Was Newtonian Cosmology Really Inconsistent?[1]

Peter Vickers

Centre for History and Philosophy of Science, Department of Philosophy, University of Leeds, UK

ABSTRACT

This paper follows up a debate as to the consistency of Newtonian cosmology. Whereas Malament (1995) has shown that Newtonian cosmology *is* not inconsistent, to date there has been no analysis of Norton's claim (1995) that Newtonian cosmology *was* inconsistent prior to certain advances in the 1930s, and in particular prior to Seeliger's seminal paper of 1895. In this paper I agree that there are assumptions, Newtonian and cosmological in character, and relevant to the real history of science, which are inconsistent. But there are some important corrections to make to Norton's account. Here I display for the first time the inconsistencies—four in total—in all their detail. Although this extra detail shows there to be several different inconsistencies, it also goes some way towards explaining why they went unnoticed for two hundred years.

---

1. Introduction

There is now a substantial literature devoted to inconsistencies in science, with examples ranging from the early calculus of Newton and Leibniz to Bohr's

[1] Published as Vickers (2009).

theory of the atom to, most recently, classical electrodynamics. Norton (2002) introduces two different approaches to inconsistencies: the 'content driven' approach and the 'logic driven' approach. Preferring the latter are several authors (see Meheus 2002) who suggest that, when faced with an inconsistency in a given body of assumptions, scientists either *do* (descriptive claim) or *should* (normative claim) adopt a non-classical, paraconsistent logic which saves them from deriving anything and everything by *ECQ*.[2] Norton, in response, suggests a 'content driven' approach where the inconsistency *is* or *should be* handled by '[reflecting] on the specific content of the physical theory at hand', whilst maintaining classical deductive logic (2002, p.192).

Both sides of this debate are concerned to address the question of what scientists do or should do when faced with inconsistency. This is of clear importance, not least because it could give us important clues as to how we might progress in the face of *current* conflicts, such as that between general relativity and quantum field theory. However, the focus of this paper lies entirely outside of this debate, and addresses a different aspect of inconsistency in science which has been largely neglected. Just as important as the pragmatic question of *what to do* when faced with an inconsistency is the epistemic question of how scientists *come to know* about inconsistencies in a given body of assumptions in the first place. In particular we may ask the questions,

(i) What is it about the *scientific community* which prevents an inconsistency from being noticed?
(ii) What is it about the *science* which prevents an inconsistency from being noticed?

There are far fewer papers dedicated to these kinds of question, but research here could also carry importance for current science: if we could accelerate the identification of conflicts in science we could accelerate science itself.

The focus of attention in this paper will be what is usually called (an old version of) 'Newtonian cosmology'. In 1895 Seeliger made the remarkable claim that a set of natural (Newtonian) assumptions concerning forces in the universe—assumptions which had been in place since Newton himself, for over two hundred years—are mutually inconsistent. This stimulated much debate over the years and decades which followed, with the latest additions made by Norton (1993, 1995, 1999, 2002) and Malament (1995). But questions (i) and (ii), above, remain largely unasked. What was it about the *nature* of the inconsistency in Newtonian cosmology which meant that it went unnoticed by

[2] *Ex contradictione quodlibet*: From the inconsistent assumptions derive a contradiction 'A&~A'. From here infer 'A' and '~A'. From 'A' infer 'AvB' for *any* arbitrary 'B'. From '~A' and 'AvB' infer 'B'.

the scientific community for two centuries?

If the derivations were long and complicated, and involved advanced mathematics or even mathematics not available to scientists in the relevant period, then our question would be answered immediately: the reason the contradiction remained hidden would be that it was exceedingly difficult to derive. But one look at Norton's reconstructions shows that no such quick answer will be forthcoming. For one thing, there are at least *two different* inconsistencies, so there is double the chance of noticing the problem. But even more remarkably, it would seem that in each case a contradiction follows from a few basic assumptions in a few simple steps. The inconsistencies are, as Malament puts it, 'so close to the surface that they are hard to miss' (1995, p.489). This in itself seems to contradict the fact that many great scientists *did* miss the inconsistencies for a period of two hundred years! Otherwise we would apparently have to admit either that scientists made a serious commitment to what they knew to be impossible, or that they were blind to some of the most obvious consequences of their beliefs.

This suggests that the reason the inconsistency remained hidden lies with the scientific community rather than the science itself. However, as this paper sets out to show, there are several complications to work through to understand the inconsistencies properly: they are not as simple and straightforward as Norton and Malament make out. Thus, after a brief section (§2) in which I introduce 'Newtonian cosmology' and the relevant assumptions, I turn to the details of the inconsistency claims which have been made. Four different inconsistencies are distinguished, which are grouped into two *types* of inconsistency discussed separately in §3.1 and §3.2. This analysis uncovers certain complications in the science which then help to answer question (ii) in §4. Some factors pertinent to question (i) are also brought to light. §5 is the conclusion.

2. The concept *Newtonian cosmology*

What is the theory of 'Newtonian cosmology'? It would be a mistake to suppose that Newtonian cosmology 'exists' somehow, perfectly formed, in a textbook somewhere, and that all we have to do is identify it and see if it is inconsistent. Rather, what we really find are various assumptions, often differently stated by different individuals, sometimes saying the same thing and sometimes something slightly different. Some of these assumptions are clearly about the universe as a whole, others about the local universe, others of a more metaphysical nature. Still other assumptions may never be articulated, such as that space is Euclidean or that the night sky is dark. There is often no correct way to articulate such assumptions precisely, and no specific set of them which together constitute 'Newtonian cosmology'. What, then, do we mean when we refer to ‘Newtonian cosmology’? What do we mean when we say it is inconsistent?

People mean different things by the word ‘theory’. The question is not who is right but which conception of ‘theory’ is the right one to use for a given investigation.[3] The present concern is to investigate the inconsistency of Newtonian cosmology, but still this is too vague to tell us which assumptions we should be concerned with: we must ask what we are hoping to *show* by the inconsistency of the theory. For example, we will choose different assumptions, and aim at something quite different, if we want to show that the assumptions *used* by scientists were inconsistent (but that the inconsistency somehow didn’t affect their reasoning) or if we want to show that the assumptions *believed* by scientists were inconsistent (but that they somehow didn’t notice it). The more interesting claim for present purposes is the latter, because what seems to be the case is that there is a set of assumptions which are inconsistent, all of which were believed to be true—or at least important *candidates* for the truth—by a significant number of relevant individuals. However, focusing on belief in this way still doesn’t enable us to identify a set of assumptions for investigation. There are many different species of ‘belief’, and even if there were not it isn’t clear which beliefs should be grouped together to count as ‘Newtonian cosmology’.

Thankfully there is another way to proceed. Instead of trying to group together the assumptions for which inconsistency would be interesting, one can instead group together the assumptions which *are* inconsistent, and then investigate how interesting that inconsistency is. Of course *some* sense of which assumptions are going to be important is required: there has to be some reason why the assumptions will be interesting *as a group*. The reason here, as made clear in Norton’s papers, is the role they play in answering a single question:

> (Q) What is the net gravitational force on a test particle at a given time at an arbitrary place in the universe?

When this question is asked certain assumptions are naturally drawn together. Norton (1995) introduces various such assumptions as follows:[4]

(a) Newton’s three laws of motion.

---

[3] Cf. Vickers (2008), where I analyse the disagreements between Mathias Frisch, Fred Muller and Gordon Belot concerning the consistency of classical electrodynamics in terms of their differing conceptions of ‘theory’. See also Wilson (2009), where three different conceptions of ‘Classical Mechanics’ are discussed.

[4] See pp.513-514. In labelling the assumptions I follow Norton’s lead for ease of cross-reference.

(b) Newton’s inverse-square law of gravitational attraction.

(b') Poisson’s equation with gravitational attraction described in terms of the potential φ.

(c) Matter in the universe is distributed homogeneously (when viewed on a large enough scale) in an infinite Euclidean space.

(d) There is a determinate net gravitational force on a test mass at any given time.

(d') The gravitational potential φ is homogeneous.

It turns out that the question (Q) can be answered in different, contradictory ways depending on which of the given assumptions are emphasised. There are essentially four different methods of reasoning which can be employed, which will be introduced in the forthcoming analysis in the following order:

1. Use Newton’s law of gravitation
2. Use Poisson’s equation
3. Use symmetry considerations
4. Use the gravitational potential

Proceeding in this way means that we can investigate the inconsistencies without getting into the messy meta-ontology which accompanies such questions as ‘What is the theory?’ and ‘What is Newtonian cosmology?’ Asking such questions presumes a simplicity to the history of science which doesn’t exist.[5] Instead the analysis can proceed in terms of the given question (Q), various methods of answering that question, and the assumptions which those methods draw upon. Any use of terms such as ‘theory’ and ‘Newtonian cosmology’ in what follows should be taken as shorthand for an analysis in terms of sets of assumptions. The connection between the given assumptions and the real history of science will be considered in §4.

