The Necessary Structure of the All-pervading Aether: Discrete or Continuous? Simple or Symmetric?

Forrest, Peter

Peter Forrest

The Necessary Structure of the All-pervading Aether

Discrete or Continuous? Simple or Symmetric?

Philosophische Analyse
Philosophical Analysis

Herausgegeben von / Edited by

Herbert Hochberg • Rafael Hüntelmann • Christian Kanzian
Richard Schantz • Erwin Tegtmeier

Band 49 / Volume 49

Peter Forrest

The Necessary Structure of the All-pervading Aether

Discrete or Continuous? Simple or Symmetric?

ontos verlag

Frankfurt | Paris | Lancaster | New Brunswick

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

North and South America by
Transaction Books
Rutgers University
Piscataway, NJ 08854-8042
trans@transactionpub.com

United Kingdom, Ireland, Iceland, Turkey, Malta, Portugal by
Gazelle Books Services Limited
White Cross Mills
Hightown
LANCASTER, LA1 4XS
sales@gazellebooks.co.uk

Livraison pour la France et la Belgique:
Librairie Philosophique J.Vrin
6, place de la Sorbonne; F-75005 PARIS
Tel. +33 (0)1 43 54 03 47; Fax +33 (0)1 43 54 48 18
www.vrin.fr

P.O. Box 15 41, D-63133 Heusenstamm
www.ontosverlag.com

ISBN 978-3-86838-166-5

2012

Printed on acid-free paper
FSC-certified (Forest Stewardship Council)
This hardcover binding meets the International Library standard

Printed in Germany
by CPI buch bücher.de

Introduction

This book is an investigation into the necessary structure of the *aether* – the stuff that fills the whole universe. (I assure readers that it is not the sort of aether that could have a current in it – the aether wind that Michelson and Morley famously failed to detect.[1])

Assuming what I call the aether exists, we lack *knowledge* of its structure. In fact one of my aims is to exhibit the immense variety of structures that, for all we know, it could have. Here I use the words 'know' and 'knowledge' with neither litotes nor hyperbole – we lack knowledge because the available arguments do not establish their conclusions 'beyond all reasonable doubt'. Nonetheless, we can form reasonable beliefs about its structure. I shall argue that it has no point parts: either the aether is composed of *granules* that are extended atoms; or it is point-free (gunky) in that every part is the sum of parts of less extension (diameter).[2] In particular, the aether does not have the structure of orthodox Space-time.[3] To say it had that orthodox structure would be to

---

1 To have a current, a part of aether that exists now must be identical to some part that existed a short while ago. I deny that this can happen. If we simplify the exposition by assuming that Time is discrete then the lack of current is because: (1) a part of the aether at one time is not strictly identical to any part of the aether at the preceding time; and (2) a part of the aether x at one time is caused to exist by the sum of all the aether at the previous time that is in the past light cone from x.

2 By an *atom* of aether I mean a part u of the aether that is not the sum of the proper parts. If we assume classical mereology every atom is a *simple,* that it has no proper parts. By a *granule* I mean a part u of the aether that is not the sum of parts of lesser quantity (hypervolume). It is not analytic that granules are atoms, but on my preferred hypotheses they all are.

3 Because mathematicians use the term 'space' freely to refer to systems with properties that are reminiscent of geometry, I am adopting the convention that 'Space' is written with an upper case 'S' when used literally. Then, for the sake of uniformity, 'Space-time' and 'Time' also have upper cases.

say that the aether was the mereological sum of uncountably many points, and that every non-empty set of points had a mereological sum.[4]

The granules or gunk disjunction is more specific than the grit (i.e. discrete) or gunk thesis I have previously defended (2004). For I am able to argue against the thesis of Point Discretion, namely that every part of the aether of finite diameter has finitely many point parts.

Previously, I have suggested that it an empirical question as to whether Space-time (or, better, the aether) is continuous and hence, I say, gunky, or not (Forrest, 1995). I fear the situation is more complicated than that. To be sure, the empirical confirmation of (some variant on) String Theory would provide a strong case for a continuous theory. Without such confirmation, the answer will depend on how we weigh up two competing intuitions: (1) that the non-contingent structure is highly symmetric; and (2) that the non-contingent structure is simple. I shall argue that a highly symmetric non-contingent structure leads to a continuous point–free hypothesis. Considerations of simplicity would otherwise support granules, and specifically the hypothesis I shall call Pseudo-set Granules. The current state of physics is highly speculative. But, for what it is worth, it supports a continuous point-free aether hypothesis unless we have a rather strong preference for simplicity over symmetry.

In neither of these two cases are there any points. It is traditional, however, to think of Space, and hence Space-time, as continuous and made up of points. As already indicated, I shall consider, but reject, the thesis that the aether is likewise continuous and made up of points. I draw the conclusion that the aether should be distinguished from Space-time, which is either a fiction, or real but a construct. Not much hangs on this, however, and readers may prefer to identify the aether with Space-time and draw the conclusion that the tradition is incorrect.

In this Introduction I say why I believe the aether exists, I say why we should not assume at the outset that it is the same as Space-time, I sketch the main arguments of the book, and I consider some metaphysical preliminaries.

---

[4] The Orthodoxy states rather more than this, ascribing to Space-time the structure of a topological manifold, but this manifold structure is not the object of my criticism.

1. The aether

I used to be persuaded by Graham Nerlich's *The Shape of Space* (1994) that Space-time and its parts were mind-independent substances, and that they were not dependent on spatio-temporal relations between other things. Subsequently, I have come to believe that our universe is made up of aether and that Space-time, if it exists, is best understood as a structured set of *properties*, namely point locations. So I stipulate that if there is some fundamental kind of stuff that fills our universe and if there is only one such fundamental kind of stuff it is to be called the *aether*.[5] It follows that *substantivalists* who, as I once did, hold that Space-time is itself a substance, and who deny there is anything other kind of fundamental stuff filling the universe, should identify it with the aether.[6] So for them there is no change of topic.

To justify the topic of this book, I need first to argue that there is precisely one kind of fundamental stuff filling the whole universe, and then argue that we should initially be open-minded as to whether it is the same as Space-time. First, then, I argue that there is some kind or kinds of stuff filling our universe. I claim no originality here and begin by noting the way that the topological properties of the universe – its *shape*, for short – have explanatory power. Thus Nerlich (1994) points out that the shape of the universe includes its being globally orientable, and so it is not possible for someone – call her Alice – to take a trip that will turn her into a mirror image of her former self when you she comes back.[7] If

---

5 We might want to treat particles as holes in the aether. If so the aether need not fill our universe. This is a variation I shall subsequently ignore, treating particles as temporally long but spatially thin parts of the aether characterised by some special property, such as electric charge.

6 What do I mean by 'stuff'? What do I mean by 'fundamental'? By 'stuff' I mean a homoemerous substance, that is a substance all of whose parts are of the same kind, provided we do not treat differences of mereological structure, of shape and of size as differences of kind. A fundamental kind is one whose members do not depend ontologically on other kinds.

7 If, to our surprise, the universe turns out to be non-orientable, then the example must be changed, but the argument still works.

the universe had a twist in it like a Möbius strip and so were not orientable, then such a trip would be possible.[8]

As this example shows the shape of the universe explains otherwise mysterious facts. Hence we should not treat shape as a mere convention: it is a genuine property of some substance, the universe maybe. Suppose, however, that the universe is made up of at most a countable infinity of objects such as particles and strings, with gaps between them. Then the shape of the sum of these entities is not the same as that of Space-time: the latter has all the gaps filled in. Hence we should ask whether the substance that has the shape with explanatory power is the sum of all the objects or whether it is something larger. Consider Alice's journey from Earth in year 2100 to Earth in year 2120 via – why not? – Twearth in 2110, where the years are as recorded by Alice. If she makes this journey then indeed Earth-2100 is connected to Twearth-2110. But what we are explaining is whether or not she would comes back a mirror image if she *should* make the trip and return. And if no one makes this trip then Earth-2100 might not even be connected to Twearth-2110 in the universe, because the potential paths connecting them might all have to pass through gaps. Hence the shape with explanatory power is a property of something that has the gaps filled in. Nerlich takes this something to be Space-time. For the sake of neutrality let us just call it a universe-filler, bearing in mind that all-pervading fields would be universe-fillers that many physicists already posit. Then Nerlich's argument provides a case

8 Another example of *shape*, due to Vesselin Petkov (2007) is that the number of dimensions has explanatory power, as Kant noted (1910, vol 1: 24). For instance, the electrostatic force obeys the inverse square law not an inverse cube law, as it would do if there were 4 (macroscopic) spatial dimensions. Or, even more elementary, if Space had an extra dimension then light would leak into the extra dimension, and so its intensity would not obey the inverse square law. Given a relational theory of Space-time, there must be at least three spatial dimensions, but nothing constrains us to have as few dimensions as we can make do with. Thus, if four coordinates <t, x, y, z> suffice for Space-time we could still use five <s, t, u, v, w> if we liked and take the whole physical universe to lie in the hyperplane given by the equation $s + t + u + v + w = 0$. The minimum number of coordinates that we can use does not explain why light does not leak into an extra dimension.

for the existence of at least one universe-filler. Here I note that Nerlich assumes the shape of Space-time is contingent, but that does not affect the argument, provided it is conceivable that it has different shapes. For necessary truths can be used to explain other necessary truths. The next stage of the argument for the existence of the aether is to examine and reject various objections to Nerlich's argument.[9]

One way of objecting to Nerlich's argument might be to adopt the common-sense idea that when we talk of properties of the universe we include the gaps, just as when someone buys a bushel of grain the gaps help make up the volume. (Not even the prophet Amos would complain that customers were being cheated provided the grain was pressed down.) This amounts to an appeal to a counterfactual conditional: if the gaps were filled, then the universe/grain would have such and such a property. Any such proposal is threatened with circularity: until we know what shape the universe is supposed to have, we do not know what the gaps are, but it is precisely by filling in the gaps that we obtain the shape of the universe. To avoid circularity, there has to be some more or less natural, less or more conventional, stipulation as to how the gaps are to be filled in. We could, for instance, stipulate that the volume of a heap of grain is the smallest volume of any convex region that the grain making up the heap would fit into. Now, I do not deny that in some cases there is a suitable stipulation, and in the symmetric case, discussed in Chapter Seven, it may well be possible to use the symmetries to fill in the gaps. Such ways of stipulating how the gaps are to be filled in undermine, however, the explanatory power of the shape of the universe. For whether or not Alice could have returned, as a mirror image is not a matter of stipulation, however natural the stipulation might be.

Even a natural gap-filling stipulation undermines the explanatory power of the universe's shape, then. In addition, it is metaphysically possible that no natural stipulation will succeed. I concede that in the actual universe there may well be enough particles, strings etc to ensure that there is suitable Space-time if we fill in the gaps in a natural way. But we have been explaining features that are either implied by the laws

---

[9] These have been extensively discussed in the literature. For fairly recent surveys see (Callender 2002) and (Nerlich 2003).

of nature alone or implied by the laws together with the very early history of our universe. The explanation should hold, therefore, for variants on the actual universe. In particular we may suppose that all the objects are contained in a region of finite spatial diameter. In that case, what shape is Space-time meant to have? Space might be infinite or it might be finite but unbounded. In the latter case it could indeed be a higher dimensional analog of the sphere or the torus (and so orientable) or else have a twist in it like the Möbius strip or the Klein bottle (and so non-orientable). There is danger that we will adopt a *convention* that Space be hyper-spherical in this case, making it true by convention that Alice would not return as a mirror image. I take that to be a reductio ad absurdum of the suggestion that Space-time is orientable by natural convention.

Another way of objecting to Nerlich's argument might be to explain the shape of the universe as the result of various necessary symmetries of the universe. For instance, if the universe has the shape of (four dimensional) Minkowski Space-time this might be explained by noting that light-cone structure has the Poincare´ group of symmetries. I am an enthusiast for using symmetries to characterise the shape of the aether, but symmetries presuppose some structure for them to be symmetries of. In Chapter Seven I concede that this does not require there to be something that fills the whole universe – a universe made up of enough objects (particles strings or branes) might suffice, provided these objects had the same number of dimensions as Space, and provided it was necessary that no part of space-time was too far from any object. Such necessity is not plausible. For we are here considering an alternative to a universe-filler and so there are assumed to be some parts of Space-time that are true vacua, that is, not occupied by anything. Assuming there are some true vacua, why should it not be possible to have one too large for this proposal to succeed?

A third way of objecting to Nerlich's argument is to try to paraphrase statements apparently about Space-time as about the actual *and* possible things. One problem with this is that merely possible things spatially related to actual things are strange beasts indeed, even more peculiar than David Lewis' possible worlds. Another is that there are too many possible objects to fit into finitely many dimensions, even if possible objects can overlap, like W. V. O. Quine's *possible fat men* (1948:23-24). This problem is most easily stated by considering Space and ignoring Time, but it can be adapted to Space-time. For every positive integer n there is a possible object Obj(n) with disjoint parts $a_1, \ldots, a_n$ each distance one unit from every other one. But for any finite number of dimensions m, there is some n such that no Obj(n) could be part of an m dimensional space. For example if Space is flat and 3 dimensional then $n \leq 4$. (For three dimensions, we can find small regular tetrahedra but not their 4 dimensional analogues, the pentatopes.)[10]

A variant on the third objection is to rely on the many worlds interpretation of quantum theory, which might be taken to show that each location is occupied in some world. This is, however problematic, because the same location would then be occupied by things with incompatible properties. Consider, for instance, the puzzling case of the electron that goes through two slits in a screen. (Or else consider some more realistic analog of this famous thought experiment). On the many worlds interpretation, there are worlds in which the electron goes through one slit and worlds in which its counterpart goes through the other.[11] How wholesome compared to the Copenhagen mystery mongering! But mystery returns with a vengeance if we think of a single location that both has and does not have an electron in it. Therefore, as might seem obvious anyway, the different worlds are in different locations. It follows that the excellent many worlds interpretation requires a larger Space-time and hence we still have the problem of the gaps.

There is an argument, then, for there being one or more kinds of universe-fillers. But there might be many universe-fillers, such as two distinct fields.[12] Now fields may be thought of as the aether with various

---

10 More generally, by the Nash Embedding Theorem (Nash 1956) we can embed any n dimensional Euclidean manifold in a 2n + 1 dimensional Euclidean space. So Obj(2n + 3) could not be part of the m dimensional manifold.

11 Not quite! I say that the electron is the sum of parts in many worlds, rather than having distinct counterparts, but that is a detail.

12 More precisely, two fields whose values cannot be zero or for which any assignment of a zero value is arbitrary, because the theory has a gauge group.

field properties, and particles may be thought of as parts of aether with some properties concentrated near a curve in the aether. Again if we accept String Theory, and so add some extra 'compactified' dimensions to Space-time, resulting in say 10 dimensions, then the strings are parts of the aether with some special property concentrated near a two dimensional surface. In that sense there would only be the one *fundamental* universe-filler, which I call the aether. Now, talk of the fundamental cries out for further explication in terms of ontological dependence, which is controversial. Fortunately, in this case, it is fairly clear in what sense fields are less fundamental than the aether. For consider the standard philosophical example of the lump of clay and the statue made up of that clay. And consider the case in which the clay neither exists before or after the statue. Then they share their non-modal properties, but the difference is that some shape-property that the clay has accidentally (i.e. not essentially) the statue has essentially: the lump of clay might have been a vase, the statue could not have been. Whatever we say of the clay and the statue we can say of the aether and the field.[13] And, whatever we say, we may infer that the clay/aether is more fundamental than the statue/field. If you are otherwise sceptical about talk of the more fundamental, or one thing existing in virtue of others, then this could be a stipulation. If you go so far as to identify the clay and the statue (by denying the existence of non-trivial essential properties) then please identify the field with the aether. Whichever way this topic is handled, it remains the case that we can apply Ockham's Razor to prefer just one fundamental kind of universe-filler to more than one.

In addition to the above Ockham's Razor argument, there would be an analogue of the mind/body problem if there were two fundamental universe-fillers, call them aerthur and maertha. There would have to be a co-incidence relation between some parts of aerthur and some parts of

---

This ensures that there is a difference between having no field and having a field of zero value.

13 For what it is worth, I deny that there are any modalised properties, of the form *essentially F*. I hold that the statue is the lump *qua* having various shape properties, and that if x has property F, x *qua* F is *essentially* F even if x is only accidentally F.

maertha. This is problematic in two ways. First a part of the aerthur with suitable properties would be a concrete thing, as would a part of the maertha with suitable properties. Initially we resist the thesis that two things could coincide, and it is only when examples such as the lump of clay and the statue are drawn to our attention that we might concede that thesis. So the proposal that an aerthurian thing might coincide spatially with some maerthan thing remains counter-intuitive. The other problem with the co-incidence of aerthurian things and maerthan things is that it makes a mystery of location. If there is some fundamental universe-filler then, ignoring complications, the location of an object, a particle say, composed of that universe-filler just is the part of the universe-filler it occupies, which is quite without mystery. But that account cannot be applied to two fundamental universe-fillers, for which there would be an otherwise redundant primitive relation of co-location.

