) are 4.80823 times larger than those calculated for a sphere without expansion and $r_{\text{max}} = ch^{-1}$.

It is commonly known that the ordinary baryonic matter accounts for at most 5% of the critical density of a BB universe, $\rho_c = 3H^2(8\pi G)^{-1}$. If the contributions to the potential of the universe are 4.8 times larger than those calculated for a Hubble sphere, the same density will account for 24% of the potential that is necessary to bring the velocity of escape from the universe to $c$. There remains a discrepancy of 76%, but the suggested approach also predicts a substantial additional energy density to be present, mainly in form of gravitons, as briefly discussed in Section 6.

The isotropic dilatation (a stretching action), which is described by the Hubble acceleration $cH$, has the effect of reducing the slopes and gradients of gravitational fields everywhere. This is a small effect, which becomes dominant only at accelerations that are still smaller than $cH$. The acceleration $a_{\text{red}}$ that is given by a gradient that is reduced by this action can be calculated simply as

$$a_{\text{red}} = \frac{a}{1 + \frac{cH}{a}}. \quad (14)$$

If the inertial force that needs to be overcome in order to impart an acceleration $a$ on a body is entirely due to the acceleration of the rest of the universe in the opposite direction, we might, thus, expect this force to be disproportionately reduced at $a \ll cH$, as described by Eq. (14). Observations tell us that there is such a reduction, but that it becomes dominant below a cut-off acceleration $a_0$ that is still lower than $cH$. This can be understood and explained as follows: If everything but a test body accelerates uniformly in the same direction, the acceleration of a distant mass ‘seen’ by the test body is not $a_{\text{red}}$ but its dilated equivalent, $a_{\text{red}}(1+z)^{-1}$. This suggests that Milgrom’s $a_0$ is an averaged dilated view of the Hubble acceleration $cH$. 
Figure 2. Potentials per unit radius in a homogeneous isotropic universe shown as a function of the radial distance (in Hubble length units $cH^{-1}$) from an observer. Naive contributions to $\Phi$ [as if Eq. (11) was valid] (dotted line) and those in a cosmology in which Eq. (13) is valid (continuous line). These are less negative by the factor $D/D_{\text{exp}}$. The dashed line shows the contributions to the equivalent potential $\Phi_{\text{equ}}$ of Eq. (15), which are less negative by an additional factor of $(1+z)^{-1}$.

In order to calculate $a_0$, we have to find the weighted mean value of the factor $(1+z)^{-1}$ by which accelerations appear reduced from a distance, and to multiply it by $cH$. The weighting must be proportional to the contributions to $\Phi$ by each shell, see Eq. (13) and the continuous line in Figure 2. The dashed line in the figure shows the so weighted, less negative contributions to an equivalent potential

$$\Phi_{\text{equ}} = -4\pi G \rho \frac{H}{c} \int \frac{r^2}{z(1+z)} \, dr.$$  \hspace{1cm} (15)

In Figure 2, $\Phi_{\text{equ}}$ is represented by the area between the abscissa and the dashed line. It is found to be smaller than $\Phi$ (the area between the abscissa and the continuous line, extended to $r = \infty$) by a factor of 0.168093 ($\Phi/\Phi_{\text{equ}} = 5.94910$). The same result is obtained by calculating the mean $r$ of the distribution shown by the dashed line and finding the value of $(1+z)^{-1}$ for this $r$. We get $a_0 = 0.168093 cH$.

Since MoND does not fix the interpolating function between the regimes $a << a_0$ and $a >> a_0$, several such functions have been tried. The empirical data are not clear enough to allow a reliable decision between the two main alternatives, the “standard” and the “simple” interpolating function [24]. These do not give exactly the same optimal value for $a_0$, but the range of $cH/a_0 = 6.1 \pm 1.5$ accommodates both, and the calculated value of 5.94910 falls well within this range. Our considerations single out the simple interpolating function as the correct one. If the inertial force
goes towards $F = ma$ at $a >> \alpha_0$, is given by a gradient in the field seen by an accelerated body, and
slopes are dilated in the way described, the equation for the inertial force becomes

$$F = \frac{ma}{1 + \frac{0.168093}{cH} \cdot a}.$$  \hspace{1cm} (16)

This explains the observed galaxy rotation curves and their successful description by MoND. In
distinction from this phenomenological approach, the present one is deductive and based on a
well-founded alternative to Newtonian Mechanics that remains to be fully disclosed.

Eq. (14) is exact, and Eq. (16) tho the given digits of the constant, in reference systems that are
(1) unaccelerated with respect to the Universe (approx. if $a << cH$), (2) at rest in the Universe (approx.
if $\nu << c$), and (3) locked to a place from which the velocity of escape to infinity is zero (approx. if $v_{\text{esc}}$
$<< c$). A more comprehensive treatment is desirable but exceeds the scope of this paper.

6. Discussion

The statistical properties of a persistent universe remain constant over time. This raises a range
of important questions to which hardly any attention has been paid before, since they do not arise in
a transient universe. These questions concern the recycling of energy via a cosmic ocean of photons
(microwaves), gravitons and neutrinos, which interact with each other in ways that need to be
studied. If energy is conserved, which is a first principle, there must be a cosmic matter and energy
cycle [74, 75]. Since this is not so in a transient universe, the topic is never touched in BB cosmology,
while it is of central importance if the PCP holds.