3. How was Newtonian cosmology inconsistent?

Inconsistency, of course, means that a contradiction follows. When the noted

[5] Equally there should be no question of whether we are really discussing ‘Newtonian cosmology’ (Norton 1993) or ‘Newtonian gravitation theory’ (Norton 2002).

methods of reasoning provide contradictory answers to our question (Q) one of two principle kinds of contradiction results:

(i) The net force on a test particle at a given time is both **F** and **G** where **F**≠**G**, or

(ii) The net force on a test particle at a given time is both *determinate* (some vector quantity) and *in*determinate (in a sense to be clarified).

These will be tackled in §3.1 and §3.2 respectively. The precise inconsistency then depends on which of the given assumptions (a)-(d') the contradiction is said to follow from. Thus §3.1 and §3.2 are each split into further subsections.

3.1 A contradiction of forces

By 'contradiction of forces' I simply mean that the given assumptions can be used to derive the following contradiction:

(C1) The force on a test mass is **F** and the force on a test mass is **G**, where **F**≠**G**.

This splits into three different claims depending on which of the assumptions (a)-(d') are used to make the derivation.

3.1.1 …using Newton's law of gravitation?

Norton's original paper (1993) shows the first possible method of reasoning, and claims that we have a contradiction of forces from assumptions (a), (b) and (c):

(a) Newton's three laws of motion.

(b) Newton's law of gravitational attraction.

(c) Matter in the universe is distributed homogeneously (when viewed on a large enough scale) in an infinite Euclidean space.

In greater detail, by (b) we mean:

(b) The force of gravity $\mathbf{F}_i$ on a test body m_t at $\mathbf{r}$ due to another body m_i at

$\mathbf{r}_i$ is given by $\mathbf{F}_i = G\dfrac{m_t m_i}{|\mathbf{r}_i - \mathbf{r}|^3}(\mathbf{r}_i - \mathbf{r})$.

This gives us the magnitude and direction of the force on our test mass m_t due to *one* other mass m_i. But in this paper we are asking what the *net* gravitational force is. By (c), since we are supposing the universe to be infinite and the mass distribution to be homogeneous, there will be an infinite number of masses. Further, (b) comes with no caveat that it doesn't hold beyond a certain distance $|\mathbf{r}_i - \mathbf{r}|$. Thus we must infer that *every* mass in the universe has some effect (however small) on our test mass. Thus there are an infinite number of terms in our sum, and the net gravitational force is represented as the sum-total of all the contributory forces: $\mathbf{F}_{net} = \sum_i \mathbf{F}_i$.[6]

At this point, since we are drawing on assumption (c), it is worth pausing to consider what is really meant by the homogeneity of the universe *when viewed on a large enough scale*. Of course nobody ever believed that the universe is *totally* homogeneous, but rather that if you take any arbitrary region of space *R* of a given large volume *V* then you will always find the same total amount of mass there (with small deviations from some mean value, which get smaller for volumes of space larger than *V*). This actually tells us next to nothing about the density of matter in the vicinity of a given test particle (it could be sat on the surface of a black hole, or be several hundred light years from the nearest massive particle). All we know is the *total* amount of matter in an arbitrary region of space *R* of volume *V*, which may include our particle. But this uncertainty in the local matter distribution doesn't transfer to an uncertainty in the force on such a particle, at least insofar as Norton's 'lines of force argument' (1993; 2002) is concerned. All the argument requires is the constant density of matter over different regions *R* at *some* scale (however big *V* has to be to achieve this constant density).

Continuing Norton's argument, from the infinite sum we can apparently get different answers depending on how we compute it.[7] If we first consider a spherical region of the universe upon which our particle is sitting—of any given volume *V* or greater, and situated on any side of our particle—then we get a force **F** towards the centre of that sphere. It can then apparently be shown that the force due to *all other* masses amounts to nothing, since they can be

[6] For some purposes it is convenient to turn this sum into an integral according to $\sum_i \mathbf{F}_i = \int_V \mathbf{F} dV$, where instead of summing over all (discontinuous) masses we sum over all (continuous) points in Euclidean space, as in Malament (1995, p.491). This won't be important here.

[7] For the full argument see Norton 1993 and 2002.

grouped into spherical shells, concentric with the centre of the original spherical region, each of which has no net effect on our particle. Thus our infinite sum turns into $\mathbf{F}_{net} = \mathbf{F} + \mathbf{0} + \mathbf{0} + \mathbf{0} + ... = \mathbf{F}$. But the size and direction of the original sphere, and thus the force **F**, was completely arbitrary. Thus Norton claims that the theory is ‘logically inconsistent in the traditional strict sense… [because] we can prove within the theory that the force on a test mass is both some nominated **F** and also not **F**, but some other force.’ (1993, p.413; 2002, p.186).

In fact no such contradiction can be legitimately derived. Malament (1995) criticised the reasoning as follows:

> What Norton presents as an argument for inconsistency is better understood as just a vivid demonstration of non-convergence. (A perfect analogue of his argument might be used to “prove” that, for every integer n, the infinite sum 1 – 1 + 1 – 1 + … is equal to n.) … Rather than asserting that Newtonian theory makes inconsistent determinations of gravitational force … Norton should have asserted that it makes no determination *at all*. (1995, p.491, original emphasis)

In other words, Norton should have noted that not all infinite sums have an answer. For example, mathematicians in the late 17th and 18th centuries rigorously debated whether the infinite sum 1-1+1-1+…, known as ‘Grandi’s series’, is equal to 1, 0, ½, or something else. Today, with the benefit of hindsight, we can look upon these struggles as mere historical curiosities, and say instead that since the series in question is not convergent it has no sum. The question is mathematically well-posed but has no answer, just as a question can be grammatically well posed and have no answer.

The key failure of Norton’s argument can be seen by the fact that he groups together the effects of masses in certain regions of space in order to get a result. This is equivalent to *bracketing* together terms in Grandi’s series in order to get a result. But since this bracketing can be done in more than one way, if this were legitimate you could also show Grandi’s series to sum to two different values. The most obvious two are as follows:

$$1-1+1-1+... = (1-1)+(1-1)+... = 0+0+... = 0$$
$$1-1+1-1+... = 1-(1-1)-(1-1)-... = 1-0-0-... = 1$$

So does Grandi’s series sum to *both* 0 and 1? Since this sort of bracketing is mathematically illegitimate, one must stop at the unbracketed series and conclude that it is equal to *no* value.

These facts take on particular significance in the light of Norton’s 1999 derivation. He argues (p.274) that the infinite sum can be written as follows:

$$\begin{aligned}\mathbf{F}_{net} &= G\pi\rho\Delta r\hat{\mathbf{x}} - G\pi\rho\Delta r\hat{\mathbf{x}} + G\pi\rho\Delta r\hat{\mathbf{x}} - G\pi\rho\Delta r\hat{\mathbf{x}} + G\pi\rho\Delta r\hat{\mathbf{x}} - \ldots \\ &= G\pi\rho\Delta r\hat{\mathbf{x}}(1-1+1-1+1-\ldots)\end{aligned}$$

where $\hat{\mathbf{x}}$ is a unit vector in *any* nominated direction. This time he considers *hemi*spherical concentric shells, first on one side of the particle (in the direction of $\hat{\mathbf{x}}$) and then the other (in the direction of $-\hat{\mathbf{x}}$), which build up to infinity. But the outstanding question is: why is *this* particular style of summation legitimate and the others not so? Hasn't Norton once again illegitimately grouped together (in hemispheres) the effects of large numbers of masses to achieve his sum? Hasn't he introduced brackets into the reasoning?

This is true to some extent. The *real* infinite sum, without brackets, is a close approximation to a 3D version of Grandi's series.[8] Strictly speaking Norton shouldn't group together the effects of large numbers of masses in hemispherical shells as he does in his 1999 paper. Crucially, however, this time the grouping does not affect the divergence of the summation. The introduction of his brackets is analogous to the following manipulation of Grandi's series:

$$1-1+1-1+\ldots = (1-1+1)-(1-1+1)+(1-1+1)-\ldots = 1-1+1-1+\ldots$$

In other words, the divergence of the sum is preserved by the bracketing. This is the achievement of Norton (1999). Thus although strictly speaking the infinite sum should *not* be set equal to **k**(1-1+1-1+…)—where **k** is the relevant constant vector—the indeterminate nature of the force $\mathbf{F}_{\text{net}}$ is preserved by Norton's 1999 analysis, whereas it *isn't* preserved in his 1993 and 2002 analyses. With Norton's 1999 analysis we can be sure that the original series *is* divergent, because the re-ordering and bracketing of a convergent series would never leave us with a non-convergent series.