I conclude that the fundamental universe-filler is unique. Now, it would be nice if we only needed the geometric properties of the aether as in the geometrodynamic program that seeks to model particles as topological features of the aether. The symmetries that are required by a continuous aether hypothesis would exclude this, but it might turn out to be correct on the hypothesis of granulated aether. And in that case we could say the aether is the only universe-filler. But even if, as in the symmetric case, we need to posit further properties and relations I have argued that there is only one *fundamental* universe-filler, which is then somewhat like Aristotelian prime matter.[14]

Our universe is filled, then, with aether, which we might or might not identify with Space-time. One reason for not automatically identifying the aether with Space-time is that in some cases the latter may be considered to depend on the spatio-temporal relations between parts of the aether, resulting in a relational theory of Space-time combined with realism about the aether. For example, let us assume a point-free, 'gunk'

---

14 James Franklin informs me that my aether corresponds to Aquinas' 'materia signata', 'materia subjecta dimensioni' (In Boeth. de Trin., Q. iv, a. 2), or 'materia sub certis dimensionibus' (De Nat. Mat., iii)). Franklin points out that the aether, as I understand it, cannot be prime matter, because to have a universal twice the matter must already have some structure.

aether, every part of which has at least one proper part. Given fairly weak constraints on its mereology and topology, we may then describe 'points' in a Whiteheadian fashion (Whitehead 1929 Ch. 2, Roeper 1997, Forrest 2010). These 'points' are not parts of the aether but set-theoretic constructs, namely ultrafilters with respect to the interior parthood relation, «, where x « y just in case for some small positive δ every part of the aether disjoint from y is at least δ from x.[15] Centuries of mathematical and physical practice have some authority as to how we use the term 'Space' and hence 'Space-time'. It may well be analytic, therefore, that Space-time, if it exists, is composed of points. In the case of point-free, 'gunky' aether, I grant realism about Space-time, with one proviso, but claim it depends for its existence on the aether.[16] The proviso is that we should be realists not so much about Space-time points but about the properties that would ordinarily be described as *being located at such and such a point*. But this is minor and systematic adjustment to genuine realism about Space-time.[17]

Another case in which we should distinguish the aether from Space-time is if, for the sake of a dynamic theory of Time, we grant the reality of the present (and perhaps the past) aether but not the future aether. In that case it is intuitive that Time exists (in the basic tenseless

---

[15] A *filter* is a set W of parts of the aether such that: (1) if x ∈ W and x « y then y ∈ W; and (2) if x ∈ W and y ∈ W then there is some z ∈ W such that z « x and z « y. (Note that there is no empty part of the aether so there is no guarantee that x and y even have a common part.) An *ultrafilter* is a filter W such that if W ⊆ V and V is also a filter then W = V. (See Forrest 2010.)

[16] I stipulate that by *realism* about Xs I mean belief in the mind-independent existence of Xs. Realism about Xs does not, therefore, imply the belief that the Xs are fundamental entities. That is, realism about Xs is compatible with believing that all (or some) of the Xs depend for their existence on some things that are not Xs.

[17] I take it that an extended thing is located *at* every point in its location and located *in* every region of Space-time of which its location is part.

sense of ‘exists’) in the future and in the past, and so the aether only occupies part of Space-time.[18]

The chief reason for distinguishing Space-time from the aether is, however, the traditionally a priori character of *local* geometry. (It is noteworthy that Euclid’s Fifth postulate, long recognised as more problematic than the other axioms, concerns the non-local feature of parallel lines never meeting.) To be sure, the unification of Space and Time into Space-time required a replacement of geometry by chronogeometry, as Alexander Alexandrov called it, but traditionalists may still make the a priori knowledge claim that Space-time is a topological manifold, and a fortiori made up of uncountably many points. Now I have no objection to a priori knowledge claims, provided they are granted to be defeasible, but I shall be arguing that our intuitions (themselves, I suspect, a priori) about any stuff filling the universe, enable us to reject what I call the Orthodoxy, namely the thesis that the aether has the structure traditionally ascribed to Space-time. There is a case, then, for the conclusion that the aether is not the same as Space-time.

I fear that if the aether is discrete then Space-time is a fiction: it is *as if* parts of the aether have locations in a continuous point-based Space-time. For example, consider the 3 dimensional case and suppose the aether is made up of simple parts that are represented as tetrahedra. Then the fictitious Space is the sum of fictitious tetrahedral sums of points. We could have taken this to be a construct not a fiction if we were realists about the properties such as *being at the centre of a tetrahedral simple*. But if the necessary structure of the aether is point-free I am reluctant to hypothesise even a non-instantiated property that is analysable in terms of the centre of something simple.

If we had an intuition that Space-time was real then the need to treat it as a fiction would be a defect in the discrete theories of the aether. I shall not, however, take that as a serious objection. First, I do not find relational theories of Space and Time counter-intuitive, and I re-

---

[18] This is problematic in the case of granulated aether, but for the symmetric case discussed in Chapter Seven we can construct the future Space-time given the aether’s past existence. Alternatively, we may construct both past and future given an extended present aether, say a layer of Planck time thickness.

quired arguments such as Nerlich's to be persuaded that relational theories were incorrect. Second, we should be wary about combining an intuition that something of a certain kind exists with claims that it is analytic that things of this kind have rather specific properties. For typically our intuitions are less precise than our concepts. So if you claim that it is intuitive that Space-time is real you should allow a conceptualisation of Space-time loose enough to permit its identification with the aether.[19]

More important than the above reasons for distinguishing the aether from Space-time is adopting a methodology in which we do not automatically assume identity, so as to protect the study of the aether from traditional geometry. To be sure, we might reject the a priori knowledge claims about Space-time, even if demoted to intuitions, but, in case we do not, it is good to set up a barrier preventing such geometric intuitions contaminating the discussion of the universe-filler.

The distinction between the aether and Space-time is corroborated by the practice of speculative physicists, who often talk of Space-time 'emerging' from some more fundamental structure. Initially, I took such remarks to imply the existence of that whereof we can only speak mathematically – I call it SAD.[20] But on closer examination they seem to be talking of the aether, taking for granted a traditional structure for Space-time. Thus David Rideout, in his dissertation on Causal Set theory under the supervision of Rafael Sorkin (Rideout 2001, p.2), writes:

> The Causal Set program postulates that Space-time is a macroscopic approximation to an underlying discrete causal order. The other familiar properties of a Space-time manifold, such as its metrical geometry and Lorentzian signature, arise as 'emergent' properties of the underlying discrete order.

I interpret the 'underlying discrete causal order' as Point Discretion, and as such a theory of the aether itself. What is said to 'emerge' is Space-

---

19 At the risk of annoying some readers I say that the prime example of this is the divine. It is intuitive that there is something divine, but it is rash then to rely on a claimed analytic truth that the only divine thing would be God – an agent of unlimited power and knowledge.

20 This is no acronym – I judge such a posit to be 'sad stuff'.

time, which is, therefore, not the aether.

I have explained why my topic is aether rather than Space-time, but I repeat that there are some hypotheses on which the aether just is Space-time, for instance, that in which the aether has simple point parts in one to one correspondence with the coordinate quadruples used to represent Space-time.

Neither points nor point locations should be confused with the co-ordinate quadruples used to represent them. Clearly a change of coordinate axes results in a different assignment of coordinates to a given point, or a given point location.

2 A synopsis of the argument

My aim is to discover the necessary structure of the aether, and in Chapter One I survey the hypotheses. A standard philosophical method would be to argue from our intuitions for some hypothesis and against the others. But there is a complication. At the Rutgers 2007 MMT Conference, I used this method, arguing from premises I found intuitive to the conclusion that Space-time (or, as I would now say, the aether) had the structure of Point Discretion, only to have it pointed out to me that this too is contrary to our intuitions.[21] Because our intuitions are jointly inconsistent, my method is first to exhibit this inconsistency (Chapter Two) and then raise the question as to which of our intuitions about the structure of the aether we should discard.[22] By the end of Chapter Three I shall have drawn up a provisional list of the more likely hypotheses. I divide theories of the aether into four using the twofold classification: continuous/discrete, point-based/point-free. I choose one example from each class, the four *exemplars* as I call them, and other hypotheses are conveniently considered as variants on these exemplars. These exem-

---

21 I confess not to remember the name of the graduate student who pointed this out to me. No point has hypervolume and on Point Discretion a unit hypervolume of the aether is the sum of finitely many points and so would have to have non-zero hypervolume, which is inconsistent.

22 Chapter Two substantially overlaps a revised version of my paper at the Rutgers MLT conference (Forrest 2012).

plars are the hypotheses that are judged best at an early stage in the inquiry.

Two of the four are discrete theories: Point Discretion and Extended Simples. Point Discretion is the hypothesis that any part of the aether of finite diameter is the sum of finitely many points. Extended Simples is the hypothesis that the aether is the sum of disjoint extended granules, and that these granules are not merely atomic but simple (that is, without proper parts.).

The remaining two exemplars are best described by stating which sets of coordinate quadruples represent parts of the aether. There is no empty part so nothing corresponds to the empty set. In both the exemplars, although not on some variants, sets that represent parts of the aether correspond to all the non-empty members of a certain *Boolean* algebra X of sets. (To say X is a Boolean algebra of sets is to say that the complement of any member of X is also in X and that the union of any two members of X is also in X.[23]) One of the continuous exemplars, the Borel Continuum, is a point-based theory, in which X consists of all the *Borel* sets of coordinate quadruples. These Borel sets form not just a Boolean algebra but a Boolean σ-algebra,. That is, given any set U in X the complement of U is also in X, and given any sequence of members of X, U_1, U_2 etc, indexed by the positive integers, the union of the U_j is also in X. The Borel sets make up the smallest σ-algebra containing all the open sets of coordinate quadruples. For any coordinate quadruples <t, x, y, z> the singleton {<t, x, y, z>} is a Borel set. So there is a point corresponding to any quadruple. In fact it is not easy, although it is possible, to show that there are any non-Borel sets of coordinate quadruples without relying on the Axiom of Choice.

This leaves one other exemplar, Arntzenius Continuum. In this case the Boolean algebra X of sets used to represent parts of the aether is the same as for Borel Continuum, but one part of the aether is represented by many Borel sets. Borel sets U and V represent the same part of the aether if the sets U - V and V - U are both of Lebesgue measure zero.

---

23 Here, as elsewhere, I take the aether to be 4 dimensional unless the contrary is explicitly asserted. So the complement of a set of quadruples Y is the set of all quadruples not in Y.

Because the empty set represents nothing, we stipulate that all and only the Borel sets of positive measure represent parts of the aether. So we may think of the Arntzenius Continuum as obtained from the Borel Continuum by first ignoring parts of zero hypervolume and then identifying those that differ by zero hypervolume.

I shall consider, and reject, the Orthodoxy, by which I mean the hypothesis that parts of the aether are in one to one correspondence with the non-empty sets of all quadruples of real numbers. Borel Continuum and the Orthodoxy agree on what points there are but disagree in that the former restricts just which sets of points have a mereological sum.

In Chapter Three I make a provisional comparison between the exemplars and some variants before considering any structure additional to mereology, quantity and extent. According to that comparison, the ranking is as follows: Point Discretion, then Extended Simples, then Arntzenius Continuum, and in fourth place, Borel Continuum, which is, nonetheless, preferable to the Orthodoxy. This is provisional in two ways. The first is that among the (defeasible) intuitions I hope readers share there is the requirement that the aether have further structure, such as a topological one. I consider this structure in Chapter Four, noting that point-set topology may be adapted not merely to the point-free ('gunk') hypotheses but also to Extended Simples and other granule hypotheses. There is, however, a serious problem in combining topology with Point Discretion. A corollary is that Point Discretion will only be tenable if there is some necessary structure that renders topology redundant. A highly symmetrical structure could serve that purpose but I shall argue in Chapter Seven that this leads to other difficulties for Point Discretion, thus completing my case against that hypothesis.

Thus far, the discussion will be 'armchair' in the sense of conducted independently of physics. In the remaining chapters I discuss ways in which physics might affect the conclusions reached in Chapter Four. Chapter Five presents the case for my disjunction: the aether is either symmetric and continuous, or it is granulated. The case is based on the problem of characterising a differentiable manifold.

In Chapter Six I take a closer look at discrete theories. I show that if, improbably in my judgement, String Theory or some variant turns out to be correct then we should reject discrete theories. The reason for this is that String Theory and its variants rebut the underminer for the prima facie persuasive Argument from Scale Invariance, namely that discrete theories posit an arbitrary scale for distance. The underminer is based on the significance of the Planck length. Another current research program, Loop Quantum Gravity, initially seems to support a discrete theory but on closer examination fails to. A third program, which has not been reported on as widely, is Causal Set theory. This is a commendably simple approach and is usually based on Point Discretion, but it can be adapted to granule hypotheses such as Extended Simples.

I end Chapter Six by considering some metaphysical arguments in favour of discrete Time, and I argue that these do not give further support to discrete theories of the aether.

In Chapter Seven I consider how we can characterise suitably symmetric aether. Finally I speculate what a 'theory of everything' would look like in the two cases being considered, namely granules and symmetry. And I provide a consideration to suggest that the symmetry intuition should trump the simplicity one.

3. The metaphysics of structure and characterisation problems

The aether has layers of, metaphysically or nomologically necessary, structure. There is the mereological structure, the topological and, according to most theories of physics, a differentiable structure. We seek not merely to describe this structure in a way that enables us to distinguish various hypotheses, but in a way that enables us to understand what the aether is.

I would be nice to say that this requires us to describe the aether's essential, as opposed to accidental, properties. Or maybe it is to describe the aether in an intrinsic rather than an extrinsic way. Many readers will, however, deny that there can be accidental but non-contingent properties, and reject the intrinsic/extrinsic distinction in the case of necessary properties. So neither of these is an accommodating ways of thinking about the necessary structure of the aether. In their place, I shall rely upon the way the characterisation helps us understand what is being characterised. Hence we should avoid back-to-front characterisations, that is, those that reverse the order of explanation. Thus we should, for the most part, avoid characterising the coordinate-independent in terms of coordinates or the frame independent in terms of relativistic frames of reference. Likewise we should, for the most part, avoid offering characterisations using set-theoretic constructions. I use the qualification 'for the most part' because we have to judge these characterisations on a case by case basis. For instance, if the best metaphysics cum physics requires there to be a unique 'privileged' relativistic frame then there is no general requirement to avoid using this choice of frame in characterising various structures. Again, sometimes what we are characterising clearly is a set, so a set-theoretic construction does not reverse the order of explanation. For example, if asked to characterise a *group,* in the mathematicians' sense of course I would say that a group is a set whose members. . . . By contrast, I think the 'definition' of real numbers as sets of rationals is back-to-front, because the rationals are real numbers and so we are attempting to characterise the reals generally in terms of sets of some of them.[24]

This requirement sets up what I call the *characterisation problem* for a given structure: how to describe it without cheating, that is, reversing the order of explanation. Suppose, for example, we want to characterise a regular tetrahedron. We could explain, to a child, say, what we are talking about just by showing examples of (almost) regular tetrahedra, contrasting with examples of clearly irregular tetrahedra, prisms, square pyramids, cones, cubes and so on. The child will soon have a concept of regular tetrahedron. But how do we characterise it? One proper way is to say a regular tetrahedron is a finite convex solid that has triangular faces of the same area, meeting in threes at each of four vertices and having six edges of the same length. It is back-to-front to say that a regular tetrahedron is the regular solid with the fewest number of faces, because that compares it with other regular solids, whose characterisation is no more fundamental.

But is it back-to-front to say that the regular tetrahedron is the convex polyhedron whose group of (orientation preserving) rotational sym-

---

24 For example, I would reject the characterisation of π as $\{x: \forall y((0 < y < x) \supset (\sin y > 0))\}$, where x and y range over rationals, and '$\supset$' is the material conditional.

metries is isomorphic to a certain twelve-member group, namely A4?[25] I hold that this is a satisfactory characterisation because symmetry is an essential component of *regularity*, in the sense in which Thaetetus discovered that there were precisely five *regular* convex polyhedra.[26] I have two things to say to readers who disagree. The first is that this bears on the choice between simplicity and symmetry regarding the structure of the aether. If characterisation in terms of symmetry is cheating then it may turn out that you are committed to a granulated aether (grit) hypothesis. (See Chapter Seven.) The second is that our disagreement might reflect a disagreement over universals. For a distance-and-orientation preserving rotation would seem to be a two-place universal. For many pairs of faces stand in the very same rotation relation. I follow David Armstrong(1978) in taking universals to be constituents of objects (and more generally states of affairs). Hence the rotational symmetries that a regular tetrahedron possesses are some of its constituents. So it is not back-to-front to characterise the regular tetrahedron's structure in terms of these symmetries. By contrast those nominalists such as Lewis who think of a relation as a set-theoretic construct out of the things related should say that symmetry-theoretic characterisations of structure are cheating, because it is back-to-front to characterise something in terms of the sets it belongs to.[27]

The characterisation problem will be used in Chapter Five, where it will be raised for the case of differentiable structure.