It is of particular interest that there must be very many neutrinos out there, since once emitted,
it is likely to take many Hubble times until a neutrino will be absorbed again. Neutrinos will,
however, like photons, incur a loss of energy due to the cosmic redshift, and this is substantial
already within one Hubble time. This energy will remain within the cosmic ocean in some form. It
can be expected that the major part of the energy of photons and neutrinos will be transferred to
gravitons and ultimately be absorbed by matter in deep potential wells. If we assume that gravitons
originate with an energy $h\nu$ and $\nu$ is given by astrodynamic periodicities, a transfer of energy from
neutrinos and photons to gravitons, whose original values of $\nu$ are much lower, is to be expected.
Such a process is apt to prevent a heat death of the universe.

The idea that massive neutrinos might contribute substantially to the coherence of galaxy
clusters has been voiced before [27, 26]. In a BB universe with its limited age, this requires the
presence of primordial neutrinos, while the amount of neutrinos produced in known processes is of
course much larger in a persistent universe. Their presence, as well as that of gravitons, is predicted.
The major open question is how these are absorbed again, so that the comparatively high energy
content of the cosmic ocean remains constant over time and the temperature of the CMB at 2.7 K.

Beside opening a range of new questions, a neo-Epicurean approach like this also closes a wide
range of questions, primarily those of cosmogonic and related kind. “The early Universe” can no
longer be a topic. It is embarrassing that there are thousands of papers with this phrase in their title.
The prior state of physical cosmology is depressingly unsatisfactory. It demonstrates repeatedly
the undesirable lock-in effects of path dependence in science. It appears that the whole community
engaged in “concordance cosmology” is blinded by preconceptions. Consider just the data that have
led to invoking “dark energy”. An unprejudiced analysis of these data immediately suggests that the
redshift factor ($1+z$) is a simple exponential function of distance, but this is only told by researchers
who, for whatever reason, are not on the mainstream path. Those on the mainstream path did notice that the discrepant empirical data, but instead of reconsidering their tenets, they attribute the discrepancy to the action of a phantom, and they are only bothered about not knowing anything else about this phantom. This may be due to a lack of awareness of the main point in Section 2 of this paper: it is not scientifically justified to rely on entities and processes that are fictitious within the frame of existing knowledge, i.e., neither observed nor predicted. The confidence these deserve, their explanatory power and their epistemological value are no larger than zero. When empirical data are at variance with model predictions, the challenge needs to be addressed by other means than by invoking fictitious assumptions.

The routine assumption that standards of measurement do not expand in BB cosmologies can also be understood as an instance of path dependence. It reflects the idea of the rigid ruler of CM, which continues to be tacitly relied on in the frame of theories in which such rulers no longer exist.

The drawbacks of path dependence show themselves also in the activity of innovators. Sciaima [56] still treated GR and the expansion of the universe as givens, although the idea he investigated, inertial induction, has consequences that speak against both. Later, among the two alternatives to MoND, modified gravitation and modified inertia, only one was pursued. It was the one that can be realised by keeping CM or GR and adding some new fields to them ad hoc [29], not the one that would call the foundation of both theories into question (modified inertia).

In Section 5.3 it was shown that inertial forces are reduced disproportionately at low accelerations if they are due to the relative acceleration of the rest of the universe. While this needs to be reflected on and elaborated more deeply, it requires no ad hoc assumption. It requires just allowing the cosmic redshift to affect gravitational and electromagnetic fields and signals in the same way. If such a generalisation works, it is clear that fundamental progress has been made. Therefore it is particularly refreshing that the present approach leads to a numerical value of $cH/a0$ that lies within the range suggested by empirical analyses of galaxy rotation curves [24]. It is capable of explaining these, while MoND just describes them.

The confidence criteria derived in Section 2 apply to the tenets as such and to their yield, provided that the theory is sufficiently elaborated and that no errors (conceptual, logical, mathematical) are present. In BB cosmology, such an error has been identified in the vacant delimitation between the expanding and the non-expanding domain. The tenets our non-speculative cosmology explicitly relies on are very solid. It remains to be scrutinised whether any unwarranted implicit assumptions or errors have crept into the reasoning and whether the approach also holds for other observables than those considered here. While it emerged as an alternative to BB cosmology, it involves a Machian alternative to GR that still needs to be elaborated or perhaps merely identified among all the prior proposals that were ignored because of their unorthodoxy.

The original aim of this paper was to disclose the nature of the aberrations from the path to knowledge which tend to be conserved in what Kuhn [4] called “normal science” and of which Concordance Cosmology provides a variegated range of illuminating examples. The analysis of these may be inspirational in the evaluation of theories and models in other fields, where analogous aberrations may have occurred. Here, it was not intended to present an elaborate more well-founded cosmological theory, but still to suggest an alternative path along which such a theory may possibly be arrived at. This path leads to a physical cosmology that is more in accord with the ancient world view of Epicurus (ca. 341–270 BC), Lucretius (ca. 99–55 BC) and Giordano Bruno (1548–1600) than
with the one arrived at by proceeding on the path indicated by Newton and Einstein even when
Nature tells us that this is the wrong path.

Acknowledgments

References