A further point is in order here. Malament says that what Norton presents is a vivid demonstration of non-convergence. However, more specifically what Norton presents is a vivid demonstration of *alternating* non-convergence. Only when signs in a divergent series alternate is it possible to make quantities cancel out, and achieve various different finite answers through bracketing and re-ordering. Thus there is a sense in which alternating series *don't* diverge, since 'diverge' usually means 'diverge to infinity'. However you bracket and re-order an infinite series which diverges to infinity you get infinity (whether

[8] In fact the *real* infinite sum could never be presented. Grouping of some sort *has* to be introduced, because in order to consider the universe as homogeneous we have already grouped together large portions of it.

positive or negative). Only with alternating series is it possible to achieve any number of finite answers for the sum of the series. To make this distinction clear, in what follows the word 'indeterminate' will be preferred over 'divergent' to describe the sum of Grandi's series.

With this clarified we can still agree with Malament that assumptions (a), (b) and (c) make no determination of the net force. But this isn't because the force is *divergent*, in the sense of 'diverges to infinity'. If it were divergent then we could say that the assumptions predict an *infinite* force. Rather, it is because we reach a force balanced between convergence and divergence, an *indeterminate* force. Norton apparently accepts Malament's criticism, since he is moved in his reply (1995) to add a further assumption to (a), (b) and (c), and introduce a somewhat different contradiction, an 'indeterminacy contradiction', as we'll see in §3.2.

3.1.2 …using Poisson's equation?

A second method of reasoning involves Poisson's equation. Norton never suggests that we get a contradiction of forces using Poisson's equation directly, but it will be useful later on to consider precisely why the theory isn't inconsistent in this way. Fleshing out assumption (b') we have:

(b') The force of gravity **F** on a body m_t at **r** due to the mass distribution in a given volume V is given by $\mathbf{F}(\mathbf{r}) = -m_t \nabla \phi(\mathbf{r})$, where $\varphi(\mathbf{r})$ is such that $\nabla^2 \phi(\mathbf{r}) = 4\pi G \rho(\mathbf{r})$, where G is a constant and where the gravitational potential $\varphi(\mathbf{r})$ and the mass density $\rho(\mathbf{r})$ are continuous scalar fields on V.

To find the net force on a test mass we no longer need to sum up all the individual forces, but can simply derive the force from the potential field $\varphi(\mathbf{r})$. From assumption (c) ρ is constant in space, so instead of '$\rho(\mathbf{r})$' we can simply write 'ρ', with the added proviso (as we saw in the last section) that the density is only constant for regions of space of a given volume *V* or greater. Poisson's equation then becomes $\nabla^2 \phi(\mathbf{r}) = 4\pi G \rho$, where $\varphi(\mathbf{r})$ now refers to the gravitational potential at 'points' of space **r** which actually pick out regions *R* of volume *V* or greater. Since Poisson's equation is a differential equation we need to integrate, and when you integrate you inevitably incur constants of integration. Thus the so-called 'canonical solutions' of Poisson's equation are,

$$\phi(\mathbf{r}) = \frac{2}{3}\pi G \rho \left|\mathbf{r} - \mathbf{r}_0\right|^2 .$$

Here $\mathbf{r}_0$ is the constant of integration.[9] If we differentiate this equation to test whether it satisfies Poisson's equation we get the right result $4\pi G\rho$ *whatever* $\mathbf{r}_o$ is, because it simply disappears during the calculation.

We can now move to the force on our test mass using $\mathbf{F}(\mathbf{r}) = -m_t \nabla \phi(\mathbf{r})$. We find,

$$\mathbf{F}(\mathbf{r}) = m_t \frac{4}{3} \pi G \rho (\mathbf{r_0} - \mathbf{r}).$$

Once again, it should be emphasised that this really means the average force **F** in a large region of space *R*—picked out by **r**—of volume *V* or greater. We get a different value for this average force depending on how we choose $\mathbf{r}_0$, so we can get two different, contradictory average forces in the same region *R* by choosing $\mathbf{r}_0$ in two different ways. But it would be a gross error to go on to suppose that the theory is inconsistent for *this* reason. The theory *just doesn't tell us* what $\mathbf{r}_0$ is, so we must leave it as an unknown constant. We certainly can't just arbitrarily choose it to be two *different* things. To emphasise that the theory leaves us guessing we could write **F**=?, because $\mathbf{r_0}$=? (this will be useful for comparison later).

3.1.3 …reasoning from symmetry

Given the tools at our disposal we've thus-far seen two different ways of reasoning when faced with the question, 'What is the net force on a given test particle?' We can use Newton's inverse-square law of gravitation or we can use Poisson's equation. A third possible method of reasoning, and perhaps the most obvious (particularly to non-scientists), is to use symmetry considerations.

There is an intuition that if the universe is really infinite and Euclidean, and has a homogeneous mass distribution (with the qualifications noted above), then it will be exactly the same vis-à-vis the average force on a test mass in any given region *R*.[10] In other words, it is assumed that the average force *cannot differ* for any two such regions, since they are identical in the relevant respects. Of course this can't follow from the cosmological assumptions (c) alone, since

[9] Norton (1995) calls these the 'canonical solutions', and distinguishes them from the most general solutions which include another constant of integration. This constant is quickly eliminated (see Norton 1995, p.513), and at any rate does not affect any argument here (see fn.12, below).

[10] Focusing on regions rather than points of space obviates the need to account for local variations in the matter distribution, and allows us to compare the results of this method of reasoning directly with that which draws on Poisson's equation, as seen below.

no reference is made to *force* there. From (c) it only follows that every 'point' of the universe is identical vis-à-vis the surrounding *mass*. So we need to add another assumption to change the intuition into a valid claim that every 'point' in the universe is identical vis-à-vis the gravitational force.

The following assumption will do:

(e) Gravitational force is caused by *all* mass and *only* mass.

If gravitational force is caused by *all* and *only* mass, then the fact that every 'point' in the universe is identical vis-à-vis the surrounding mass distribution will mean that every 'point' in the universe is identical vis-à-vis those factors relevant to the gravitational force. Which will mean that the average gravitational force is identical at every 'point' in the universe. It is often assumed that this means that the force must be everywhere zero. What reason, the argument goes, could there be for the force to point in one direction rather than another? This really needs an additional 'no preferred direction' assumption, which can be ignored for present purposes. All that is required to reach contradiction is that the force, whatever it is, doesn't differ from 'point' to 'point'. Following the discussion in §3.1.1, it should be noted that this is also consistent with the force being everywhere indeterminate.[11]

We can now finally achieve a contradiction of forces by comparing this method of reasoning with the one seen in the previous section. From symmetry, drawing on assumptions (c) and (e), we have inferred either that **F**=**k** at every 'point' **r** of the universe (for some $\mathbf{k} \in \Re^3$) or that **F** is everywhere indeterminate. Either way **F** will not differ from 'point' to 'point'. But we saw in the previous section that we can draw on assumptions (b') and (c) to conclude that the average force at a given 'point' **r** will be

$$\mathbf{F}(\mathbf{r}) = m_t \frac{4}{3}\pi G\rho(\mathbf{r_0} - \mathbf{r}) .$$

In order to satisfy Poisson's equation $\nabla^2\phi(\mathbf{r}) = 4\pi G\rho$, $\mathbf{r}_0$ *must* be some real vector quantity: $\mathbf{r}_0 \in \Re^3$. But *whatever* vector quantity we choose we find that the force **F** *will* differ from one 'point' **r** to another **r'**. In fact whatever the choice of $\mathbf{r}_0$ we find that the average force on a test mass will be **k**, for any

---

[11] Malament (1995, pp.493 and 509) argues that a homogeneous mass distribution does *not* entail a homogeneous force field. However, this is only when one takes "gravitational force" to be a gauge quantity with no 'direct physical significance' (as Malament puts it). But before 1900 "gravitational force" certainly was presumed to have physical significance, so the entailment holds true for the purposes of this paper.

given **k**, in exactly one region in the universe. And the *difference* between regions increases as the *distance* between the regions increases. Using this fact we reach the following contradiction, resulting from reasoning from (b'), (c) and (e) in two different ways:

> (C2) The average force on a test mass in any two arbitrary, widely spaced regions of the universe *R* and *R'* (of volume *V* or greater) will not differ (or differ by a negligible amount), and the average force on a test mass in any two arbitrary, widely spaced regions will differ significantly.[12]

Although this *is* a contradiction, it is not immediately obvious what it means in empirical terms. We will not find the 'force on a test particle' **F** being, impossibly, both two different things at a single time (in other words we don't get contradiction (C1)). Whereas the 'force on a test particle' is something we might say 'exists', the *average* force in a given region of space is a non-existent abstraction, just as the average family (with 2.4 children) is an abstraction. The real empirical difference here lies with the large scale movements of matter over time: with one story there are no large scale movements, whilst with the other there will be a large scale acceleration towards the 'point' $\mathbf{r_0}$.