---

25 Here A4 may in turn be characterised as the von Dyck group D(2, 3, 3), that is, the unique-up-to-isomorphism group with three generators a, b and c such that $a^2 = b^3 = c^3 = abc = Id$.

26 Mathematicians use words such as 'regular' or 'normal' rather freely to pick out some well-behaved species of a mathematical genus, but I am concerned with the pre-theoretic idea of a regular polyhedron as opposed to some mathematical explication.

27 Although the claim that sets depend on their members is intuitive, there is, in addition, an argument for it. Sets are, some say, required to be well-founded in order to avoid paradoxes, and this requirement is only justified by the principle that a set depends on its members not vice versa.

This discussion of symmetries leads on to my quasi-realism about universals. By that I mean I shall refer to properties and relations with apparent commitment to their being universals. I take such quasi-realism to be the starting point for the debate between realists and nominalists. We speak and think as if there are universals. If such speech and talk can be paraphrased without commitment to universals then so be it, but what I object to is a double standard, in which some philosopher happily talks as if there are universals except when doing metaphysics, when it becomes problematic. I would like to emphasise, however, that by 'realism' about universals I mean the thesis that universals exist, not the thesis that they are fundamental, which I call fundamentalism about universals.

This brings me to the tricky issue of ontological dependence. I often say that some things, the Xs, exist in virtue of others, the Ys. I take this to be synonymous with saying that the Ys depend ontologically on the Xs or that the Xs are the ontological grounds for the Ys. I do not assume that readers agree with these judgements, all things considered. But I do assume that we have intuitions about dependence, such as Kit Fine's (1995) example of the singleton {b} depending on its member, b. In Chapter Two I use an intuition about dependence to support the premise that every region contains a connected part. The only exemplar that is inconsistent with this is the Arntzenius Continuum. In Chapter Three I show, however, that this results in no advantage for Borel Continuum' over Arntzenius Continuum. This illustrates the way in which I am supposing we have intuitions about ontological dependence without the conclusion reached requiring those intuitions to be veridical.

An apparent exception to this caution in reliance upon intuitions about dependence was my claim that there is only one fundamental kind of all-pervading stuff, made above. But there I was at pains to use the clay/statue example to enable readers to avoid explicitly saying that the less fundamental exists in virtue of the more fundamental.

In spite of my official caution, I am an enthusiast for ontological dependence, which is more than supervenience (McLaughlin and Bennett 2011). If the X's depend ontologically on the Ys then, I grant, the X's supervene on the Ys. But ontological dependence, unlike supervenience, is anti-symmetrical: if the Xs depend on the Ys then the Ys do not depend on the Xs. If asked to add anything, I would say that ontological grounding is much like causation. In fact I think ontological grounding is the same as being a sustaining cause. The chief difference between it and efficient causation is that the necessity involved in ontological grounding is metaphysical, not merely nomological.

And this brings me to my next metaphysical preliminary. I stipulate that anything that is metaphysically necessary is also nomologically necessary. With that stipulation in place I can assert that by the necessary structure of the aether I mean the nomologically necessary structure. Hence I say nothing contrary to the hypothesis that there are metaphysically possible worlds in which the aether is continuous, in which it is discrete, in which there are points, and in which there are none. Indeed, nothing I say is contrary to the thesis that there are metaphysically possible worlds with no aether at all.

4 Kantian objections to latter day metaphysics

I am no scholar and a fortiori no Kant scholar, but as an intuition-wallah I should say something about my appeal to many and varied intuitions, and it is convenient to do this by considering some Kantian objections. This is especially appropriate because Kant began the *Critique of Pure Reason* by considering Space and Time and someone might hold that he was right about them, even if he got carried away later on, using his discovery to cure every wooden leg – from induction to religion. As I interpret him, Kant held that various propositions about (Time and) Space are true because they are made true by our ways of (introspecting and) imagining/perceiving. If Kant is right, there is no Space-time in itself, and we should only be concerned as to how we may represent things spatiotemporally. The representation starts with the idea of Cartesian coordinates for Space, and we may develop it to consider the representation of Space-time by quadruples of real numbers, then restrict this representation to overlapping neighbourhoods as in the theory of differentiable manifolds, and if necessary consider n-tuples of reals with $n > 4$.

Even those who are otherwise unconvinced by Kant might well consider that this representation of Space-time using quadruples of real numbers has much to commend it and reflects our situation in the world, carrying around as each of us does a set of axes whose origin is here-now and has four directions: future, forwards, upwards and to the left (or, if you prefer, to the right). If Kant is correct, then I take this to be an additional reason to distinguish the aether about which I am a realist, from the Space-time, the structure of which could be taken as imposed by our imposition of the representing quadruples.

To be sure, Kant would go on to deny the possibility of knowing the aether as it is in itself. I concede that we neither *know* the aether, nor *know that* the aether exists. Now, we may contrast knowing something with knowing *that* something is the case.[28] The latter is the topic of Gettier's fame and something that analytic philosophers have obsessed about; the former is something of which direct perception by sight and touch are the paradigms, resulting in metaphors of seeing or feeling reality when discussing non-perceptual knowledge. Plato held that there is non-perceptual knowledge of things in themselves, but Kant may well be right and Plato wrong in this regard. Contemporary philosophers, however, talk of *intuitions* – but not in the same sense as Kant. These could perhaps thought of as akin to direct perception, 'seeing through a glass darkly' maybe, if we mix St Paul with Plato. But they don't have to be. Instead I take intuitions to be beliefs that are: (1) not obviously grounded in experience, (2) not inferred from other beliefs, and (3) resilient. The third clause is intended to exclude thoughts the believer has but recognises as silly, such as that something nasty will happen because it is Friday 13th. Resilience does not, however, ensure that intuitions are indefeasible.

I readily grant that intuitions are not knowledge. This is not primarily because they are uncertain. For even on those occasions where a consilience of intuitions results in beliefs very close to certainty or where the intuition is self-evident, there is something lacking that non-inferential knowledge is usually considered to have, namely counterfactual dependence on the way things are (Goldman 1976, Nozick 1981). Consider, for instance, an intuition that I rely upon and which I call Hume's Razor (Forrest 2009): Necessities are not to be multiplied more than necessary. The idea is that you should be reluctant to *hypothesise* that some truth is necessary. If Spinoza was right and all truths are nec-

---

[28] The two kinds of knowledge overlap in those cases in which knowing that p is knowing the state of affairs that p.

essary then, presumably the rest of us would still consider it extravagant to posit necessities unnecessarily.

Kant also argued that a priori knowledge claims – unless restricted in scope – lead to contradictions – the infamous antinomies. Kant's illustrations of this may be construed as cases of conflicting intuitions, and merely show that not all intuitions count as knowledge. I also note that there is scope for resolving antinomies by making distinctions. Consider, for instance, an antinomy that is marginally relevant to this work. Kant argues both that the world has no beginning and that it must have one.[29] Whether or not you find Kant's somewhat obscure arguments persuasive, you might perhaps agree that Space-time has no beginning, but hold that the physical universe has one. Maybe you hold that the Big Bang is the universe's birth rather than a mid-life crisis. Or you might be persuaded by the Kalam argument as presented by William Craig (Craig 2000). Any reason for believing aether has a beginning, but Space-time does not, provides us with a reason for distinguishing aether from Space-time.

There remains the question of justification: why trust our intuitions? Well some of them are self-evident, that is they are obvious in a resilient way – they stay obvious even when argued about and reflected upon. But some important intuitions, including the reliability of intuition itself, are not self-evident. Notoriously, David Hume drew our attention to the way reliance on induction was not self-evident (not 'intuitive' in his sense of the word). Nor are appeals to Ockham's Razor, or to simplicity more generally. Should we therefore adopt the sceptical position that only self-evident intuitions are to be trusted? No, for as Alvin Plantinga (1984) has pointed out that sceptical position can no more be justified than the intuitions it rejects.

I conclude that any attempt at the justification of intuitions that are not self-evident is bound to rely on intuitions that are not self-evident either, and so be circular (see Hales 2000). In that case, it is tempting to try to justify some but not other intuitions: those without which the sci-

---

[29] I explicate the concept of a *beginning* of X to mean an interval of time, in some part of which X exists but it is not the case that X existed before that interval.

ences would be undermined; those that gave our ancestors a greater chance of survival and reproduction; or those that are attributed to divine providence. But that does not avoid the circularity of justifying intuitions intuitively. So I can do no other than assert that reasoning from intuitions is on the whole reliable.[30] And if you want to pick and choose, asserting the reliability of some but not others, then I am puzzled but I cannot argue against you. Let me be more explicit: I assert that intuitions are to be trusted unless they are defeated, that is either *undercut* or *rebutted* (Pollock 1967). Undercutting, or as I say, *undermining*, occurs if a good case can be made for their having arisen in an unreliable way, rebuttal if they clash with other beliefs, such as other intuitions. My assertion could be thought of as faith – faith in human reasoning powers – or as a Jamesian 'passional choice'.[31] Because the undefeated intuitions form the basis of metaphysical reasoning this defence of intuition is also a charter for a *metaphysicians pride* movement. (To all philosophers who are still ashamed to be called metaphysicians, I say, 'Come out of the closet!')

It is customary to assert your own intuitive judgements and leave it up to others to agree or disagree. I shall follow this convention while noting the peculiarity of ignoring the intuitions of others. I hope, however, that there is a consensus concerning intuitions about the aether, even if none can be found when it comes to politics and religion. By a consensus, in this context, I do not mean that all share my intuitions, merely that none of my intuitions are judged counter-intuitive by others or vice versa.

5. Metaphysics and the philosophy of physics

One distinguished metaphysician expressed surprise that an investigation into Space and Time could contain an a posteriori component. My shift

---

[30] When I say 'can do no other' I mean by way of justification. Whether we accept the reliability of intuition as a mystery or explain it (in terms of God) is another question.

[31] For a recent defence of 'passional choices' see (Bishop 2007).

of topic from Space-time to aether deals with that. The opposite is expressed by the following remark by a philosopher of physics:

> The use of alleged pre-theoretic, a priori intuitions in investigating issues of ontology is inappropriate in the context of formulating interpretations of physical theories.

There are two distinct criticisms implicit in this quote. The first concerns the suspicion that the intuitions relied upon are not pre-theoretic but reflect, say, immersion in Newtonian physics. My response is that intuitions are defeasible and they may, on occasion, be undermined by an examination of how we came to have them. That some intuitions have been undermined is, however, a bad reason for their wholesale rejection. To be sure some may find the bad reasoning intuitive but that intuition is self-refuting.

The other criticism in the quote asserts that there is an intellectual activity of 'formulating interpretations of physical theories' that is best done without resort to intuitions. The criticism illustrates the gap separating metaphysics from philosophy of physics. The latter concerns rigorous proofs and, as the quote indicates, distrusts arguments based on intuition. I have already defended my reliance on intuitions and I challenge philosophers of physics to explain how we can do without them if we are to investigate the reality that scientific realists believe – correctly I say – to explain the phenomena.

Consider, for example, that paradigm of philosophy of physics, the Hole Problem This problem, which delayed General Relativity from 1913 to 1915, was re-introduced by John Stachel and subsequently presented as an argument against realism about Space-time by John Norton and John Earman. (Norton 2008). If the Space-time is real there is a fact of the matter as to whether a given macroscopic object, a rock say, has a given location. We then note that the states defined by the distribution of the energy-momentum and the gravitational field can neither be inferred from observation nor determined by Einstein's equations. This is quite general, but is most easily illustrated in the case in which we expect Special Relativity to be a good approximation, nearly 'empty' Space-time with just a few small well-separated rocks in it. (Of course, I say that Space-time would not be nearly empty but rather full of the aether.) General Relativity permits a state described as just such an approximation to Special Relativity, in which the aether is almost flat everywhere. There are, however, infinitely many other solutions including those in which the aether is almost flat outside the ‘hole’ and highly curved inside it. (The problem arises because General Relativity implies six independent equations but the gravitational field requires ten scalar functions to specify it.) Other things being equal, it is reasonable to conclude that two ‘states’ that are indistinguishable in this way are in fact two descriptions of the same state. But if we do draw this conclusion, then there is no fact of the matter as to whether a given event, say the collision of two rocks, has a location, for such locational ‘facts’ depend on which ‘state’ is used to describe the one true state.

This argument relies heavily on an intuition, namely that other things being equal it is reasonable to conclude that two ‘states’ that are indistinguishable in this way are in fact two descriptions of the same state. And this is a metaphysical intuition, one about the things in themselves. To be sure, it could be dressed up as a rule of scientific enquiry, partly constitutive of the scientific method. I am inclined to agree, but treating our reliance upon an a priori metaphysical intuition as constitutive of scientific method in turn shows that science has not and cannot emancipate itself from metaphysics.

I conclude that the gap between metaphysics and philosophy of physics should never have arisen.

1. A Survey of Hypotheses about the Aether's Structure

The obvious features of the aether are: (1) its mereological structure, namely that it has parts and that some parts of the aether are parts of other parts; (2) the way parts of the aether vary in *extent*; and (3) the way they vary in *quantity*. The aim of this chapter is to survey hypotheses about these elements of structure. The survey will exhibit the great variety of the hypotheses between which I will be choosing, and so counter the blancmange prejudice, namely the assumption that the aether has no interesting structure.

The survey will also, I hope, provoke readers into provisionally judging them in an a priori way, and so overcome any empiricist scruples they might have. I say 'provoke' because it is customary not to bother with 'silly' or 'crazy' hypotheses unless you intend to defend them. But this custom covers-up our reliance upon the a priori intuitions used to reject them, as does the careless dismissal of such hypotheses on grounds of simplicity. At the point where readers exclaim, 'spare us any more hypotheses' I urge them to be honest about their reliance on the (defeasible) a priori.[32]

My survey is carried out subject to certain restrictions. One of these is that *extent* is explicated as diameter; bearing in mind that Relativity seems to imply that diameter is frame relative. Likewise the *quantity* of a region will initially be explicated as its hypervolume – the 4 dimensional volume analog – although in the next chapter I shall argue that this is not always the case. These restrictions are, therefore, simplifications for the purposes of exposition.

To survey the hypotheses I use the method of coordinate representation. There are other methods, such as describing the actual structure

---

[32] Recently Laurence Bonjour (1998) has defended rational insight as a source of a priori beliefs. This suggests a remarkable faculty. Possibly we have such a faculty, and if so the only explanation I find satisfactory is divine providence. But I see no need to posit it. For I hold to an epistemology in which *any* belief whatever is rational unless there are grounds for questioning it that due diligence either has or should have revealed.

by noting that adjoining some fictions results in a more familiar structure. That may be compared to Hartry Field's program of science without numbers, where numbers are presented as fictional additions to the physical entities (Field 1980). Because the method of representation is easier, I shall ignore the method of fictions except for some very special cases, notably the fictional addition of the empty region.

I begin with a brief account of the mereology of the aether, partly to introduce ideas that will be needed later, but chiefly to combat the blancmange prejudice mentioned above. Although no issues will be settled in this chapter, the sheer variety of hypotheses shows that there is a question to be answered as to the mereological structure of aether and that it goes beyond the debate over whether there are aether atoms .

1. Aether mereology[33]

I begin by restricting attention to *full* parts of the aether. The contrast is with, for example, the putative part of the aether consisting of 50% of the aether everywhere. Ockhamist appeals to simplicity support the position that all parts of the aether are full parts, but I shall have occasion to consider the alternative in Chapter Five. In accordance with this restriction to full parts I define a *region* to be a full part of the aether, and concentrate on the mereology of regions. To avoid confusion, I do not call parts of Space-time regions.

Mereology concerns the part/whole relation, $x \leq y$. As is standard, I stipulate that everything is part of itself, $x \leq x$, and define *proper parthood*, $x < y$ by: $x < y$ if $x \leq y$ and $x \neq y$. I take it as self-evident, and no doubt analytic, that parthood is transitive, that is, a part of a part is a part, and that proper parthood is anti-reflexive.

The other axioms are metaphysical hypotheses, even if we grant their non-contingency. For instance, because of the restriction to one kind of stuff, the aether, the mereology is taken to be extensional, that is

---

[33] For expositions of mereology see Peter Simons (1987) and Achille Varzi (2009).

no two non-atomic regions can have the very same regions as proper parts.[34]

An *upper bound* of some regions, the Fs, is defined to be any region of which every F is a part. A *lower bound* of the Fs is any region that is part of every F. Two regions without a lower bound are said to be *disjoint*. Otherwise they are said to *overlap*. One of the differences between the part/whole relation on the aether and many other orderings is that there is something, œ, the whole aether, of which every region is part, but there is no empty thing, ø, that is part of every region. The other difference is that because the relation of being disjoint is of intuitive importance there are two rival ways of explicating the non-technical idea of the combination of two or more region into a larger one. I shall stipulate that the word 'sum' is reserved for this non-technical idea of the combination operation and not be used as synonymous with *fusion*, which is thus one potential explication of summation. Although the definition of a fusion of the Xs is familiar it is somewhat convoluted: a fusion of the Xs is some region that overlaps all and only the regions that overlap some X. It is customary to speak of *the* fusion, assuming uniqueness. Uniqueness of fusion follows from the principle of Weak Supplementation, namely that if x is a proper part of y then there is some part z of y disjoint from x. Weak Supplementation will be discussed further below.

More straightforward than fusion is the idea of the *join* $x \vee y$ of x and y, or more generally the join of the Xs, $\vee X$. It is the least upper bound, that is, an upper bound (of x and y or the Xs, respectively) that is part of every other upper bound. If there is a join thus defined, then it is unique. Joins are one possible explication of sums. Likewise we may define the meet ($x \wedge y$, or $\wedge X$) as the greatest lower bound, that is, a lower bound of which every lower bound is part. Clearly disjoint regions have no meet. It is, however, intuitive that overlapping regions always do. If there is a meet then it is unique.

Some hypotheses about mereology

---

[34] Controversies about extensionality (Varzi 2009) arise because of the way one thing can be constituted by another, as in the lump of clay and the statue.

One way of assuming more than the minimum amount of mereology is to make the plausible supposition that any two overlapping regions have a meet and any two regions a join. Likewise it is intuitive that both the distributivity laws hold.[35] To state the laws neatly we take the meet of disjoint regions to be the fictional empty region, ø. If we adjoin this fictional region ø, then I say we are considering a *lattice* of regions, reserving the term 'mereology' for the system of real regions. So a *distributive lattice* of regions is one in which any two members have a join and a meet and the distributive laws hold: $x\wedge(y\vee z) = (x\wedge y)\vee(x\wedge z)$ and $x\vee(y\wedge z) = (x\vee y)\wedge(x\vee z)$. A mereology whose lattice is distributive will be called a *distributive mereology.*

Another operation of interest on a lattice is the complement ¬x of a region x. This is the join of all the regions disjoint from x. Clearly ¬ø = œ and ¬œ =ø. If the distributive laws hold then the complement of a join is the meet of the complements, and the complement of a meet the join of the complements. If in addition, for any x, ¬¬x = x, then a distributive lattice is said to be *Boolean,* as is the corresponding mereology.

We may also define the difference x – y as the join of all the parts of x disjoint from y. In a lattice $x - y = x\wedge\neg y$.

It is intuitive that the mereology of the aether is Boolean as are all the four exemplars. The cases of infinite joins and meets are not so clear.

If finite meets distribute over arbitrary joins, the mereology becomes a *complete Heyting* lattice if the fictional empty region ø is adjoined.[36] I call this a complete Heyting mereology. It may be thought of as the result of weakening classical mereology by abandoning Weak Supplementation.

---

35 The necessary and sufficient condition for either of the two distributivity laws is that (even including the fictional empty region ø) there are no five regions that instantiate the mereological relations represented by the Hasse diagrams M5 and N5. As a corollary either distributivity laws implies the other (Grätzer 1971).

36 The categories of complete Heyting algebras, frames and locales have the same objects, which I call complete Heyting lattices, but differ in their morphisms. See (Johnstone 1982.)

Classical mereology (restricted to regions) or *general extensional mereology* as it is widely called, is obtained by adjoining to the transitivity and anti-reflexivity of parthood the one, deceptively simple, axiom that any regions have a unique fusion. This implies that the unique fusion is a join. In fact it shows that if we adjoin the fictional empty region the mereology becomes a complete Boolean algebra. The principles of classical mereology imply those of complete Heyting mereology but the converse is not the case. The difference between classical mereology and (non-classical) complete Heyting mereology may be illustrated in a heuristic way by considering the case in which every region has positive diameter but there is a point particle. Then in classical mereology the particle should either be inside a given region or outside it, while in complete Heyting mereology the particle could as it were be sitting on the fence between two regions.

The case for classical mereology is based upon the intuitive principle of Weak Supplementation, namely that if x is a proper part of y then there is some part z of y disjoint from x. From this it follows that the join of some regions must be a fusion, and that any fusion must be the join. Classical mereology follows from Weak Supplementation together with the principle that any regions have a fusion. Although I consider Weak Supplementation somewhat intuitive, I want to contrast it with a much firmer intuition, Interior Part Supplementation:

> If x is an *interior* part of y then there is some part z of y disjoint from x.

To say that x is an *interior* part of y is to say that x is *separated* from every region z disjoint from y. Separation may in turn be characterised in the continuous case by saying that two regions x and y are separated if they are a positive distance apart. (See Chapter Two, Section Three, for further discussion of separation). In the next chapter I consider a premise (Premise Eight: Hypervolume Supplementation) that is more intuitive than Weak but less intuitive than Interior Part Supplementation.

Suppose we do not assume Weak Supplementation but assume the lattice obtained by adjoining ø is a complete Heyting lattice. Then the complement ¬x of x is defined, as above, as the join of all the y disjoint from x. It follows that x ≤ ¬¬x. If x = ¬¬x then x is said to be *regular*. The regular members of a complete Heyting lattice form a complete Boolean lattice (although the join in the subsystem of regular regions is not always the same as the join in the system of all the regions). One way to obtain a classical mereology, then, is fictionally to identify regions x and y that are *classically equivalent* in the sense that $\neg\neg x = \neg\neg y$.

Explicating summation

Initially it is not clear whether the non-technical idea of the sum should be explicated as join or fusion, so it is worth considering the difference, if there is one. To be sure in classical mereology any regions have a unique fusion, which is their least upper bound, so there is no difference, but classical mereology, although intuitive, is defeasible.

I now argue that the sum is a join that is also a fusion. First, I note that if the join of the Xs exists then it overlaps every region that overlaps some X. Hence, if it fails to be a fusion, that is because it overlaps some region disjoint from every X, and hence has a part disjoint from every X, which is not what we expect from the sum of the Xs.

Next I consider an example. Because I am here considering the general concept of summation I am not required to restrict myself to the aether but I may consider Time in isolation from Space. In that case, one hypothesis we might consider is that all Time's parts are *intervals,* represented by sets of real numbers of the form $\{t: a < t \leq b\}$, where a < b and we allow as special cases $a = -\infty$ and $b = \infty$.[37] Call this hypothesis Intervals Only. It generalises to the aether, where it becomes the hypothesis that all regions must be convex, but I find that much less intuitive than Intervals Only for Time. Now consider the two intervals c and d represented by $\{t: 1 < t \leq 2\}$ and $\{t: 3 < t \leq 4\}$. They have a join e represented by $\{t: 1 < t \leq 4\}$. But they have no fusion. Intuitively they have no sum because the sum, if it existed would have to be represented by $\{t: 1 < t \leq 2\} \cup \{t: 3\ t \leq 4\}$, but on Intervals Only there is no such part of Time. This example supports the requirement that the sum of regions should be a fusion.

On the other hand, it is intuitive that the sum of regions contains every one of them as parts, so the sum, if it exists, must be an upper

---

37 The use of the half-open intervals is not essential for the example.

bound. If it is not a least upper bound then why should not some smaller one deserve to be considered a sum? Intuitively a sum, if it exists, should be unique. Joins are unique but fusions not always so. Therefore the sum should be the join. I explicate the idea of the sum, then, as a join that is also a fusion. One intuitively appealing requirement for mereology is, then, that any Fs have a sum, and hence a join that is also a fusion of the Fs. This will be defended in the next chapter. It is, however, weaker than classical mereology, which requires that fusions be unique. In fact it holds in a complete Heyting mereology, in which there can be two regions u and v, where u is a maximal proper part of v. In that case u is a fusion of v and vice versa. Rather than take this as counter-intuitive, I think we should grant that, outside classical mereology, fusion fails to explicate summation.

Simples and other atoms

Given classical mereology any region x ≠ œ has a unique complement. This enables us to divide œ into atoms and gunk, where an atom is a part that is not the sum of two disjoint parts and any part of gunk has a proper part. In classical mereology atoms coincide with *simples*, that is, parts of œ with no proper parts. In that case, either there are no atoms, so aether is gunk, or the aether is the sum of atoms, in which case there is no gunk, or the aether is the sum of two portions, the sum of all atoms and the sum of all gunk. I reject the latter hypothesis on Ockhamist grounds. The only further question of mereological interest is then just how many atoms there are or how many parts to gunk. If, however, mereology is non-classical but the aether is the sum of atoms, then by definition all simples are atoms but it is an open question whether all atoms are simples.

The size of a mereology

'How big is the mereology of regions?' Here the only answers that we are tempted to, but should not, ignore are those in which there are too many parts to have a cardinal number assigned. For instance, suppose the aether, œ, is a hypersphere and is made up of points (regions of zero diameter). Then each point might be an exact replica of the whole of œ, with its own diameter function, which might then be considered to be of infinitesimal amount. We might also consider the hypothesis that what we think of as the whole of aether is but a point in another, qualitatively identical region. By itself this does not provide an exotic mereology, it is gunk with no more regions than there are real numbers. Now, however, let us go further and suppose that for every limit ordinal α there is a point corresponding to every sequence of 'worlds within worlds' indexed by the ordinals less than α. The resulting hypothesis is one of *hypergunk,* (Nolan 2004, Hazen 2004) in which there is an absolute infinity of regions, that is, more than that of any set.

A survey of hypotheses about mereology

Even before we consider diameters and volumes we have, then, quite a variety of hypotheses about the structure of the aether. Of these the ones that I take to be most significant are hypotheses about gunk versus atoms, and about the number of regions along with the following hypotheses, which I have already described.

Distributive mereology: If we adjoin ø, any two regions have a join and a meet and the distributivity principles hold.

Boolean mereology: A distributive mereology, whose lattice is Boolean

Summation mereology: A lattice mereology in which any regions have a join that is also a fusion.

Complete Heyting mereology: A summation mereology where the following infinite distributivity principle holds in the lattice:

> For any x and any Fs, the meet of x with the join of the Fs is the join of the meets of x with individual Fs

Classical mereology: A mereology for which any regions have a unique fusion, so if ø is adjoined the lattice is complete Boolean.[38]

---

[38] The initially plausible hypothesis that classical mereology holds for the aether does not imply that it holds more generally. I have suggested that sets can be replaced by pseudo-sets, which have as their pseudo-members parts. Thus, {b}*, the pseudo-singleton of b, has b as a proper part that has every proper part of {b}* as a part, and so there is no part x ≠ {b}* such that {b}* is the sum of b and x (Forrest 2002).

A classical mereology must be complete Heyting and Boolean. A complete Heyting mereology must be a summation mereology. And a summation mereology must be a distributive mereology. A Boolean mereology that is also a summation mereology must be classical. A summation mereology that satisfies Weak Supplementation must be classical.

All these hypotheses about mereology have *curmudgeon* variants, obtained from the same lattices but with the qualification that not only ø but also some joins of disjoint regions are fictional. These fictions include all the joins of the regions that do not touch. The intuition behind these curmudgeon variants is that if a 'region' is disconnected then 'it' is in fact many regions, not one. In the next chapter I undermine this intuition. I also consider that this intuition is defeated by the following Scattered Object Argument. The first premise is that there are scattered material objects. The second is that a scattered object is constituted by a scattered region. So there are scattered regions. Although quantum theory clouds our understanding of such things, I claim that familiar objects, including ourselves, are scattered, consisting of nuclei that never touch each other, together with electrons that never touch the nuclei. This supports the first premise. The second premise could be rejected on the grounds that a material object is constituted by many regions not a single one. To that I reply that this introduces a mysterious mode of constitution to be contrasted with the way in which a single region constitutes a material object.[39]

The Scattered Object Argument is not totally persuasive, so it is worth bearing in mind curmudgeon variants. Of the exemplars, one, Arntzenius Continuum has no curmudgeon variant, because it implies that there are regions without connected parts. (See Chapter Two.) So curmudgeons should reject rather than modify it. Otherwise, all the regions are sums of connected ones and curmudgeons may treat them as fictions. Apparent reference to disconnected regions may then be paraphrased as plural reference to the connected parts.

Another family of hypotheses is obtained by proposing that a countable family of regions has a join and a meet but not all uncountable

---

[39] The only difference between the material object and the region is that some accidental properties of the region are essential properties of the object

families do. The case for the existence of arbitrary sums is based on the intuition that unrestricted principles are more plausible than restricted variants. I note that as a reason for preferring unrestricted summation.

I list these hypotheses because they are supported by distinct intuitions, but the reader might supplement the list using Peter Simon's thorough discussion (1987). My purpose is to emphasise the variety of hypotheses, and so open up the question of the mereological structure of aether, rather than taking it for granted that classical mereology holds. As far as I can see, the case for classical mereology is based on two premises: (1) Weak Supplementation, which is intuitive but not as firmly so as the analog based on interior parthood, mentioned above; and (2) Universal Summation, which has an advantage over Countable Summation, but only because unrestricted rules are simpler than restricted ones. The resulting support is genuine but not especially strong.

2. Coordinate representation

If the aether has the structure of a manifold, it is the join of suitable overlapping regions each of which has a coordinate representation, but there may not be a representation of the whole aether.[40] If, however, we are concerned only with the local structure of the aether that complication is irrelevant. So, one way of examining this local structure is to represent regions by sets of quadruples. I require the diameter and hypervolume of a region to be approximately that of the representing set. This mention of diameters and hypervolumes might alarm readers for neither is fundamental to the description of regions. Diameter is frame-relative, and hypervolume may be replaced by the greater-hypervolume-than ordering. But my initial aim is to use our intuitions about hypervolume and diameter in order to rank hypotheses, delaying their further discussion

---

40 I shall have more to say about manifolds in Chapters Four and Five. A 4 dimensional aether-manifold is the sum of (mereo-) topologically open regions u_j, j = 1, 2 etc, where, for each j, u_j is represented by an open subset of $\mathfrak{R}^4$, the topological space of all quadruples of reals, in such a way that the open parts of u_j are in one to one correspondence with some of the open subsets of $\mathfrak{R}^4$. We may always suppose that for each j, u_j is represented by the whole of $\mathfrak{R}^4$.

until Chapter Four. That is quite compatible with diameter and hypervolume being derived from more fundamental relations. To be sure, the diameter depends on which relativistic frame is used, but that is of no concern in this chapter because we are representing regions using coordinate quadruples – any such representation already requires a choice of coordinate axes, and hence a frame.

Concentrating on the local structure if necessary, I take the aether, œ, to be represented using some or all the coordinate quadruples in some open convex set Œ, so $Œ \subseteq \Re^4$ the set of all quadruples of real numbers. Readers may choose to consider the special case in which $Œ = \Re^4$, with no loss of generality relevant to this chapter.

$\Re^4$ is equipped with a *norm*, that is, the distance from <0,0,0,0>, defined by $||\langle t,x,y,z\rangle|| = \sqrt{(t^2 + x^2 + y^2 + z^2)}$, and with the Lebesgue measure μ. I use the technical phrase 'Lebesgue measure' in place of 'hypervolume' because I am considering sets of quadruples rather than regions. The basic idea is that the Lebesgue measure of a set of quadruples is the hypervolume of a region in a fictitious 4 dimensional Euclidean space whose points are represented using Cartesian coordinates t, x, y and z. [41]

The *distance* between quadruples ξ and η is $||\xi - \eta||$. If $X \subseteq \Re^4$, then |X|, the *diameter* of X = $\sup\{||\xi - \eta||: \xi\in X, \eta \in X\}$, that is, is the least upper bound of $\{||\xi - \eta||: \xi\in X, \eta \in X\}$ or $+\infty$ if there is no upper

---

41 More formally, the Lebesgue measure μ is uniquely characterised as follows:

1. If a, b, c, d, e, f, g and h are real numbers with $a \le b$, $c \le d$, $e \le f$, $g \le h$ then $\mu(\{\langle t,x,y,z\rangle: a \le t \le b, c \le x \le d, e \le y \le f, g \le z \le h\}) = (b - a)(d - c)(f - e)(h - g)$.
2. The Lebesgue measure is a non-negative real number or $+\infty$
3. If $X \subset Y$ and X and Y have Lebesgue measures then $\mu(X) \le \mu(Y)$.
4. If $X \subset Y \subset Z$ and if X and Z have the same Lebesgue measure then Y has a Lebesgue measure.
5. Lebesgue measure is countably additive on pairwise disjoint sets. That is, for any countable sequence X_m of sets of quadruples such that $X_m \cap X_n = \emptyset$ if $n \neq m$, $\mu(\cup X_m) = \sum\mu(X_m)$, provided all the X_m have a Lebesgue measure.

bound. It follows that if ξ is a quadruple, $\|\xi\| = |\{\xi, <0,0,0,0>\}|$, so unless $\xi = <0,0,0,0>$ $\|\xi\| \neq |\{\xi\}|$.