This is closely related to Norton's so-called 'inhomogeneity contradiction', as introduced in his 1995 paper (p.514). However, instead of a contradiction of forces he introduces the contradiction,

> (C3) The gravitational potential φ is homogeneous and *it is not the case that* the gravitational potential φ is homogeneous.

This is only achieved by introducing a new assumption:

(d') The gravitational potential φ is homogeneous.

Unlike the other assumptions used so far, commitment to (d') by scientists is dubious. φ isn't a physical thing after all, but is just a mathematical tool which

---

[12] Einstein provided a possible way forward here. He found that he could make things work by altering Poisson's equation to $\nabla^2\phi - \lambda\phi = 4\pi G\rho$, so that a constant solution for φ is possible: $\varphi = -4\pi G\rho/\lambda$ (he found that no such solution is possible with the normal version of the equation, even if one draws on the most general solutions—see fn.9, above). This then gives an average force of **F**=**0** everywhere, consistent with symmetry considerations. See Norton (1999) for this and other ways in which the assumptions can be modified to avoid the problems.

intermediates between the 'real' masses and forces.[13] The suggestion seems to be that (d') follows from the homogeneity of the mass distribution, but this is to mix up the physical and the mathematical. The argument is surely that since the universe's mass distribution is symmetrical the universe cannot differ from region to region in certain physical respects. The latter include the (physical) force, but not the (unphysical) potential.

Here it is better to turn to Malament (1995, p.492), who frames the difficulty in terms of the homogeneity of the *force* field. Thus a slightly different contradiction is suggested:

> (C4) The gravitational field $\mathbf{f}$ is homogeneous and *it is not the case that* the gravitational field $\mathbf{f}$ is homogeneous.

where $\mathbf{f}$ stands for $\mathbf{f}(\mathbf{r})$, the force per unit mass at $\mathbf{r}$. This is closely related to (C2), but differs from it in two respects which are worth noting. First, Malament has left out the fact that $\mathbf{f}$ must stand for the *average* value of the force field in large regions (thus (C4) makes things seem simpler than they really are). However, this isn't really a difference since Malament has simply left it implicit. A more substantial difference comes in the fact that Malament does not allow for the fact that a force everywhere indeterminate is compatible with symmetry constraints (unless, somewhat implausibly, he intends this possibility to be covered by the word 'homogeneous'). In other words one can only reach Malament's contradiction (C4) by drawing on an extra assumption, assumption (d), which blocks the possible indeterminacy of the force. But the present analysis shows that drawing on this extra assumption is not necessary to reach contradiction. Leaving assumption (d) aside, from symmetry it follows that *if* such a force field exists *then* it is homogeneous. Poisson's equation then brings contradiction by telling us that such a force field *does* exist, and that it is *in*homogeneous, which affirms the antecedent and denies the consequent of our conditional.

Summing up, what I have shown is that we do get a genuine contradiction of forces (C2) from (b'), (c) and (e). In fact we might even say that we reach the contradiction by (b') and (c) alone, since it might be argued that (e) is embedded within (b'). However, this doesn't affect the main point here: the challenge isn't to find inconsistency in as few assumptions as possible, but to find inconsistency in assumptions which probably were committed to, in the relevant historical period, as serious candidates for the truth. There can be little doubt that assumptions (b'), (c) and (e) meet this criterion (see §4 for more on

[13] The potential was introduced by Laplace in the 1770s, and was considered a mere computational tool from the very beginning (see Cat 2001, p.402ff. and Grattan-Guinness 1990, p.332).

this point).

3.2 An indeterminacy contradiction

The contradiction of concern in this section will be,

> (C5) There is a unique gravitational force on a test mass and *it's not the case that* there is a unique gravitational force on a test mass.

This isn't quite the contradiction Norton presents in his 1995 paper, but it is surely what he *means* to present. It is worth pausing to clarify things here, since he doesn't correct his mistake in his 1999 and 2002 papers.

Norton responds to Malament's objections (as seen above in §3.1.1) in his 1995 paper. He accepts that there is no contradiction of forces after all, and instead brings to our attention what he calls an 'indeterminacy contradiction'. On p.513 he adds the following assumption to (a), (b) and (c):

(d*) There is a unique gravitational force on a test mass fixed by (b) and (c).

Now, since Malament is right about the non-convergence of the sum, one *cannot* derive a unique gravitational force on a test mass from (b) and (c). Thus it might be supposed that we have a contradiction here:

> (C6) There is a unique gravitational force on a test mass fixed by (b) and (c) and *it's not the case that* there is a unique gravitational force on a test mass fixed by (b) and (c).

But on closer inspection we don't have this contradiction after all. If we accept Norton's (d*) we have,

(a) Newton's three laws of motion.
(b) Newton's inverse square law of gravitation.
(c) Matter is distributed homogeneously and isotropically (when viewed on a large enough scale) in an infinite Euclidean space.
(d*) There is a unique gravitational force on a test mass fixed by (b) and (c).

(d*) gives us the positive contradictory of (C6), so it only matters that we can derive the negative contradictory. However, even if it follows from (b) and (c) that, *It's not the case that* there is a unique gravitational force on a test mass.

this *isn't* the contradictory we want. To establish (C6) we need to add '…fixed by (b) and (c)' on the end. But since the assumptions in question don't refer to '(b)' and '(c)' at all this is an impossible task. (C6) can't be derived from (a)-(d*) after all.

It is clear what has happened here. In specifying (d*) Norton has accidentally mixed up the theory and the meta-theory. He actually meant to add,

(d) There is a unique gravitational force on a test mass

which leads to contradiction (C5), as we will see in the next section.

3.2.1 …using Newton's law of gravitation

We saw in §3.1.1 that, as Malament claims, assumptions (a), (b) and (c) tell us that the net force on a given test mass is undetermined. But (d) tells us that the net force on a test mass *is* determined. And the introduction of (d) should not be dismissed as the ad hoc introduction of the required contradictory. In fact, the introduction of (d) is merely the explicit mention of an assumption which is already an integral part of Newton's three laws (a). Take Newton's first law, for example. In its original form it states, 'Every body perseveres in its state of being at rest or of moving uniformly straight forward, except as it is compelled to change its state by force impressed.' This is equivalent to 'a body is either at rest or moving in a straight line, or accelerating due to an impressed force'. These are the only options, so a body is either experiencing a force (accelerating) or it is not (straight line motion, or rest). In other words, there is always a determinate force on a body, whether it be something or nothing. Thus the introduction of (d) is not the introduction of a new assumption at all, but is part and parcel of (a). Thus the indeterminacy contradiction (C5) is meant to follow from (a), (b) and (c).

There is an important distinction to make here. Certainly from (b) and (c) we end up with an infinite sum which is indeterminate, from which we cannot achieve an answer to the question 'what is the force?' But can we conclude from here that there is *no* unique gravitational force on a test mass? That is, we have failed using (b) and (c) to determine what the force is. But couldn't it still be the case that there is *some* unique force, and that we could determine what it is by another method, using different reasoning or bringing in other considerations?[14] This is crucial, because if we cannot make the further assumption that *no force exists* then we cannot get to the negative contradictory in question, and the inconsistency claim falls down.

The two possible inferences can be distinguished as follows:

(I) from indeterminacy infer that no solution has been reached.
(I*) from indeterminacy infer that there *is* no solution.