$\Re^4$ has a *topology*, namely a family of *open* subsets T of $\Re^4$ containing Ø and $\Re^4$ itself, such that T is closed under the operations of finite intersection and arbitrary union. T is specified as follows: a set U is open just in case, for any $\xi \in$ U there is some positive number ε such that if $\|\eta - \xi\| < \varepsilon$ then $\eta \in$ U. A *closed* set is one whose complement is open. It follows that both the empty set $\varnothing$ and the set of all quadruples, $\Re^4$ are both open and closed. No other sets of quadruples have that distinction. The *closure* of a set is the intersection of all the closed sets that include it, and the *interior* of a set is the union of all the open sets included in it. The closure of any set is closed and the interior of any set is open. A set is said to be *regular open* (also called *perfectly open*) if it is the interior of its closure. A non-empty open set is said to be *connected* if it is not the union of two disjoint non-empty open sets. So $\Re^4$ is itself connected.

A *representation* of the aether is a mapping Φ assigning to any region u one or more non-empty sets of quadruples of real numbers $\Phi(u)$. For the moment I consider the case in which the representation is a single-valued function, that is, it assigns a single set of quadruples to a region. I require that diam(u) $\approx |\Phi(u)|$. I require that if u has a volume and if $\Phi(u)$ has Lebesgue measure $\mu(\Phi(u))$, then hvol(u) $\approx \mu(\Phi(u))$. Here diam(u) is the diameter of u and hvol(u) is the hyper-volume of u, the 4 dimensional analog of ordinary volume. The symbol '≈', used for approximate equality, is required in place of '=' for two reasons. The first and most straightforward is that if the aether is curved there will be a distortion due to the curvature. We may restrict attention to a region chosen so that this is no more than 1% and largely ignore this distortion as irrelevant. The second is more serious. Something peculiar might go on when we reach the Planck barrier. Suppose K $\approx 10^{44}$, so (1/K)sec is of the order of magnitude of Planck time. For instance Point Discretion might be correct with some K^4 points in a hypercube of side one (light) second. In that case there could be further distortion so that, in (light) second units, diam(u) is within 1% of $|\Phi(u)| \pm 1/K$. Likewise we might suppose that hvol(u) is within 1% of $\mu(\Phi(u)) \pm 1/K^4$.The details do not matter much and I continue to write diam(u) $\approx |\Phi(u)|$ and hvol(u) $\approx \mu(\Phi(u))$.

The representation should preserve the part/whole structure: for all parts u, v of œ, if $u \leq v$ then $\Phi(u) \subseteq \Phi(v)$ We should not assume, however, that the representation is *faithful,* that is a one to one correspondence, in the sense that if $u \neq v$ then $\Phi(u) \neq \Phi(v)$. For if there are points u and v an infinitesimal distance apart then there is a coordinate quadruple $<t,x,y,z>$ such that $\Phi(u) = \Phi(v) = \{<t,x,y,z>\}$.

Even if it is faithful, the coordinate representation might not preserve joins and meets. The most familiar example is Tarski Continuum,, the hypothesis that the regions are represented faithfully by precisely those sets that are non-empty *regular open* sets of coordinate quadruples (Tarski 1956). One of the nice features about Tarski Continuum is that the regions satisfy the principles for classical mereology because any regions, even infinitely many, have a join and that join is the unique fusion of the regions. In fact for any non-empty set of regions W, $\Phi(\vee W)$ is the interior of the closure of $\cup\{\Phi(w): w \in W\}$. But the representation does not always preserve joins, as the following example shows. In this case we suppose the aether œ is infinite and the set representing it, Œ, is the set of all quadruples, but the example copuld be restricted to the case in which Œ is an open ball of coordinate quadruples. The set of quadruples has a left open half $L = \{<t,x,y,z>: x < 0\}$ and a right open half $R = \{<t,x,y,z>: x > 0\}$ with a hyperplane $H = \{<t,x,y,z>: x = 0\}$ separating them. So Œ = $L \cup H \cup R$ Given Tarski Continuum, L represents a region u, R represents a region v and H fails to represent, because it is not open. The interior of the closure of $L \cup R$ is $L \cup H \cup R$ = Œ. So œ = $u \vee v$, even though Œ = $L \cup H \cup R \neq L \cup R$. A similar result holds if all regions are represented by sets of quadruples of 4 dimensions. Both these hypotheses violate Weak Supplementation, as Hud Hudson notes (2005: 50-56).

Should we prefer hypotheses with greater restrictions on regions? In that case we should prefer Tarski Continuum to the hypothesis that the regions correspond to all non-empty open sets, which in turn should be preferred to the hypothesis that they correspond to all four-dimensional regions. And in that case all these hypotheses should be preferred to the

Orthodoxy that all non-empty sets correspond to regions. Or should we argue that the fewer and simpler the constraints the better? That would support the Orthodoxy. Neither, I say! For both ways of arguing assume there exist possible points corresponding to all coordinate quadruples and that our hypothesis is about which of these possible points are actual. We should not forget that the sets of coordinate quadruples are representations of the parts of œ and so hypotheses about these sets do not describe the intrinsic structure of the aether. The criteria of simplicity apply to the intrinsic structure not to the way that structure is represented.

3. The Axiom of Choice and Banach Tarski.

The Axiom of Choice is intuitive, as shown by the way that it requires practice even to notice its use. But the occurrence of non-measurable sets (e.g. sets of coordinate quadruples lacking any Lebesgue measure, even 0 or ∞) is also somewhat surprising, and Solovay has shown that if set theory with the Axiom of Choice is consistent then so is set theory in which there are no non-measurable sets (1970). I ask, then, how rejecting the Axiom of Choice would affect the investigation of the structure of the aether. Consider, for instance, the Orthodoxy that any non-empty set of coordinate quadruples represents one and only one region, and so, in particular, every singleton represents an point. Given realism about sets of quadruples, there is a fact of the matter as to whether the Axiom of Choice holds. If it does, then the celebrated Banach Tarski theorem also holds.[42] Hence there are disjoint regions b_1, b_2, b_3, b_4, b_5; c_1, c_2, c_3, c_4, and c_5 such that each b_j is represented as having the same shape and diameter as the corresponding c_j, and yet the join of the b_j is a ball of 1 cm radius lasting for 1 second while the join of the c_j is a ball of 2 cm radius lasting also for 1 second. This result implies that not all regions have a hypervolume, provided we assume: (1) that regions that are represented as congruent have approximately equal hypervolume – within 5% is far more accurate than required; and (2) that the hypervolume

42 See Wagon 1985) for an account of the Banach Tarski theorem. In (Forrest 2004) I use Banach Tarski to argue against the Orthodoxy.

of the join of *finitely* many disjoint regions is the sum of their hypervolumes.[43] For in that case, if all the regions have a hypervolume then that of the ball lasting one second would equal that of one of twice the radius, to within 5%, which is absurd. The assumption that regions represented by congruent sets of quadruples have volumes within 5% follows from the requirement that the diameters are represented to within 1% ± 1/K, and the obvious assumption that K is greater than 1000. For then the error in hypervolumes is less than $(1.1)^4$%, which is less than 5%. Curmudgeons who insist all regions are connected may still obtain regions to which no hypervolume can be assigned. They will differ from b_1, b_2, b_3, b_4, b_5; c_1, c_2, c_3, c_4, and c_5 by regions of zero hypervolume, but they will not be disjoint.

The Banach Tarski theorem does not hold if we suppose Solovay's Axiom in place of the Axiom of Choice. In that case the Orthodoxy implies that we may assign a hypervolume to any region, as we intuitively expect. The Orthodoxy then has an additional intuitive advantage over the two continuous exemplars, Borel and Arntzenius Continuum. For on the Orthodoxy all regions are composed of aether points using the *one* basic operation of summation. By contrast, on the hypotheses of Borel and Arntzenius Continuum regions are composed of fundamental regions ('globules') using *two* basic operations, summation and difference.

Initially, then, the Orthodoxy combined with Solovay's Axiom might well be taken as a serious contender for the structure of the aether, and defeat the Axiom of Choice. I have, however, two reasons for rejecting this defeater, leaving the Axiom of Choice as a defeasible but *undefeated* intuition.

My first reason for taking the Axiom of Choice to be undefeated is that the case for there being some non-measurable sets, and hence the case against Solovay's Axiom requires only a special, and especially in-

---

43 Typically we suppose that (hyper)volume is represented by Lebesgue measure, which satisfies *countable* additivity. It is noteworthy, therefore, that the Banach Tarski theorem only requires (hyper)volume to be finitely additive, which is intuitively much more secure than countable additivity.

tuitive, instance of the Axiom of Choice.[44] For that axiom has been assumed at just one point in the argument, namely that given a suitable equivalence relation there exists a *cross section* – a set that contains just one member from each equivalence class. To assume that for *any* equivalence relation we could find a cross-section would be tantamount to assuming the Axiom of Choice itself. But the especially intuitive case being considered is that in which the equivalence relation is 'natural' in sense that Lewis uses that term, as opposed to 'artificial'. (Here an artificial term is one that would not be used to state a hypothesis if we are to assess its simplicity.)

The other reason why the Axiom of Choice is an undefeated intuition is that its potential defeater, the Orthodoxy without the Axiom of Choice, is not as attractive as might initially appear. The Orthodoxy attracts us because, if it is correct, all regions are sums of points, which permits the assertion that all other regions depend for their existence on points. Because points may reasonably be taken to be simples this is in accordance with the thesis that everything depends ontologically on simple things, their properties, and their relations. That may in turn be

---

[44] We could use the proof of the Banach Tarski theorem to illustrate this but an easier result due to Giuseppe Vitali suffices to show where an instance of the Axiom of Choice is assumed. Consider a circle, of unit radius, and say that two points are equivalent if one may be obtained from the other by a rotation by a whole number of radians. That is, two points on the circle are equivalent if for some integer n, there is a path of length n going around the circle, perhaps many times, connecting the two points. Now consider a set X containing just one member of each equivalence class. For any integer n, let $X(n)$ be the result of rotating X by n radians. Because there are 2π radians in a full circle and because π is irrational, $X = X(n)$ if and only if $n = 0$. The whole circle is the union of all the $X(n)$. If X has a countably additive measure invariant under rotation, such as Lebesgue measure, then the whole circle has length that is the sum of countably many equal quantities, namely the measures of the $X(n)$. Let this measure be η. Then 2π is a countable infinity times η. This is impossible. For a countable infinity times η is the least upper bound of all $N\eta$ for positive whole numbers N. If $\eta = 0$ (or infinitesimal) the least upper bound is 0 (or infinitesimal, respectively). Otherwise it is infinity, but in neither case can it be 2π.

argued for by means of the intuitive premises that (1) every composite thing depends on the parts of which it is composed and how they are related, and (2) there cannot be an infinite regress of dependence relations. (This is the less obscure half of Kant's Second Antinomy.) Now suppose we reject the Axiom of Choice in order to save the Orthodoxy. Then we should rely on the intuition that there is a measure of the quantity of every region, namely hypervolume. For it was only the occurrence of non-measurable sets that might have made us reject that intuition. Within the scope of that intuition we have a further intuition expressed b the following principle.

Dependent Quantity:

> If something depends ontologically on some disjoint parts then its quantity is independent of the relations between those parts.

Suppose, then, that b and c depend ontologically on some disjoint parts, the Bs and the Cs, respectively, all of which are intrinsic duplicates. And suppose there is a one to one correspondence between the Bs and Cs. Then the quantity of b should equal that of c. So if regions depend ontologically on the point parts of a continuous aether, we may derive a conclusion known to have been false at least as far back as Duns Scotus, namely that regions whose points are in one to one correspondence have the same quantity (Solère 2010: 5-6.) Scotus used the example of two concentric circles. To be sure the intuitions relied upon in this argument are not very firm but they are strong enough to undermine the case against the Axiom of Choice.

In addition, I note that thesis that all regions are of maximal dimension, which I am taking to be 4, itself has some initial appeal in the context of continuous aether. In the next chapter I undermine that intuition, replacing it with the thesis that all regions depend on those of 4 dimensions. While I do not attach much weight to that intuition I note that when combined with the thesis that everything exists in virtue of simple things, their properties, and their relations, it completes the antinomy and, for what it is worth, weighs against continuous theories of the aether rather than supporting the Orthodoxy.

4. Some examples of the method of representation

The method of representation enables us to state various hypotheses about which sets of quadruples do the representing. Here are some examples, starting with the exemplars, the first two of which are discrete hypotheses.[45]

Point Discretion

The most straightforward case, Point Discretion, is that in which the aether is the sum of points that are represented by the quadruples of integers. We can generalise, although that case suffices for the discrete analog of Special Relativity. More generally, then, we suppose there is a set of quadruples D and a positive number δ such that if $\xi \in D$, $\eta \in D$ and $\xi \neq \eta$ then $||\xi - \eta|| > \delta$. It follows that if D is infinite then it is countable. The hypothesis is that all and only the non-empty subsets of D represent regions.

A variant on Point Discretion would be to consider, for instance, points represented by the quadruples of rational numbers in Œ. More generally, the points may be represented by a set of quadruples D* that is countable but for which there is no positive number δ such that if $\xi \in D^*$, $\eta \in D^*$ and $\xi \neq \eta$ then $||\xi - \eta|| > \delta$. In fact we might assume that D* is dense in Œ, that is, its closure is the whole of Œ as in the case of the quadruples of rational numbers.

Another variant is the infinite dimensional one in which points are represented by sequences of real numbers, for instance by infinite sequences all but finitely many members of which are zero.

Granulated Aether

By Granulated Aether I mean Extended Simples and its variants. On these hypotheses there are said to be significant small (presumably Planck scale) regions, the granules, represented using a set C of pairwise dis-

---

45 See (Van Bendegem 2009) for a survey of discrete theories of Space.

joint, non-empty open connected subsets of Œ.[46] It follows that C is either finite or countably infinite. Extended Simples states that all and only the unions of members of C represent regions. So every region is the sum of extended simples, which are the granules. [47] In that case the aether is connected if the interior of the closure of the union of C is connected.

Is the aether that constitutes our universe connected? I have no objection to island universes not connected to our part of the universe. It is then a matter for stipulation whether we consider them parts of *our* universe. I stipulate that they do not. The variety of disconnectedness that concerns my investigation into the structure of the aether is that which might occur at some small, presumably Planck, scale. We could, for instance, hypothesise aether simples that are represented by sets of diameter ½ whose centres are quadruples of integers. So there would be gaps between them. Either this would just be a silly way of representing Extended Simples or it would imply the triviality of the topological relation of adjacency (touching), because no two simples would touch. Assuming we have need of topology that would be a disadvantage. So in neither case need we bother with this variant.

Some readers might prefer to represent aether simples as the sets of quadruples that are the closures of the members of C, but because all we

---

46 A non-empty open set is said to be *connected* if it is not the union of two disjoint non-empty open sets.

47 As Varzi (2009) notes, several philosophers have recently defended the idea of simples that are extended not just in time but in Space. In particular I note Ned Markosian (1998), who, however, argues for the principle that every maximally connected thing is a simple, which would have the unhappy consequence that the aether has no proper parts. Simons (2004) defends simples from the Geometric Correspondence Principle (Simons 2004: 372), which I shall discuss in Chapter Three. David Braddon-Mitchell and Kristie Miller (2006) propose a connection between extended simples of what I call the aether and contemporary physics. See also Hudson (2006), Kris McDaniel (2007a), and Ted Sider (2007). One of the aims of this work is to present a disjunction between Extended Simples and a symmetric continuous aether hypothesis. I shall argue, however, that contemporary physics tends to support the latter.

are doing is *representing* regions there is no genuine difference between theories corresponding to the choice of representations. For the difference, if there were one, would concerns the boundaries of the aether simples, but a boundary would be a proper part, and so there are no boundaries of simples.

Of special interest is the case in which the members of C are polytopes, that is, higher dimensional analogs of polyhedra. In that case I say the aether simples are *polytopic*. For instance they might be represented by the 4 dimensional analogs of tetrahedra, the pentatopes. Diagram One illustrates the difference between Point Discretion and Extended Simples in a 2 dimensional case in which the extended simples are represented as right-angled triangular regions.[48]

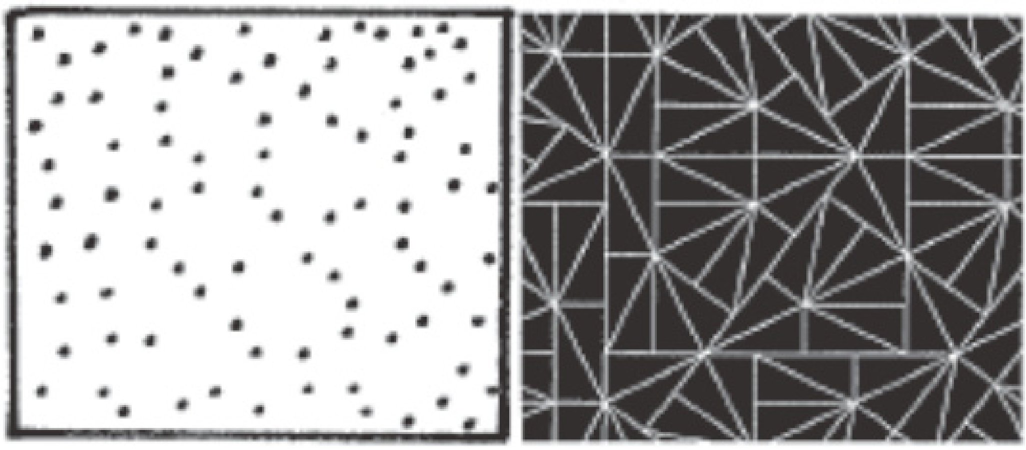

Diagram One

The most interesting variant on Extended Simples is Pseudo-set Granules, a hypothesis in which closed cells represent atoms but these atoms are not simple because the hyperfaces, faces edges and vertices of the cells also represent atoms, with only the vertices being simple.[49] This

---

48 Based on the Pinwheel Tiling (Radin 1995),

49 I call this Pseudo-set Granules, because we could develop a theory of sets, or as I called them pseudo-sets, in which the pseudo-singleton of u has u as its greatest proper part (Forrest 2002). On this theory the granule is a pseudo-singleton whose 'member' is the sum of all its hyperfaces.

fails to satisfy Weak Supplementation, but will turn out to have compensating attractions.

Borel Continuum

The Borel subsets of Œ are the members of B, where B is the intersection of all those subsets X of Œ that contain the open sets and where X is closed under the operations of countable union and the taking of complements. It follows that B contains all closed sets. Every singleton is closed, so B contains every finite and every countably infinite set of coordinate quadruples. One of the important features of Borel sets is that a Lebesgue measure (representing hypervolume) can be assigned to them. So the weird division used in the Banach Tarski theorem involves sets that are not Borel. The Borel Continuum is the hypothesis that all and only the non-empty Borel sets represent regions.

The Borel Continuum is a hypothesis in which there are points, that is, point regions. Variants that likewise have points are the *Orthodoxy* that all the non-empty set of quadruples represent regions, and the Lebesgue Continuum that all and only the non-empty measurable sets represent regions. Assuming the Axiom of Choice, I have noted that the Orthodoxy suffers from Banach Tarski trouble, and that neither the Borel nor the Lebesgue Continuum obey the principle of arbitrary fusion. They do, however, satisfy Weak Supplementation and countable fusion.

Another variant on Borel Continuum is that in which the regions are represented by all non-empty Gδ sets that is, countable intersections of open sets. All Gδ sets are Borel sets and they are of mathematical interest because every Lebesgue measurable set differs from a Gδ set by a set of measure zero. I reject this hypothesis of Gδ Continuum as inferior to Borel Continuum for two reasons, both to do with taking complements. The first is that Gδ Continuum is not Boolean. For we will be able to find a region u such that for any region v disjoint from u there will be a region w disjoint from u of which v is a proper part.[50] The se-

---

50 To prove this first note that the intersection of two Gδ sets is a Gδ set. Therefore if v is disjoint from u the representing sets of coordinate quadruples V

cond is that if Gδ Continuum held then complex regions would be composed of simpler ones as meets. Initially meets might be thought no worse than complements as a way in which some regions depend on others. But, although neither taking complements nor countable meets is an initially intuitive way of grounding some regions in others, a case can be made for taking complements, provided we are realists about absences. No metaphysician should take absences as fundamental entities, but the idea of the absence of u depending on region u is fairly intuitive. Then the difference u – v can be understood as the result of combining u with the absence of v, and, in particular ¬v = œ – v is a combination of the whole of the aether with the absence of v. The intuitive idea here is that the absence of v 'eats up' a v-shaped hole in the aether, so adding the absence of v subtracts v.

Arntzenius Continuum

This hypothesis differs from the previous ones in that it is most perspicuously stated as a many-valued representation. First we consider an equivalence relation on the Borel sets B, that of *almost identity*. This use of 'almost' is standard in mathematics and means that something holds except on a set of measure zero. So X is almost identical to Y if they differ by a set of Lebesgue measure zero, that is if X – Y and Y – X are both of zero measure. Arntzenius Continuum is obtained by representing the regions by equivalence classes of Borel sets. That is, Borel sets X and Y represent the same region just in case they differ by a set of measure zero. All and only the sets of positive measure in B represent regions. The restriction to sets of positive measure is required because a set of zero measure is equivalent to Ø and so would represent the fictional empty region ø.

---

and U respectively will have empty intersection. If $V \subset Œ – U$, there is some point p, represented by quadruple ξ, in neither u nor v. Then $W = V \cup \{\xi\}$ is a Gδ set representing w and w is disjoint from u. It follows that the required result holds provided there is a (non-empty) Gδ set U such that Œ – U is not itself a Gδ set. For example, let U be the set of all quadruples whose first coordinate is irrational.

Every measurable set differs from a Borel set by measure zero and so we could as easily have represented the regions as equivalence classes of measurable sets.[51] Using the Axiom of Choice, we could also have selected just one Borel set from each equivalence class, with Ø selected from the class of sets of zero measure, to obtain a set B* that represents the regions using a single-valued function, but that is messy.

To characterise the representation we also need to define the diameter and hypervolume of an equivalence class Z of Borel sets, which must approximate that of the region represented. Since every member of Z has the same measure, the hypervolume of Z is just the measure of any member of Z. The diameter of Z is the greatest lower bound of the diameters of members of Z. That is, we may ignore any set of measure zero when considering diameters.

The representation of Arntzenius Continuum is more complicated than that of the other hypotheses, but it would be a mistake to infer that it is a complicated hypothesis. It has the slight advantage over Borel Continuum of being a classical mereology and has the advantage over the Orthodoxy of assigning hypervolumes to all regions. It has the further advantage over Borel Continuum that hypervolume is a faithful measure of quantity. That is, there are no two regions whose difference is of zero hypervolume.

This last advantage needs to be understood correctly, because I have deemed that the infinitesimal is zero. Consider, then, the situation in a Borel Continuum in which there are regions u and v with u a proper part of v such that v – u has ‘zero’ hypervolume.[52] That could be interpreted as a case in which v - u is of infinitesimal hypervolume. The advantage of Arntzenius over Borel Continuum may then be expressed as a dilemma facing Borel Continuum. Either the measure of quantity fails to record any difference between u and v, because ‘zero’ means ‘literally zero’

---

51 In fact the measurable sets may be characterised thus: If W is measurable there are Borel sets X, Y and Z such that W - X $\subseteq$ Y, X - W $\subseteq$ Z, and Y and Z have measure zero.

52 Or consider the case in which u is a proper part of v and u and v have the same finite hypervolume.

or we must resort to infinitesimals. The first horn is uncomfortable because intuitively there should be a measure of the quantity of the most fundamental stuff of which the universe is made. When this stuff was identified with matter the quantity was thought to be mass, but the aether replaces matter, and so hypervolume is the measure in question. It should not, therefore, assign literally zero to any difference. The second horn is also somewhat uncomfortable because we have an intuitive dislike of the (actual) infinity and the (actual) infinitesimal, as shown by the widespread and, I say, deserved, acclaim for the way mathematicians such as Augustin-Loius Cauchy and Karl Weierstrass developed a rigorous theory of the calculus free from infinities and infinitesimals.

This dilemma is not a conclusive refutation, which is why I called its horns uncomfortable not sharp, but, together with the advantage of classical mereology, it provides the case for preferring Arntzenius Continuum over Borel Continuum. In Chapter Three, I shall show that an apparent advantage of Borel Continuum, arising from Chapter Two, is not genuine.

Open Gunk

A variant on Arntzenius Continuum may be obtained by taking only the non-empty *open* subsets of Œ to represent regions, and once again treating as equivalent two sets that differ only by measure zero. In that case, if Z is an equivalence class of open sets, $\cup Z \in Z$ and I call the union of an equivalence class of open sets a *maximal* open set. So the regions may be represented in a one to one fashion by the non-empty maximal open sets. This is a complete Heyting mereology.[53] I mention it largely to exhibit the range of hypotheses. It is one of a variety of hypotheses in which regions are represented by open sets and in which summation is represented by union. I call these *Open Gunk* hypotheses. They include Locale Continuum, the hypothesis I have already mentioned as rejected by Hudson because it violates Weak Supplementation. On it the regions

---

53 The proofs of this result and of the previous result that for any equivalence class Z, $\cup Z \in Z$, rely on the fact that in the topological space of n-tuples of real numbers the union of a set X of open sets is the union of some countable subset of X.

are represented in a one to one way by *all* the non-empty open subsets of Œ. I call it Locale Continuum because of its connection with the theory of *locales* (Johnstone 1982). In Chapter Three I shall explain why I reject them.

The method of representation enables is to consider other hypotheses such as the following.

Sparse Continuum

On this hypothesis, we start with the set S of open convex subsets of Œ and consider all complements and finite unions of S as representing regions. More precisely, consider S* the intersection of all the sets Z of sets of quadruples that: (1) $S \in Z$; (2) if $X \in Z$ and $Y \in Z$ then $X - Y \in Z$ and $X \cap Y \in Z$. Then the regions are represented by all and only the non-empty members of S*.

Tarski Continuum

Finally I should mention Tarski's hypothesis that all and only the non-empty regular open subsets of Œ represent regions(1956). This is not an Open Gunk hypothesis because summation fails to correspond to union of sets. It is a classical mereological theory.

Curmudgeon variants

The restriction that all regions are connected may be expressed easily enough if the regions are represented faithfully by open sets. For a non-empty open set is connected if it is not the union of two disjoint non-empty open sets. More generally, a region u is connected unless there is some positive δ such that u is the sum of regions v and w and no region of diameter less δ than overlaps both v and w. So if X is the set of non-empty sets of quadruples used to represent the regions ignoring curmudgeons, then curmudgeons may represent the regions by $X^c \subset X$, where, if$U \in X$, $U \in X^c$ unless there is some positive δ and U is the union of $V \in X$ and $W \in X$ such that no $Y \in X$ of diameter less than δ intersects both V and W.

Restrictions on the representation of the whole

I have assumed that Œ $\subseteq$. $\Re^4$. A variant is obtained by increasing the number of dimensions from 4 to k ≥ 9 and representing curved aether without approximation in $\Re^k$. We can do this by relying on John Nash's Imbedding Theorem (Nash 1956). In that case, Œ is a 4-dimensional differentiable manifold. We may use the diameter as defined in $\Re^k$ but the measure is that for the manifold not that for $\Re^k$, which would assign zero to œ.

Another restriction is to the union of the boundaries of non-overlapping polytopic open sets, of the sort used to represent Extended Simples. In that case the aether has only 3 dimensions. We could go further and take Œ to be the union of the 1 dimensional edges of the polytopes. Yet another restriction, based on discrete Time would be to take Œ to be the union of countably many space-like aether hypersurfaces corresponding to a discrete sequence of moments.

Fine structure

The method of representation directs our attention to the fine structure, if there is any, because that is what the representation ignores. I shall not be concerned much with fine structure because Ockham's Razor inclines me to deny there is any. Moreover, hypotheses that I will be comparing can be stated without commitment to whether or not there is fine structure. It is, however, instructive to state some hypotheses concerning fine structure as part of my overall purpose in this chapter, namely showing that there is a great variety of ways the aether might be, for all we *know*.

One kind of fine structure would occur if each region x contains a certain quantity of aether and if x has parts that have the same extension as x but contain only some of x's aether. In that case even regions that have no proper subregions might fail to be simple having parts consisting of a proportion between 0 and 1 of the aether. This will be briefly considered in Chapter Five.

Another kind of fine structure occurs if the regions are represented, perhaps faithfully, as sets of quadruples of non-standard numbers, say of the form $x + x'\iota + x'\iota^2$ where x, x' and x'' are standard real numbers and ι an infinitesimal. Call this the *fine representation*. Then there is a *coarse* *representation* obtained by ignoring the infinitesimals. There is an equivalence relation that holds between regions u and v just in case they have the same coarse representation. If there are points, that is regions of zero coarse diameter, then the sum of all the points in a given equivalence class is a maximal point. Consider, for example, the maximal point with coarse representation {<0,0,0,0>}. Its fine representation is $\{< t'\iota + t''\iota^2, x'\iota + x''\iota^2, y'\iota + y''\iota^2, z'\iota + z''\iota^2>: t', t'', x', x'', y', y'', z', \text{and } z'' \text{ are standard real numbers}\}$. So in this case the maximal point has the structure of an eight dimensional space. If we allowed powers of ι up to and including the k^{th}, then the maximal points have 4k dimensions.

Another way of hypothesising fine structure is to start with a fine representation of the regions as sets of quintuples <w; t, x, y, z> where $<t, x, y, z> \in \Re^4$ but $w \in$ M, some set with a topology on it, for instance a manifold. In that case if there are points, then each maximal point has the structure of M, which could be of any number of dimensions. Each maximal point might for instance have the structure of a circle so *w* is an angle.

This survey of fine structure is not intended as exhaustive but merely sufficient to add to the enormous variety of structures we can conceive of the aether having. It should be noted that if there is fine structure a point is not an atom. Because I shall largely ignore fine structure the important question is not whether the aether is atomic or gunk but whether it has points or not.

5. Extent.

I am delaying most considerations of structure in addition to the mereological, but some intuitions about extent and quantity will be required for the next chapter. We may measure the extent of a region by means of a diameter function, which will assign either ∞ or a non-negative real number |x| to each region x. It must satisfy:

Diam 1: If $x \leq y$ then $|x| \leq |y|$; and

Diam 2: If x and y overlap, and if they have a join $x \vee y$, then $|x \vee y| \leq |x| + |y|$.

It is worth pausing to think about plausible hypotheses in which the join is not a fusion to check that Diam2 is rightly stated using the join not the sum. On the Intervals Only hypothesis about Time, joins are not always fusions but Diam 2 is still intuitively correct, as it would be if, contrary to intuition, we supposed that all regions are convex.

Unless the aether is discrete, there is an argument against treating diameter as a primitive property, analogous to the arguments of Robin le Poidevin (2004) and Tim Maudlin (2007: 86-89) against a primitive relation of distance. The argument is that Diam2 is a necessary constraint on the scalar diameter property, but one that lacks explanation. Necessities should not be multiplied more than is necessary, so we should not treat Diam2 as primitive. My response is that we start with the linear ordering due to a relation of *being of greater extent*, about which we can say that no part has greater-extent than the whole.[54] Then, paralleling the methods of measuring extent with rulers, we quantify this obtaining a diameter function (arbitrary up to a scale factor).[55] Regarding Diam2, suppose we had a measure of extent, call it quasi-diameter, that satisfies Diam1 but not Diam2. Then we could define an associated distance relation between regions u and v as the greatest lower bound of the sums of the quasi-diameters of chains of overlapping regions joining u to v.[56] Then the diameter of a region w is the least upper bound of the distances between parts of w. This method of characterising the diameter could fail if the distance function assigned zero to regions that did not have zero quasi-diameter, but otherwise we obtain a measure of extent that satisfies Diam1 and Diam2. Because Special Relativity implies that the diameter is frame-relative it is unlikely, though, that diameter, or even the greater extent relation, is a fundamental feature of the aether. So the details of deriving the diameter function as a way of measuring extent

---

54 It is a *linear ordering* in the sense that: (a) it is transitive; and (b) the derived relation of being neither of greater nor of lesser extent is an equivalence relation.

55 There is a well developed theory of how quantitative scales can be derived from linear orderings given various plausible constraints (Suppes and Zinnes, 1963) (Luce and Suppes 2002).