If one makes the weaker inference (I) then the reasoning continues as follows. The fact that we can't figure out what the force is from (b) and (c) can be represented by a question mark (cf. §3.1.2, above):

$$\underbrace{\sum_i \mathbf{F}_i = ?}_{(b),(c),(I)}$$

On this understanding, *all options are still open*—there might be some other way to determine what the unique force on a test mass is. There is then no contradiction with (d). We might formalise (d) thus:

$$\underbrace{\exists \mathbf{k} \in \Re^3, \mathbf{k} = \mathbf{F}_{net}}_{(d)}.$$

In other words, there exists some vector quantity which equals the net force on a test particle.[15] If we believe that (b) Newton's law of gravity holds then we can add,

$$\underbrace{\mathbf{F}_{net} = \sum_i \mathbf{F}_i}_{(b)}.$$

Placing these beside each other we have,

$$\underbrace{\exists \mathbf{k} \in \Re^3, \mathbf{k} = \mathbf{F}_{net}}_{(d)} \quad \underbrace{\mathbf{F}_{net} = \sum_i \mathbf{F}_i}_{(b)} \quad \underbrace{\sum_i \mathbf{F}_i = ?}_{(b),(c),(I)}$$

Now, by substitutivity of identicals, we can fill in the question mark and write:

[14] This is suggested by Norton's introduction of (d*). (d*) says that there is no force *fixed by (b) and (c)*, not that there is no force.
[15] Of course before the introduction of 'real numbers' and the like physicists would have talked vaguely about 'quantities', but that doesn't affect the argument at hand.

$$\underbrace{\exists \mathbf{k} \in \Re^3, \mathbf{k} = \sum_i \mathbf{F}_i}_{(b),(c),(d)}$$

There is no contradiction here. We couldn't find an answer to our indeterminate sum using (b) and (c), but there is *some* answer, yet to be discovered.

Today, two centuries of mathematics tells us that this is the wrong way to think about indeterminate sums. Not only do we *get* no answer when we are faced with an indeterminate sum, we find that there *is* no answer, there *cannot* be an answer, as stated by (I*).[16] As far as the sum of gravitational forces goes, this means that,

$$\underbrace{\forall \mathbf{l} \in \Re^3, \sum_i \mathbf{F}_i \neq \mathbf{l}}_{(I^*)}.$$

With this in place we really do have our contradiction. We now have the following three equalities:

$$\underbrace{\exists \mathbf{k} \in \Re^3, \mathbf{k} = \mathbf{F}_{net}}_{(d)} \quad \underbrace{\mathbf{F}_{net} = \sum_i \mathbf{F}_i}_{(b)} \quad \underbrace{\forall \mathbf{l} \in \Re^3, \sum_i \mathbf{F}_i \neq \mathbf{l}}_{(b),(c),(I^*)}$$

From the substitutivity of identicals we can then write,

$$\exists \mathbf{k} \in \Re^3, \forall \mathbf{l} \in \Re^3, \mathbf{k} \neq \mathbf{l}$$

To be logically rigorous, we could now perform existential and universal instantiation to reach the conclusion $\mathbf{a} \neq \mathbf{a}$. In other words it follows that some three-vector **a** is not equal to itself, a blatant contradiction.

Thus the inference we make when faced with an indeterminate sum decides whether we derive a contradiction or not. As we will see further in §4.2, the force of inference (I*) can be easy to overlook, and has been overlooked by several authors both in the distant and recent past. In fact what (I*) tells us is that absolute forces in the universe *do not exist* (that our metaphysics is

[16] Cauchy wrote in 1821, '*a divergent series does not have a sum*'. However, in this paper I am assuming only that *alternating* divergent series do not have a sum. The reason is that, if we are talking *physics* (rather than mathematics) we cannot infer that a quantity which 'diverges to infinity' is equal to *nothing*. Infinity counts as 'something' here.

wrong).[17] The point is that one must have assumed that absolute forces exist in order to ask the question ‘What is the net force on a test particle?’ But then when we employ (I*) we find that absolute forces do *not* exist, contradicting the assumed metaphysics. So when one reaches indeterminacy in the way seen here it is really just another way of reaching contradiction. This puts some meat on the bones of Malament’s suggested distinction (1995, p.489) between the theory being inconsistent and the theory being ‘unacceptable *a priori*’ (because of indeterminacy): the latter is a special case of the former.

3.2.2 …from summing the potential φ

There is one final method of reasoning we have not yet considered, and there is something of a tradition of using it to demonstrate the failures of the theory. The gravitational potential at a point $\mathbf{r}$ due to a given mass m_i at $\mathbf{r}_i$ can be expressed thus:

$$\phi_i(\mathbf{r}) = -G\frac{m_i}{|\mathbf{r}_i - \mathbf{r}|} \quad \text{(Eq.1)}$$

The net gravitational potential at a point will then be $\phi_{net}(\mathbf{r}) = \sum_i \phi_i(\mathbf{r})$. This then gives rise to a new assumption about forces in the universe:

(f) The net force of gravity $\mathbf{F}_{net}$ on a body m_t at $\mathbf{r}$ is given by $\mathbf{F}_{net}(\mathbf{r}) = -m_t \nabla \phi_{net}(\mathbf{r})$, where φ_{net} is achieved by summing the gravitational potential (Eq.1).

But the problem with this method of reasoning is that the net gravitational potential is everywhere infinite: whereas the components of *force* on two opposite sides of a test mass, being vector quantities, cancel each other out (to one degree or another), the components of potential, being scalar quantities, accumulate. This time not only does the sum diverge to infinity, but it diverges to infinity relatively quickly because the potential is a 1/r relationship, whereas masses in the universe increase with r^2.[18]

[17] This is the conclusion of Norton’s 1995 paper (see especially p.515). Note the subtle difference between saying that there are no forces (i.e. **F**=**0** everywhere) and saying that absolute forces don’t exist. In the latter case ‘**F**’ doesn’t refer, so it can’t equal anything.

[18] Cf. Norton 1999, p.273. Grandi’s series arises in the context of Newtonian cosmology for the *force* because the force is a $1/r^2$ relationship and masses in the universe increase with r^2, thus cancelling each other out.

What should we conclude from the fact that the potential diverges to infinity at every point? As noted in §3.1.3, the potential is merely a mathematical tool, used to mediate between *physical* masses and forces. If there is trouble in an infinite potential, that should only be in the fact that the *physical* consequences are unpalatable. Now, as stated, to reach the force on a test mass m_t from the potential we need to take its gradient: $\mathbf{F}(\mathbf{r}) = -m_t \nabla \phi(\mathbf{r})$. But if φ is infinite everywhere this operation isn't possible, because it is not defined for infinity. One cannot proceed to derive **F**(**r**)=**0**, reasoning that φ(**r**) is everywhere *constant*. 'Constant' refers to numerical constancy, and infinity is not a number. Thus, I suggest, not only do we find that we don't know what **F** is in this case, we find that **F** is *indeterminate*. Thus summing the potential is consistent with summing the force directly using Newton's law of gravity (b), as in §3.2.1. So this is really just the same problem of indeterminacy in a different guise.

There have been several related discussions, but nowhere has the problem been identified with *indeterminacy*. Jaki (1969) writes of a 'gravitational version of Olbers' paradox' where in the latter case, in an infinite, homogeneous universe, the light from distant stars accumulates to give an infinite amount of light at any point.[19] In the gravitational case Jaki (1979) writes,

> An infinite universe of homogeneously distributed stars or galaxies cannot exist because in such a universe the gravitational potential is infinite at any point. (p.121)

But unlike light the potential is a non-physical thing. Nowhere do we find a discussion of exactly *why* an infinite potential is impossible; nowhere is there a discussion of indeterminacy. And those whom Jaki draws on apparently think that the theory demands genuinely infinite *forces*. In particular, Jaki draws at length on Einstein (1917), who argues as follows:

> According to the theory of Newton, the number of "lines of force" which come from infinity and terminate in a mass m is proportional to the mass m. If, on the average, the mass-density ρ_0 is constant throughout the universe, then a sphere of volume *V* will enclose the average mass $\rho_0 V$. Thus the number of lines of force passing through the surface *F* of the sphere into its interior is proportional to $\rho_0 V$. For unit area of the surface of the sphere the number of lines of force which enters the sphere is thus proportional to $\rho_0 \frac{V}{F}$ or to ρ_0R. Hence the

[19] Olbers, in 1823, derived that the light at any point in the universe would be equal to k(1+1+1+1+…) for a given constant k (see Jaki 1969, p.134f.). Clearly this is not indeterminate: it diverges to infinity.