56 Each region in the chain overlaps the next, the first overlaps u, and the last overlaps v.

should not concern us too much. For the present it suffices that we have intuitions about diameters.

I define a *point* as a region of zero diameter, noting that atoms need not be points and points need not be atoms. To ensure that all atoms are points it suffices that the following holds.

Covering: Given any positive integer N and any region x, x is covered by (i.e. is part of the join of) regions of diameter less than 1/N.

Clearly Covering fails for the aether simples posited in the Extended Simples hypothesis.

There are some (mereo-)topological relations that might be considered primitive but may be defined in terms of diameter if there is a diameter function, and if Covering holds. I say that the *adjacency* or touching relation holds between regions x and y (x@y) if for every positive integer N there is some region z overlapping both x and y such that z has diameter less than 1/N. I say that regions x and y are *separated* if they are not adjacent. I say that x is an *interior part* of y, x « y if, for some positive integer N, every region of diameter less than 1/N that overlaps x is part of y. It follows that if x « y and y and z are disjoint then x and z are separated. And if x and z are separated then x « ¬z. If the mereology is classical, then every region is a complement, and so x « y if and only if x is separated from ¬y.

Adjacency is an equivalence relation on points. To ensure that the join of an equivalence class of points is a point it suffices to assume that any two regions have a join together with the following:

Remote Parts: Given any region x and any positive integer N, x has parts y and z such that any region overlapping both y and z has diameter at least |x| - 1/N.

6. Quantity

The quantity of a region is measured by its hypervolume, a function assigning either 0, ∞ or some positive real number to all or some of the regions. It is intuitive that every region has a hypervolume, but this intuition will be delayed until the next chapter. Meanwhile I am permitting the Orthodoxy, as I call it, in which the regions correspond in a one to one fashion to all the non-empty sets of coordinate quadruples. Assuming the Axiom of Choice, the Banach Tarski theorem then shows that not every region can be assigned a hypervolume. Therefore, when I state the rules governing hypervolumes I am restricting them, if necessary, to those regions that have a hypervolume (abbreviated to hvol). We have:

Superadditivity: If regions u and v are disjoint and w is a region such that u and v are both parts of w, then hvol(w) ≥ hvol(u) + hvol(v).

If u and v have a sum, u+v, *Subadditivity* should also hold, namely that hvol(u+v) ≤ hvol(u) + hvol(v).

I note, however, that summation is a rather special operation on regions, and in a survey of hypotheses we should not prematurely assume the existence of sums. In place of Subadditivity I propose therefore:

Approximate Subadditivity: If u < v, then given any positive integer N, there is some positive integer n and pairwise disjoint regions $u_0, u_1, \ldots, u_n$ such that $u_0 = u$, $u_j < v$ and $\sum\{hvol(u_j): j = 0, \ldots, n\} \geq hvol(v) - 1/N$.

Conclusions of survey

This chapter has shown that there are an enormous variety of available hypotheses about the structure of the aether, between which to choose. My procedure in this work is to begin with some intuitive, considerations and then examine the kinds of extra structure that the aether might be supposed to have, including the structures posited by current physics. My starting point will be to show that we have inconsistent intuitions about the aether and to generate a range of four hypotheses that give up just one intuition each, the exemplars.

2. Conflicting Intuitions

In the previous chapter I exhibited the variety of hypotheses about the aether's necessary structure. If there was a well supported unified physics, then my task would that of the 'under-labourer' namely one of interpretation and I could begin by describing the physics. As it is, the physics, especially quantum gravity, is as speculative as the metaphysics, and I find it convenient to discuss the structure of the aether a priori, using metaphysical intuitions, and then consider what further structure might be required by physics. Now, one way to assess the hypotheses a priori would be to go through them one by one listing their pros and cons. The sheer number of hypotheses is an obstacle to the survey-and-judge method of reaching a rational conclusion. Instead I proceed by first arguing, in this chapter, that our intuitions are inconsistent and then, in the next, discussing which to abandon.[57]

In this chapter, then, I state eleven premises, each of which is either directly intuitive or supported by intuitions. They are not jointly consistent so at least one premise will have to go.

1. Four stipulations

1. The sum of the aether regions being considered, œ, is taken to be a local region, say, the aether within one (light) year of Earth now. That is because in this and the next chapter I concentrate on the local structure of the aether, not such global features as whether it has the shape like a Klein bottle. As a corollary, when it is required I shall be assuming that hvol(œ) is finite.
2. I distinguish the hypervolume of a region from the Lebesgue measure of a set of coordinate quadruples. If we represent regions as sets of coordinate quadruples the hypervolume should not differ from the measure of the representing set except perhaps by a relatively small amount due to curvature or perhaps the discrete character of the aether. (See Chapter One, Section 2.)
3. By a *point* I mean an aether point, that is a region of zero diameter, where diameter is a measure of extent and is subject to the following stipulation.

---

[57] This chapter differs only slightly from 'Conflicting Intuitions about Space, Space-time, or the aether' (Forrest 2012), itself derived from my paper at the 2007 MMT conference in Rutgers University.

4. If ξ is a diameter or hypervolume, then either $\xi = 0$ or $\xi = \infty$, or there is some positive integer N such that $1/N < \xi < N$.

This last stipulation avoids some technicalities but is not intended to exclude fine-grained ways of measuring extent and quantity, ones that allows infinitesimal differences.[58] It is merely a stipulation about when a measure of extent or quantity is to be called a diameter or hypervolume respectively, namely that infinitesimal differences are to be ignored. As a consequence, when I argue that the sum of regions of zero hypervolume has zero hypervolume, I shall not rely on the, in any case dubious, principle that infinity times zero is infinity, for infinity times an infinitesimal could be anything from a smaller infinity to a larger infinitesimal, including all finite values.

---

[58] The reason for this last stipulation is not just that it avoids unnecessary technicality. There is, in addition, a good technical reason for it. One of our intuitions about hypervolume is that every region has a hypervolume, although maybe it is 0 or ∞. If we reject this, as we could decide to, then there might be regions that can be assigned a hypervolume zero given my stipulation but could not be assigned a finer measure of quantity. To illustrate this, suppose classical mereology holds and so every set of regions has a unique sum. Suppose also that the aether is the sum of maximal points u that are in one to one correspondence with coordinate quadruples $\langle t,x,y,z \rangle$. Suppose, however, that maximal points are not simple regions, but each is a copy of the Euclidean sphere and hence the sum of simple points in one to one correspondence with the members of $K= \{\langle x^*, y^*, z^* \rangle : x^{*2} + y^{*2} + z^{*2} = 1\}$. Then we may assign an infinitesimal measure of quantity ι to a maximal point u, and zero to a simple point. There is a fine-grained assignment of quantity to any point v that is part of a maximal point u, provided v corresponds to a *measurable* set $K_V \subseteq K$. The fine-grained measure of the point v equals the measure ('area') of K_p multiplied by the factor $\iota/4\pi^2$. But there are many subsets of K that have no measure.(This a corollary of the magical Banach Tarski theorem Wagon 1985). So if some proper part, v, of u is a point represented by such a non-measurable set, v has zero or infinitesimal extent but has no fine-grained quantity. Hence the intuition that any region of zero diameter has zero hypervolume is not adequately expressed by saying that any region of zero or infinitesimal extent is of zero or infinitesimal quantity.

2. Grounding and an intuition underminer

One of the central ideas of metaphysics is that some things are the ontological *grounds* of others, which are then said to exist *in virtue of* them or to *depend ontologically* on them. I have already mentioned Fine's example that the singleton {b} exists in virtue of b, not vice versa (1995). Another example is the claim that surfaces of objects exist in virtue of objects.[59] If we are considering regions, then we might well generalise this to the claim that all regions exist in virtue of 4 dimensional ones.

Now consider, for example, the initially intuitive thesis that the only regions are those of 4-dimension.[60] Hud Hudson rebuts a similar thesis by appealing to Weak Supplementation (Hudson 2005: 50-56). But I think the intuition that all regions are of 4 dimensions may be totally undermined, and hence no longer considered. It is undermined by diagnosing it as the misunderstanding of the intuition that all regions exist in virtue of those of maximal dimension. For the term 'real' is equivocal between: (1) existing; (2) existing independently of any (creature's) mind; and (3) being fundamental, that is not existing in virtue of any other things. Once we grasp this, the intuition that there are no surfaces may be replaced by the intuition that surfaces are not real in sense (3).

A corollary of the undermining of the maximal dimension intuition is that points are not directly counter-intuitive. That is, if the existence of points clashes with intuitions, that clash must be mediated by an argument from those intuitions.

The moral is to be careful to distinguish grounding from existence claims when formulating intuitions.

3. Intuitions about connectedness

Mereological curmudgeons rely on the intuition that every region is connected in some sense, which remains to be explicated. Hence, they

---

59 The alternative defended by Roy Sorensen (2008: Ch 2) is that to refer to a surface of an object is to make a context-dependent reference to a thin but 3 dimensional part of the object. This is plausible in many contexts, but my example assumes surfaces have zero volume.

60 Or whatever the number of dimensions the best physics requires, maybe the 10 of String Theory.

say, regions have a sum only if some connected region is a fusion of them. If I attempt to exhibit a disconnected region, such as the USA, the response will be that I have not exhibited one but rather several regions (Alaska, the sum of the 48 contiguous states, etc). This intuition may be undermined as in the previous section. For we tend to conflate the stronger claim that all Xs are Ys with the weaker claim that those Xs that are not Ys exist in virtue of the Ys. In this case we tend to conflate the claim that all regions are connected with the weaker claim that all regions exist in virtue of connected ones.

The intuition that survives the undermining is that every region that is not connected depends ontologically on connected parts. Now we have a further intuition that quite generally the Ys are grounded in the Xs only if the Ys are the sum the Xs suitably related. So we expect every region to be the sum of connected regions. Moreover that conclusion is itself intuitive. A fortiori, every region has a connected part.

I require merely a rather weak explication of connectedness in terms of separation, where, in the continuous case:

> Two regions y and z are *separated* if there is a positive integer N such that any region overlapping both y and z has diameter greater than 1/N.[61]

Here is a more complicated characterisation, which also covers the discrete case:

> Two regions y and z are *separated* if there are real numbers ξ and η that are possible values of the diameter function, such that $\xi < \eta$ and any region overlapping both y and z has diameter greater than η.

I define a *connected* region as one that is not the sum of two parts that separated. So I have:

Premise One: Connected Parts:

> Every region has a part that is connected.

---

61 The stronger topological characterisation of connectedness will be too strong in the case of Arntzenius Continuum, Tarski Continuum and Open Gunk. For on this topological characterisation a region is connected if it is not the sum of two disjoint open regions. It turns out that no region of Arntzenius Continuum is connected in this sense.

Note that any point is connected. This follows from the stipulation that we ignore infinitesimal differences. Hence Premise One holds on any point based hypothesis. It will not hold on Arntzenius Continuum.

Next we have:

Premise Two (Universal Summation):

Any regions have a sum.

My case against the curmudgeons removes the obstacles in the way of Universal Summation. On the other hand it is not directly intuitive. It may be supported in two ways. The first, due to Lewis (1991: 81-7), is to submit that there is no extravagance in positing sums along with their summands. For it is they and they are it. The second is the extravagance of restricting rules capable of universal formulation. (Call this 'Frege's razor', because Frege assumed that set-theoretic axioms were unrestricted until Russell derived his paradox.) Extravagance is not the only fault an ontology can have and if Universal Summation turns out to be inconsistent with firmer intuitions then we may reject it. Perhaps it is worth noting that Universal Summation follows from Finite Summation together with the principle that any regions sharing a common part have a sum. It is the latter that is more vulnerable.

Notice that I have stated Universal Summation regardless of whether the regions form a set or a proper class. Partly that is because I think that such considerations are intuitively irrelevant, but it is also in part because the structure that I am investigating is the local structure, say of some region œ of finite diameter, and so the only case that I need worry about is that in which the parts of œ form not a set but a proper class Ω. But that would tend to undermine intuitions about sums only because of some supposed intuition that every proper class $X \subseteq \Omega$ has too many members to have a join. If there is such an undermining intuition it is itself rebutted by the way the proper class Ω itself has a join namely œ.

4. Arbitrarily small regions

We have some intuitions concerning arbitrarily small regions. The following is typical.

Premise Three: The Diameter Hypervolume Nexus

> Given any positive integer M, however large, there is some positive integer N such that any region of diameter less than 1/N is part of a region of hypervolume less than 1/M.

I am not assuming as a premise that every region has a hypervolume so I cannot yet exclude the case in which however large N is there are regions of diameter less than 1/N with no hypervolume.

The restriction to regions of arbitrarily small diameter helps prevent scruples arising if we think of the aether as curved. For even a curved manifold is locally like a flat one, where, with one proviso, we have no trouble with the idea of diameter. And the restriction to the arbitrarily small requires considering what happens locally. The proviso is that the diameter of a region is relative to the choice of a frame of reference.[62] Fortunately if the Diameter Hypervolume Nexus holds in one frame it holds in every other.

The next three intuitions are about suitably shaped arbitrarily small regions, which I shall call *globules.* Precision is not important here, but given some topological structure we may characterise the globules as regions that are topologically equivalent to hyperballs. Instead we could characterise them as convex regions of finite diameter of 4 dimensions, or, in general, the same number of dimensions as the aether itself.[63] The reason why precision is not important is that the term ‘globule’ functions like the variable X in a Ramsey sentence: there are some regions, the Xs, such that all the premises hold if the word ‘globule’ is replaced by ‘X’.

Premise Four: Finite Globules

> There is a globule of finite positive diameter and finite positive hypervolume, namely œ.

---

62 The diameter corresponds to a metric in which the norm (length) of a 4-vector with coordinates <t, x, y, z> is $\sqrt{t^2 + x^2 + y^2 + z^2}$.

63 Region y is *between* regions x and z if for any part y′ of y there are parts x′ and z′ of x and z, respectively, such that diam(x′+y) + diam(y′+z′) ≤ diam(y′+z′) + diam(y′). A region w is *convex* if for any parts x and z of w any region y that is between x and z is also part of w. A convex region u of finite diameter is of *less than N dimensions*, if for any positive integer m there are N regions of diameter less than 1/m such that every convex region containing all N as parts also contains u as a part.

Premise Five: Arbitrarily Fine Covering.

If N is a positive integer and if u is a globule of finite diameter then u is part of the sum of at most a countable infinity of globules of diameter less than 1/N.

On the assumption that there are only finitely many dimensions a finite number of globules would suffice. This premise, like the next, relies on an extrapolation from regions we can imagine being occupied to much smaller regions that we can only imagine by scaling them up. Such 'zooming in' is not very reliable because it is a case of inductive generalisation, which we know to be not merely fallible but only moderately reliable in the absence of buttressing by means of a theory. (In so far as we treat the results of zooming-in as intuitive this is a case of defeat by explication..)

Premise Six: Arbitrarily Thin Boundaries

If u is a globule of finite diameter, and N is a positive integer, then there is a region v of hypervolume less than 1/N such that any connected region w that overlaps u, but is not part of u, overlaps v.

Convex regions such a cubes or balls satisfy this principle as would regions whose boundaries are fractals, but a region represented by an open dense set of finite measure in the topological space of all quadruples of real numbers would not have an arbitrarily thin boundary. For the boundary of its representing set would be of infinite measure. Such regions would not count as globules.

By Arbitrarily Fine Covering, any simple region must be a point, contrary to idea that there are extended simple regions. It might be objected that Greek atomists believed in extended simple objects, perhaps casting doubt on Arbitrarily Fine Covering. On the contrary, the intuition that any extended thing should have left and right halves, or more generally be divisible in each direction was produced as an argument against the Greek atoms, thus supporting my claim that Arbitrarily Fine Covering is intuitive.

5. Intuitions about hypervolume

It is true by definition that if x is part of y then the hypervolume (and diameter) of x cannot exceed that of y. In addition, we have the following, in which hypervolume is abbreviated to hvol.

Premise Seven: The Mereology Hypervolume Nexus

(1) If x is part of y, which is part of z, if x and z have hypervolumes, and if hvol(x) = hvol(z) then y has a hypervolume and, hence hvol(x) = hvol(y).

(2) If y is part of z and z is of zero hypervolume then y has a hypervolume, so y has zero hypervolume.

From this and the Diameter Hypervolume Nexus (Premise Three) it follows that no point has positive hypervolume. Rather than think of this as derived from Premises Three and Seven I suggest that the intuition that no point is of positive hypervolume is part of the intuitive basis for the Diameter Hypervolume Nexus.

If we had initially assigned finer grained hypervolumes that made infinitesimal distinctions, then the situation might arise in which a region to which no hypervolume may be assigned was part of one of infinitesimal hypervolume. The Diameter Hypervolume Nexus nonetheless assigns to such regions zero hypervolume when we ignore infinitesimals, in accordance with the stipulation made in Section 1.