> intensity of the field at the surface would ultimately become infinite with increasing radius R of the sphere, which is impossible. (Einstein 1917, p.106)

This is worth quoting in full, because to my knowledge it has not yet been made clear that this reasoning is seriously incomplete. How are the final words 'which is impossible' warranted? Rather than focusing on the potential φ he is here focused on the force field **f**, the force per unit volume at a point. It is certainly true that, in our infinite universe, there will be an infinite *component* of **f** in a given direction, but what is so impossible about this? The impossibility only comes when we consider the combined effect of all such infinite components and find that the result is indeterminate, as in §3.2.1, above. The impossibility does not lie simply in the absurdity of an 'infinite force field', as Einstein suggests.

The supposition that it is possible to derive infinite forces is not unique to Einstein. Seeliger, who finally shed light on the problems with Newtonian cosmology in 1895, supposes that there are infinite forces in a follow-up paper of 1896. He writes, 'It follows from potential theory that there must be in the *universe unlimited (infinitely) great accelerations*.' (cited in Norton 1999, p.279, emphasis in original). But this simply isn't the case. Once again, all that is shown is that there will be an infinite *component* of force in a given direction, or that the *potential* will sum to infinity. Kelvin is similarly unclear in 1901 (see Norton 1999, p.285).

In summary, we *do* get a contradiction here from summing the potential, and it is once again the indeterminacy contradiction (C5). This time we reach the conclusion because the gradient of a scalar field which is everywhere infinite is indeterminate just as the sum of an alternating divergent series is indeterminate. So what we have here is not a 'qualitatively different' type of problem, as Norton claims (1999, p.279), but just a different way of reaching the same conclusion.

3.2.3 …using Poisson's equation

Norton claims that the indeterminacy contradiction (C5) also follows from applying Poisson's equation:

> The addition of the potential φ and Poisson equation does not materially affect the indeterminacy contradiction of Newtonian cosmology. There are as many canonical solutions as there are choices for $\mathbf{r_0}$. Each distinct choice of $\mathbf{r_0}$ leads to a different force on the test body. (1995, p.514)

So Norton is claiming that indeterminacy follows from the fact that, depending on how we pick the constant of integration, we get a different result for the force. So no unique force follows from the theory, just as no unique force followed when we had an infinite sum in §3.2.1. But here as before we need to make a distinction between no unique force *following from the theory* and there being no unique force *at all*. In §3.2.1 this was expressed as two ‘strengths of inference’ (I) and (I*). With Poisson’s formulation we get an analogous pair of inferences:

(II) From an unknown constant of integration infer that the theory provides no unique solution.
(II*) From an unknown constant of integration infer that there *is* no unique solution.

But this time only the *weaker* inference (II) is legitimate. This is because the *reason why* one cannot infer two contradictory forces is different. Recall that in §3.1.1 we couldn’t infer two contradictory forces because there *cannot be* a solution to an indeterminate sum. But in §3.1.2 we couldn’t infer contradictory forces because, although there certainly *can* be a solution to an equation with an unknown constant, the theory couldn’t tell us what it was. Since we cannot make the stronger inference we cannot reach the contradictory in question and, *contra* Norton, there is no indeterminacy contradiction here.

However, when we compare the method of reasoning based on Poisson’s equation (b') and that based on Newton’s law of gravitation (b) we *do* find a conflict. We saw above that if one applies (b) to an infinite homogeneous universe, one infers that the force is indeterminate. But, as seen in §3.1.2, using Poisson’s equation instead we can move from $\nabla^2\phi(\mathbf{r}) = 4\pi G\rho$ to

$$\phi(\mathbf{r}) = \frac{2}{3}\pi G\rho|\mathbf{r}-\mathbf{r}_0|^2,$$

where $\mathbf{r}_0$ *must* be some real number (it cannot be indeterminate since then it would not be a solution to Poisson’s equation). And from here, using $\mathbf{F}(\mathbf{r}) = -m_t\nabla\phi(\mathbf{r})$, since one has a determinate potential one has a determinate force. So from (b), (b') and (c) we can infer the indeterminacy contradiction (C5) once again, where this time the determinacy of the force follows from (b').

This seems to go against Malament (1995). He *seems* to claim that (b') is actually a generalisation of (b) because when the infinite sum in question converges (b) and (b') agree, whereas (b') can also be applied when the sum *doesn’t* converge. He writes,

> There is a clear sense in which it [the “differential” formulation] is a generalization, with a wider domain of application… The “integral”

formulation is not applicable to cosmological contexts of the sort we have considered. (pp. 491 and 508)

But if (b') were a generalisation of (b) then they would *not* be in conflict. And they *are* in conflict, as I have argued above.

However, it turns out that what Malament calls the 'integral formulation' and the 'differential formulation' are not quite the same as (b) and (b'). They are the same except for one crucial interpretational difference: *gravitational force is taken to be a gauge quantity without direct physical significance*. But (b) and (b') were around long before this 20th century attitude to 'gravitational force'. Before 1900 (b') was *not* a generalisation of (b), and they were in fact in conflict. Thus Malament's analysis and the above analysis can stand side by side: one need only note that Malament is concerned with 20th century developments of Newtonian cosmology, and this paper is not.[20]

4. Why weren't the inconsistencies noticed?

In all this we find four notable inconsistencies:

(I1) C2 follows from (b'), (c) and (e). (§3.1.3)
(I2) C5 follows from (b), (c) and (d). (§3.2.1)
(I3) C5 follows from (c), (d) and (f). (§3.2.2)
(I4) C5 follows from (b), (b') and (c). (§3.2.3)

In a sense, then, Newtonian cosmology was riddled with inconsistency. Further, the assumptions in question are clearly relevant to the real history of science to *some* degree or another. This paper is not the place for a detailed history, but Jaki's history of Olbers' paradox (Jaki 1969) is a good place to start. Each of the assumptions in question enjoyed serious commitment for the relevant periods in between the years 1700 and 1900, and most were widely regarded as obvious truths.[21]

---

[20] I doubt Malament would also claim that (b), like the 'integral formulation', 'is not applicable to cosmological contexts of the sort we have considered' (that is, when the sum does not converge). As has been shown above, one certainly can apply (b) in such a context, with the result that the force on a test mass is indeterminate. To maintain that this means that (b) is inapplicable would be to use one's assumptions like a 'toolbox', where one picks and chooses one's assumptions depending on whether they lead to desirable results. Certainly *some* scientific practice proceeds in this way, but only when the assumptions one trades in are not being considered as candidates for truth.

[21] Perhaps this is less obvious regarding assumption (c), and there *were* some who doubted either the infinitude of the universe or its homogeneity, but the majority were

Thus the question of why the inconsistencies remained hidden for so long is more important than ever. This brings us back to questions (i) and (ii) introduced in §1:

(i) What was it about the *scientific community* which prevented the inconsistencies from being noticed?
(ii) What was it about the *science* which prevented the inconsistencies from being noticed?

In particular the content of the present paper enables us to provide an answer to question (ii). In addition, in §4.1, some relevant features of the scientific community are introduced to provide the beginnings of an answer to question (i).

4.1 Because the right question wasn't asked

There are a multitude of reasons why the question which leads to the inconsistencies wasn't asked. The above analysis highlights one reason in particular: the relevant question is actually rather obscure. This is made obvious by the contradiction of forces (C2) of §3.1.3. We are not asking what the *actual* force on a given body is: to answer this one would need to know, absurdly, the positions and masses of an infinite number of bodies. Rather, our question (Q) needs to be changed to,

> (Q') What is the *average* net gravitational force a test body would experience over all points of an arbitrary region of the universe R of a given volume V large enough so that the universe is homogeneous at that scale?

This could also be framed in terms of the force field **f**, as per Malament (§3.1.3), but still we would not get away from the complications of averaging. And, complications aside, it isn't immediately clear why this question is an interesting one, *except* that answering it in two different ways leads to inconsistency.

So *even if* relevant individuals had been asking pertinent cosmological questions there is some reason to suppose that the question at issue wouldn't have been asked. But the fact is that, particularly in the 19th century, the relevant individuals weren't asking cosmological questions *at all*. Merleau-Ponty (1977, p.283) refers to 'the disappearance of cosmological science as

---

convinced. For example, Bertrand Russell was convinced enough to write as late as 1897 that the infinitude and homogeneity of the universe, far from being working hypotheses (say), were scientific principles 'established forever' (as Jaki (1969, pp.184 and 220) puts it).

such in the nineteenth century, that is, the investigation of the properties of the Universe considered in its totality—until its surprising revival in the twentieth century.' It is this 'revival' which explains the title of his book of 1976 (co-written with Morando): *The Rebirth of Cosmology*. Therein he goes as far as to say that, in the 19th century, 'cosmology itself no longer existed' (p.66).

This is a remarkable claim, since there was certainly much work in astronomy and celestial mechanics during this period. But, regarding the former, 'in the course of the [19th] century astronomers were discussing the nature and internal structure of individual nebulae rather than the wider cosmological problem.' (North 1965, p.16. Cf. Jaki 1979, p.117 and Merleau-Ponty 1977, p.291). Similarly those working in celestial mechanics, such as Poisson himself, avoided cosmology entirely. For example Laplace, one of the founding fathers of celestial mechanics and active in the late 18th and early 19th centuries, never made even a single conjecture as to the structure of the universe as a whole (Merleau-Ponty 1977, p.283; Jaki 1969, p.98). And later in the 19th century, as Jaki puts it, 'The silence of Urbain J. J. Leverrier, the most celebrated French astronomer of those times… illustrated the typical aversion to cosmological problems on the part of most skilful experts on celestial mechanics.' (1969, p.157).

In the light of such facts we may consider afresh the question 'Was Newtonian cosmology inconsistent?' In the 19th century the answer should really be neither 'yes' nor 'no'; rather there is a mistake in the question, since Newtonian cosmology did not *exist* in this period in any meaningful sense. In the terms of §2, above, we may say that the relevant assumptions were not brought together scientifically because cosmological questions were not being asked which would have *required* them to be brought together. This is of course especially relevant to inconsistencies (I1) and (I4), since they both draw on Poisson's equation which was only introduced in 1813. And (I3), which depends on potential theory, is similarly mainly relevant to the 19th century (recall that Laplace introduced the potential in the 1770s).

4.2 Because of confusion about non-convergent series

Before the 19th century there is more interest in questions of cosmology, but before the 19th century is before Cauchy. Since inconsistencies (I2), (I3) and (I4) lead to the indeterminacy contradiction (C5), appreciating them depends on making the right inference when faced with a non-convergent series. Thus there is some reason to suppose that a lack of understanding of the relevant mathematics contributed to the inconsistency going unnoticed. In general terms we may say that one of the inferences necessary for the derivation of the contradiction was a peculiar type of inference, alien to the relevant individuals.

More specifically, I will provide some evidence in this section that certain individuals made inference (I) rather than inference (I*), as introduced in §3.2.1, and repeated here for convenience:

(I) from indeterminacy infer that no solution has been reached.
(I*) from indeterminacy infer that there *is* no solution.

This will also constitute my preferred explanation of the attitudes of those—Isaac Newton and Svante Arrhenius—who according to Norton favour a 'no-solution needed' solution to the inconsistency. He characterises this attitude as follows:

> They are aware of the inconsistency but ignore the possibility of deriving results that contradict those that seem appropriate… At first glance, it would seem that the physical theorists avoid logical anarchy by the simple expedient of ignoring it! (2002, p.191)

Instead I claim that they *weren't* aware of the inconsistency after all, because they only made inference (I)—'no solution reached'—and not (I*)—'no solution possible'. To decide between Norton's claims and my own a look at the primary evidence is required.

First to Newton. Was he aware of the inconsistency and chose just to ignore it, as Norton claims? In fact, although Norton does describe Newton as subscribing to a 'no-solution solution' in his 2002 paper, in his historically focused 1999 paper he suggests instead that Newton *wasn't* aware of the inconsistency. When the theologian Richard Bentley pressed Newton on the gravitational consequences of an infinite universe in 1692, Newton referred to how mathematicians handle infinities in terms of limits and convergence. Thus Norton concludes,

> Having recalled for us that there are perfectly good methods of comparing infinites by means of limits, Newton seemed not to have applied them himself to the problem at hand… It is hard to understand how Newton could make such a mistake. His mathematical and geometric powers are legendary. Perhaps Newton was so sure of his incorrect result from the symmetry considerations that he did not deem it worthwhile the few moments reflection needed to see through to a final result. (1999, p.290f.)

This story goes against the 'no-solution solution' as described in his 2002 paper. He further writes that Newton 'would surely have noticed' that there was an inconsistency if only he had applied the relevant mathematics. So it is *not* the case, then, that Newton noticed the inconsistency but chose to ignore it, as per the ‘no-solution solution’.

This suggests that we read Norton in another way. In sum he appears to be saying that *either* (i) Newton didn’t apply the relevant mathematics and so didn’t notice the indeterminacy, or (ii)—the ‘no-solution solution’—Newton *did* apply the relevant mathematics, noticed the indeterminacy, and chose to ignore it.

This paper can offer an alternative explanation. The fact that Newton is clear on how to handle *converging* infinite series is actually irrelevant insofar as Newtonian cosmology is concerned. The relevant series is infinite and *diverging*, so Newton *couldn’t* have applied the methods of limits to the problem at hand (as Norton suggests). And in fact Newton’s grasp on divergent series, and alternating divergent series in particular, was not good. In his most in-depth writings on infinite series[22]—an unpublished essay from 1684 entitled ‘On the computation of series’—Newton blatantly overlooks the fact that a divergent alternating series has no sum. Following one particular passage Whiteside’s annotation reads,

> He has, however, ignored the unpleasant fact that no unique sum is assignable to a divergent alternating series. (Newton 1971, p.611)

I take it that Whiteside is using the word ‘ignore’ in a loose sense here, and doesn’t mean to suggest that Newton saw the correct conclusion but decided to ignore it. Newton was not in the habit of ignoring what he knew to be correct conclusions.

In sum, then, a ‘third way’ seems a more plausible explanation of Newton’s attitude than either of Norton’s suggestions. This is to suppose that Newton *did* make the calculation in question, but upon coming across an alternating, divergent series made inference (I)—no solution reached—rather than inference (I*)—no solution possible. Since he *found* no solution, but didn’t conclude that there *was* no solution, he tried a different tack. As Norton notes, ‘symmetry considerations’ guided him, and he concluded that the average net force must be zero (cf. §3.1.3, above).[23] The infinities must balance after all, although apparently mathematics isn’t up to the task of showing us this.

To give a second example, Norton writes that Arrhenius ‘laid out a clear statement of the ‘no-solution solution’.’ Arrhenius wrote in 1909,

> [I]t is very much understandable that Seeliger’s argumentation is frequently construed as conflicting with the infinity of the world. *This, however, is not true.* The difficulty lies in that the attraction of a body surrounded by infinitely

[22] See Whiteside’s annotation in Newton 1981, p.267.
[23] Jaki (1969, pp.60-65) gives a nice discussion.

> many bodies is undetermined according to Seeliger's way of calculation and can take on all possible values. This, however, only proves that one *cannot* carry out the calculation by this method. (cited in Norton 1999, p.291, emphasis added)

Certainly Arrhenius did *not* think that there was an inconsistency, as Norton suggests. The confusion here seems to rest with what Arrhenius means by 'cannot'. As Norton interprets it, when Arrhenius says that 'one cannot carry out the calculation by this method' he means that, although mathematically sound, one must *avoid* that method of calculation because it leads to contradiction. This, however, leaves inexplicable why Arrhenius thinks there is no conflict. Things make more sense if we read 'one cannot carry out the calculation by this method' more literally. Arrhenius means that one just doesn't get an answer that way—one does not *reach a solution* by this method—because the sum in question is indeterminate. But this means that there still may *be* an answer, and he suggests zero (based, again, on symmetry considerations). His mistake is in not making as strong an inference as he ought to make when faced with non-convergence (he makes inference (I) instead of (I*)). This is a mistake, but it is not the mistake Norton takes it to be.

Even as late as 1954 Layzer, criticizing Milne and McCrea's neo-Newtonian cosmology of the 1930s, made this same oversight and claimed that we should infer that $\mathbf{F}=\mathbf{0}$ everywhere (Layzer 1954, p.269). McCrea put things straight the following year:

> [I]f the gravitational force is to be defined in the present manner, then *it does not exist* in the case of uniform density. Accordingly, nothing further can be inferred about this case. In particular, we may not proceed to argue, as Layzer does, that the force must be the same at every point, and thence that it must be zero. For, in order to prove that a force takes *any* value, in particular the value zero, the force has to exist in the mathematical sense. (1955, p.273, emphasis added)

What he surely means by the final remark is that, if the force really is equal to an indeterminate sum, then it can take *no* value, including zero.

Still other authors who clearly do understand non-convergence very well aren't sufficiently careful with their words to make the distinction between the force being zero (it is determined) and the force not existing (it isn't determined). Even parts of Malament (1995) are unclear on this point. In his criticism of Norton (1993) he writes,

> The integral I is not convergent, and so it is *not* the case that $I = I_1 + I_2 + I_3 + \ldots$ […] Newtonian theory … makes no determination [of gravitational force] at

all. (p.491, original emphasis)

Here we have two clear statements of the *weaker* of our two inferences (I). The stronger inference (I*) would state not only that Malament's integral 'I' is not equal to $I_1 + I_2 + I_3 +\ldots$ but that it is equal to *no quantity whatsoever*. And it would state not only that Newtonian theory *makes no determination* of the net gravitational force, but that Newtonian theory states that the net gravitational force cannot *be* any quantity. Compare this with §3.1.2, where Poisson's equation makes no determination of the net gravitational force (because we are left with unknown constants of integration), but nevertheless demands that it is *some* quantity.[24]

This subtle distinction between no force being *found* and no force being *possible* (even zero) is just the tip of the iceberg when it comes to confusion about divergent series in the relevant period, especially in the 18th century. Euler and others set the sums of divergent series equal to certain quantities right through the 18th century, and were able to reach startling *correct* conclusions by manipulating them (see Hardy 1949, ch.1). That is, setting divergent summations equal to certain values proved to be extremely fruitful. In addition, as Hardy explains, 'there is only one sum which it is 'reasonable' to assign to a divergent series: thus all 'natural' calculation with the series [1-1+1-1+…] seems to point to the conclusion that its sum should be taken to be ½' (Ibid., p.6). Apart from the fact that ½ is the mean of 1 and 0, there were some very persuasive reasons to set Grandi's series equal to ½. For example if we set

$$S = 1-1+1-1+1-1+\ldots.$$

then we might conclude that,

$$1-S = 1-(1-1+1-1+1-\ldots) = 1-1+1-1+1-\ldots = S.$$

This would mean that 1=2S, or that S=1/2. Another method was to consider the binomial expansion (discovered by Newton in the 1660s),

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \ldots.$$

This was known to converge for all x such that $0 \le x < 1$. But if it holds for all

[24] Further, Malament's statement that Newton's law of gravitation is 'not applicable' when the series in question doesn't converge (see §3.2.3, above) also suggests the weaker inference (I). And recall also Norton's introduction of assumption (d*), rather than (d), which strongly suggests the same.

such x then in the limit at x goes to 1 we find that,

$$\frac{1}{2} = 1 - 1 + 1 - 1 + \ldots$$

This latter method had already been recommended by Leibniz and was still popular one hundred years later, in the early 19th century. Poisson himself favoured this reasoning, despite living in the time of Cauchy's groundbreaking *Cours d'Analyse* of 1821, and he 'retained this staple component of his analysis throughout his life' (Grattan-Guinness 1970, p.88; see also Laugwitz 1989, p.218ff.). In fact, Grattan-Guinness claims that when Cauchy wrote in italics '*a divergent series has no sum*' it was partly aimed at Poisson (Ibid.).[25] And even as late as 1844 De Morgan still failed to appreciate that Grandi's series did not sum to ½ (see Hardy 1949, p.19f.).

Finally we may speculate as to what might have happened had someone such as Leibniz, Poisson or De Morgan noticed the relationship between Grandi's series and cosmology. Presumably, if they had followed Norton's 1999 analysis, they would have set Grandi's series equal to ½ and reached $\tfrac{1}{2} G\pi\rho\Delta r\hat{\mathbf{x}}$ for the net force (recall §3.1.1, above). But even then the *direction* of the force is indeterminate. Perhaps they would then have concluded that the only force compatible with an indeterminate direction was **F**=**0**, and that this was anyway the 'average' of $\tfrac{1}{2} G\pi\rho\Delta r\hat{\mathbf{x}}$ in all possible directions. Whatever the case, this would have made for a particularly interesting alternative history.

5. Conclusion

The complex web of interrelated assumptions which make up 'Newtonian cosmology' are thus at least partially disentangled. In the course of this analysis we learn a little more about why inconsistencies eluded us in the past, and how these inconsistencies eventually came to light. What was required was for someone with the relevant expertise to ask the right question. In addition, the fine distinction between inferences (I) and (I*) was crucial. Such issues are more important to cosmology than other science, because within cosmology only a tiny subset of conceivable experiments can actually be carried out. Advances which can be made *without* the need for experiment are thus most valuable. Similarly in the current age of science, where it is becoming more and more difficult and expensive to test theories by experiment, the more we can learn about our theories without having to turn to the laboratory the more

[25] For more on Poisson and divergent series see Grattan-Guinness 1990, p.731f.

chance we have of taking the next step forward. Reflecting on the mistakes of past science may in some small way help us to take that next step.

Acknowledgements

I am indebted to John Norton and Claus Beisbart for extended discussion and criticism, and to Oliver Pooley and two anonymous referees for asking important questions. Thanks also to all those at the Annual Meeting of the British Society for the Philosophy of Science (St. Andrews, 2008), where an early version of this paper was presented. This research could not have been completed without the help of a Royal Institute of Philosophy Jacobsen Fellowship.

References

Cat, J. (2001): 'On Understanding: Maxwell on the Methods of Illustration and Scientific Metaphor', *Studies in History and Philosophy of Modern Physics*, Vol.32, No.3, pp.395-441.

Einstein, A. (1917): *Relativity: the Special and the General Theory*. 15th edition, R. W. Lawson (trans.). London: Methuen (1954).

Grattan-Guinness, I. (1970): *The Development of the Foundations of Mathematical Analysis from Euler to Riemann*. Massachusetts: MIT Press.

_______ (1990): *Convolutions in French Mathematics, 1800-1840*. Basel: Birkhäuser Verlag.

Hardy, G. H. (1949): *Divergent Series*. Oxford: Clarendon.

Jaki, S. L. (1969): *The Paradox of Olbers' Paradox*. New York: Herder and Herder.

_______ (1979): 'Das Gravitations-Paradox des unendlichen Universums'. *Sudhoffs Archiv* **63**, pp.105-122.

Laugwitz, D. (1989): 'Definite Values of Infinite Sums', *Archive for History of Exact Sciences* **39**, 195-245.

Layzer, D. (1954): 'On the Significance of Newtonian Cosmology', *The Astronomical Journal* **59**, pp.268-270.

Malament, D. (1995): 'Is Newtonian Cosmology Really Inconsistent?', *Philosophy of Science* **62**, pp.489-510.

McCrea, W. H. (1955): 'On the Significance of Newtonian Cosmology', *The Astronomical Journal* **60**, pp.271-274.

Meheus, J. (*ed.*) (2002): *Inconsistency in Science*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Merleau-Ponty, J. (1977): 'Laplace as a Cosmologist', in W. Yourgrau and A. D. Breck (eds.), *Cosmology, History and Theology*. New York: Plenum

Press, 283-291.
Merleau-Ponty, J. and Morando, B. (1976): *The Rebirth of Cosmology*. New York: Alfred A. Knopf.
Newton, I. (1971): *The Mathematical Papers of Isaac Newton*, *Volume IV, 1674-1684*. D. T. Whiteside (ed.), Cambridge: Cambridge University Press.
_______ (1981): *The Mathematical Papers of Isaac Newton*, *Volume VIII, 1697-1722*. D. T. Whiteside (ed.), Cambridge: Cambridge University Press.
North, J. D. (1965): *The Measure of the Universe: A History of Modern Cosmology*. Oxford: Clarendon.
Norton, J. (1993): 'A Paradox in Newtonian Cosmology', in *PSA 1992*, vol.2, Philosophy of Science Association, pp.412-420.
_______ (1995): 'The Force of Newtonian Cosmology: Acceleration is Relative', *Philosophy of Science* **62**, pp.511-522.
_______ (1999): 'The Cosmological Woes of Newtonian Gravitation Theory', in H. Goenner *et al*. (eds.), *The Expanding Worlds of General Relativity*, Einstein Studies, vol.7, pp.271-323.
_______ (2002): 'A Paradox in Newtonian Gravitation Theory II', in J. Meheus (ed.), 2002, 185-195.
Seeliger, H. (1895): 'Über das Newton'sche Gravitationsgesetz', *Astronomische Nachrichten* **137**, no. 3273, pp.129-136.
Vickers, P. (2008): 'Frisch, Muller and Belot on an Inconsistency in Classical Electrodynamics', *British Journal for the Philosophy of Science* **59**, no.4, pp.1-26.
_______ (2009): 'Was Newtonian Cosmology Really Inconsistent?', *Studies in History and Philosophy of Modern Physics*, vol.40, issue 3, pp.197-208.
Wilson, M. (2009): 'Determinism and the Mystery of the Missing Physics', *British Journal for the Philosophy of Science* **60**, no.1, pp.173-193.