Another premise relating mereology with hypervolumes is:

Premise Eight: Hypervolume Supplementation.

If x is part of y and hvol(y) > hvol(x) > 0, then y has a part z disjoint from x such that hvol(z) > 0.

This is a weakening of the principle of Weak Supplementation, which is used along with the existence of sums of any regions to argue for classical mereology.. Weak Supplementation states that if x is part of y then y has a part z disjoint from x. The idea behind this is that some of y is left over when x is removed. I submit that although Weak Supplementation is moderately intuitive, Premise Eight is more so. For in the case in which hvol(x) = hvol(y) when you take x away there might well be nothing left. Likewise if hvol(x) = 0 taking it away might have no effect.

It could be objected that the similar diffidence shows that we should not assume that z is disjoint from x. Maybe they have a meet of zero hypervolume. I concede that a weaker version of Premise Eight would be even more intuitive, but since Weak Supplementation is itself intuitive and Premise Eight more so, I maintain that Premise Eight is itself intuitive – with one proviso. The proviso is that, strictly speaking, Premise Eight is intuitive only within the scope of the supposition of Premise Two. For otherwise a positive amount of stuff left over might be scattered in an infinity of regions of zero hypervolume.

It is worth recording the weaker, and even more intuitive, version of Premise Eight as:

*Premise Eight** (Weak Hypervolume Supplementation):

> If x is part of y and $hvol(y) > hvol(x) > 0$, then y has a part z such that $hvol(z) > 0$, and either x and z are disjoint or $hvol(x \wedge z) = 0$.

It is intuitive that the whole is no greater than the sum of the parts. We have, therefore:

Hypervolume Subadditivity

> If there is a family of regions, u_θ, $\theta \in \Theta$, each with hypervolume $hvol(u_\theta)$, which have a sum w, then w has a hypervolume and hvol(w) does not exceed $\Sigma_\theta vol(u_\theta)$, the sum of the hypervolumes.

By Premise Two the sum w is guaranteed to exist.

Hypervolume Subadditivity may conveniently be divided into two cases. The first concerns the sum of two regions and is the familiar:

Premise Nine: Finite Subadditivity

> If x and y have hypervolumes, $hvol(x + y) \leq hvol(x) + hvol(y)$.

Next there is the transition from the finite to the infinite case and, not quite as intuitive, the transition from the countable to the uncountable case. For this purpose I need the following

Premise Ten: Limiting Hypervolumes

> Consider some totally ordered regions (i.e. given any two of them one is a part of the other) each of which has hypervolume less than or equal to k. Then, if the regions have a sum this sum has hypervolume less than or equal to k.

From Premises Two, Nine and Ten we obtain:

Countable Subadditivity.

> Consider a sequence of regions, u_1, u_2, etc such that, for all j, u_j has hypervolume. Then the hypervolume of the sum is no greater than the sum of the hypervolumes of the u_j.

Countable Subadditivity has intuitive support, illustrated by the following cautionary tale against trying to get something for nothing. You think Countable Subadditivity is false and so fill up a sequence of regions with *twin gold*, where in the world being considered twin gold has arbitrarily small parts that are still twin gold, but twin gold is as precious as gold is in our world. You have to fill the regions up more and more rapidly so as to fill them all up, but this is no problem because they are smaller and smaller. Worried by curmudgeons you are also careful to connect each of the globules to the next by a thin thread using a total of only 10cc. The sum of the hypervolumes of the regions is 100cc, so at the end of the operation you have used only 110 cc of twin gold. You believe, though, that Countable Subadditivity fails and hope that as a result there is a region w full of twin gold, with hypervolume 200 cc of twin gold. But you got it wrong: it turns out that the regions in question do have a sum which is full of twin gold but it cannot be assigned any hypervolume at all! Both your hope and the unfortunate outcome are absurd. So this is not a metaphysically possible world after all.

Premise Ten is intuitive, but may also be argued for by relying upon a principle that I have not taken as one of the eleven premises, namely that every region has a hypervolume. Consider the sum of totally ordered regions each of hypervolume no greater than k, and assume this sum has a hypervolume. If that hypervolume is greater than k we may ask, as we did when considering Hypervolume Supplementation, where the extra stuff comes from. It comes from nowhere, which is counter-intuitive. To appreciate the intuitive strength of the assumption that every region has a hypervolume – perhaps zero, perhaps positive but finite, or perhaps infinite – consider the idea of a liquid, it might as well be molten twin gold, which, unlike gold has point parts that are made of the same stuff. Although contrary to science we can easily imagine that there is such stuff. (And by the principle that necessities are not to be multiplied more than is necessary, we should be reluctant to suppose that such stuff is impossible.) Now consider a region with no hypervolume but full of this liquid. Then pour it out into a beaker. Does it vanish? Are you drowned in an infinity of the stuff, or what? The supposition is almost as absurd as the story about trying to increase the quantity of twin gold. The intuition here is that hypervolume is the measure of the quantity of some stuff, in this case the aether. Because the measure is coarse-grained there may well be regions of zero hypervolume but not ones without a hypervolume at all.

6. Premise Eleven and the Axiom of Choice

Lastly I propose the following.

Premise Eleven:

> If every totally ordered set of parts of u of zero hypervolume has a join of zero hypervolume then the join of all the parts of u of zero hypervolume is itself of zero hypervolume. (A set of parts X is totally ordered if for every u and v in X either u is part of v or v is part of u.)

Premise Eleven is a hypothetical and should not be confused with the categorical principle that any sum of a totally ordered set of parts of zero hypervolume is of zero hypervolume. For Premise Eleven holds of the Orthodoxy (combined with the Axiom of Choice) but the categorical principle does not. This premise follows from Zorn's Lemma, itself provable using the Axiom of Choice, which I defended in the previous chapter.

7. **The sum of all the zeros**

If a region u has a part of zero hypervolume then, I shall argue, the sum of all its parts of zero hypervolume is a part of zero hypervolume. By the combination of the Mereology Hypervolume Nexus (Premise Seven) and the Diameter Hypervolume Nexus (Premise Three) any point is of zero hypervolume. Hence this result excludes any theory that implies that every region is the sum of points, including the Orthodoxy, namely that regions are represented by all the non-empty sets of coordinate quadruples.

First we show that u has a maximal part of zero hypervolume, v. For given any totally ordered set W of parts of u with zero-hypervolume, by Premise Two W has a sum w, and by Premise Ten w has zero hypervolume. So by Premise Eleven u has a maximal part of zero hypervolume, v.

Now consider any regions of hypervolume zero – call them the Z. By Premise Two they have a sum u. We have just shown that u has a maximal part of zero hypervolume, v, that contains all the parts of zero hypervolume and hence the Z. So v is an upper bound of Z and v is part of u. But by the characterisation of the sum as a least upper bound that is also a fusion, u = v, establishing the result that any sum of regions of zero hypervolume is itself of zero hypervolume.

I anticipate the response that for two millennia geometers have relied upon points, and so we should reject any intuitions that might be deployed against point-based theories. I have three responses. First, I am committed to abandoning one of the eleven premises together with one of their supporting intuitions. So point-based theories are not excluded by my result. Second, the intuitions are about the aether and I have already conceded that that Space-time is point-based. Finally, suppose an appeal is made to the tradition of Euclidean geometry, to the development of Cartesian coordinates, and to the Nineteenth Century perfection of geometry, all in support of the Orthodoxy. Then I reply that the Banach Tarski theorem was a surprising discovery and that had it been discovered by, say, Archimedes, the Orthodoxy would not have been so widely accepted.

This result that any sum of regions of zero hypervolume is itself of zero hypervolume excludes any point-based theory of the aether, such as the Orthodoxy that every set of coordinate quadruples represents a region. But that is not all. It excludes the existence of *supersponges,* of which the Menger-Sierpinski sponge might be familiar Weisstein 2009). (By a supersponge for a region u of positive hypervolume, I mean some part v of u that is of positive but strictly less hypervolume than u, such that there is no connected part of u that is of positive hypervolume but disjoint from v.)

Now consider a region u of positive hypervolume and v a supersponge for u. Let w be the sum of all the parts of u of zero hypervolume. By Premise Eight (Hypervolume Supplementation) there is a part x of u, disjoint from v, of positive hypervolume. Suppose x has a properly connected part z of positive hypervolume. Hence it has no connected parts, contradicting Premise One (Connected Parts). Therefore to complete the argument for the clash of intuitions I need only argue for the existence of supersponges.

8. Supersponges

First suppose that all regions are represented by sets of coordinate quadruples. In that case for any positive integer M, for any integer N ≥ M, and for any coordinate quadruple of rational numbers χ, there is a region whose representing set is included in the ball center χ radius 1/M and includes the ball center χ radius 1/N. Assume for convenience that œ has hypervolume at least 1 unit. That there is a supersponge may be demonstrated as follows.

There are countably many planes obtained by varying all but one of the coordinates of the quadruples of rationals, and they may be arranged as a sequence σ_k k = 1, 2 etc. Then there is a region w_k whose representative includes σ_k but whose total hypervolume is positive but less than $(½)^k$. The sum of all the w_k is a supersponge for the whole of œ. Moreover if we could assume that the set of coordinate within 1/N of a quadruple of rationals χ represents a region then we do not need the Axiom of Choice to show the supersponge exists.

More generally, we can show the existence of supersponges without assuming coordinate representation, by relying upon Premises Four, Five, Six, Nine and Ten. First, by Premise Four (Finite Globules) we may suppose that u is a globule of finite hypervolume and diameter. Let its hypervolume be 1 unit. Then by Premise Five (Arbitrarily Fine Covering) for every positive integer N, u is part of the sum of at most a countable number of globules of diameter less than 1/N. So we may consider a sequence of globules u_j, j = 1, 2 etc such that for any positive integer N, u is part of the sum of the members of the sequence of diameter less than 1/N. Then for every globule u_j there is a region v_j as in Premise Six (Arbitrarily Thin Boundaries) of hypervolume no greater than $(½)^{j+1}$. By Countable Subadditivity (which follows from Premises Nine and Ten) w, the sum of the v_j, has hypervolume no greater than ½. Because globules are of positive hypervolume, the region w is the required supersponge.

Summary of the conflict of intuitions

From the eleven premises we obtain two results: (1) regions of zero hypervolume have a sum of zero hypervolume; and (2) there exists a supersponge for some region of finite positive hypervolume. Now (1), (2), Premise Three (Diameter Hypervolume Nexus) and Premise Eight (Hypervolume Supplementation) form an inconsistent tetrad, so one of the Premises One to Eleven must be false. In the next chapter I use this result to discover some hypotheses that do not run counter to more intuitions than necessary.

3. Which Intuition to Abandon?

Some readers may be exasperated, considering it obvious which premise or premises to abandon. My response is that often when an intuition is undermined we are tempted to deny that we ever had it. This is especially the case when we have a *Hume moment,* reflecting upon an intuitive belief and then thinking, 'Why would I believe *that*?' I take this to be a *partial* undermining of a genuine intuition. If readers disagree, assuming that in such cases we never had the intuition, then one of the eleven premises of the previous chapter should be taken not so much as supported by intuitions but as a premise that the benighted author thought was. The problem remains, though, 'Which premise lacks support?'

By considering how we might undermine one or other of the eleven premises, we have a fairly systematic way of finding a list of theories from which to choose the best – or to express indifference between, if that is the correct conclusion. I say 'fairly' systematic because an intuition that I judge robust might perhaps only seem that way to me because of a lack of insight and imagination.

After a preliminary, concerning the fine structure of regions, I consider a range of hypotheses that violate only the less firm intuitions. Then I provide some, albeit inconclusive, considerations in favour of my claim not to have omitted a hypothesis that would have deserved serious attention.

I am prepared to declare a provisional winner, namely Point Discretion. In Chapter Six I argue against it because of a non-locality problem that arises if we combine it with the premise that approximations to Special Relativity are nomologically possible.

1. Fine structure

In order to restrict the number of hypotheses about the structure of the aether, I shall ignore the fine structure that can arise in those hypotheses, such as Point Discretion, Borel Continuum, and Lebesgue Continuum, according to which there are points. I have already mentioned fine structure in passing, but here I need to note its impact on the exemplars.

By a *point* I mean an aether point, a region of zero diameter; by a *simple* I mean something without parts. It is not, I say, intuitive that every point is simple, although if there are complex points then we might well say that they have infinitesimal rather than zero diameter, but I am assimilating infinitesimals to zero.

If some points are complex then maybe all are, and so points could be gunk, that is without simple parts. Whether or not that is the case, we may define an equivalence relation on points. For by the standard properties of diameters, together with the firm intuition that if two regions overlap they have a sum, it follows that being part of the same point is an equivalence relation on points.

If the members of an equivalence class have a join then it is itself a point, for if a region is the join of points then its diameter should be the least upper bound of the distances between its point parts.[64] In that case every point is part of a maximal point, where a maximal point is not part of any other point. Maximal points must be disjoint. Obviously if all points are simple then all points are maximal points.

The equivalence relation between points may be extended to regions that are represented as sets of coordinate quadruples. In that case regions are equivalent if they are represented by the same set.

The simplifying assumption that fine structure is to be ignored needs to be remembered in two contexts. The first is that what we consider a point might, perhaps, in fact be many equivalent points. The second occurs because I mean by a *discrete* hypothesis one in which every region of finite extent is deemed to have finitely many parts if we ignore fine structure. Therefore some discrete hypotheses such Point Discretion have variants in which maximal points have infinitely many parts.

Once we make the simplifying assumption that equivalent regions are identified, then there are four exemplars: a point based theory with finitely many points in any region of finite diameter, namely Point Discretion; one with uncountably many points, namely Borel Continuum; one with extended simples of positive hypervolume, namely Extended

---

64 The distance between points u and v is just the diameter of the sum of u and v, or, if there is no sum, the greatest lower bound of the diameters of regions containing both u and v.

Simples; and a gunk theory, namely Arntzenius Continuum, in which there are no points but regions of arbitrarily small but positive hypervolume and diameter.

2. Point Discretion

The first two premises are intuitive once curmudgeons have been silenced. In any case we may state curmudgeon variants if we have to. Later I shall consider hypotheses that violate Premise One (Arntzenius Continuum) and Premise Two (Borel Continuum). Initially more promising is the violation of Premise Three, permitting Point Discretion. That hypothesis is easily stated if we ignore any fine structure. In that case we may treat the points as simples and Point Discretion states that every region of finite quantity (hypervolume) is the sum of finitely many points. A variant that might conceivably be proposed in the context of General Relativity is that in which the points are not simple but each has the structure of Minkowski Space-time.[65] But if we ignore fine structure, this will not be distinguished from Point Discretion.

The initial problem with Point Discretion is the way that some or all points would, it seems, have to have positive hypervolume, which is not merely contrary to Premise Three but absurd. For surely the hypervolume of a region of diameter x can be no greater than that of a hypercube of side x, which is x^4. So the hypervolume of a point must be less than $(1/M)^4$ for any positive integer M and hence zero.

There are two ways we might try to undermine this intuition. The first is to suggest that quantity only emerges as the number of points becomes so great that they get blurred as it were. I find this unsatisfactory because the quantity was intended as an answer to the question, 'How much?' and if we are realists about the aether there should be an answer to the question in all cases, or at very least in all except perversely complicated cases, such as might arise with non-measurable sets of points but not in the discrete case.

---

65 This would enable us to define velocities for particles, but I do not think we need velocities. If aether is discrete what we need is a propensity for one state to be followed by another, not a velocity.

Instead, therefore, I undermine the intuitions about hypervolume by denying that hypervolume is always the right way of answering the quantity question, ‘How much?’ In the case of Point Discretion the right way of answering it is to alter the question to ‘How many?’ I agree that every region has a quantity, but the quantity is the number of points in it. In that case hypervolume emerges as a convenient substitute for the correct but unknown measure.[66] For large regions of not too complicated a shape the emerging hypervolume is approximately proportional to the correct measure of quantity.

This use of approximate hypervolume in place of the precise measure of quantity is familiar in other contexts. Consider a typical fungible, something made up of grains but for which we use a mass term. Consider again the good Biblical example of a bushel of wheat. Prophets were rightly angry with merchants who cheated the poor giving them less than the bushel they had paid for. But they never complained that the buyer had paid for all the gaps between grains and got only air in return. The moral is that in circumstances where the precise question is ‘How many?’ we often make do with ‘How much?’

Point Discretion is initially an excellent hypothesis, but I shall argue, in Chapter Five, that there are problems in explaining how Point Discretion can approximate a differentiable manifold in the case in which the aether is not necessarily symmetric. The symmetric case, discussed in Chapter Seven, is interesting in that Minkowski Space-time has a discrete analog with the points arranged in a regular fashion. Although this seems promising I consider it refuted by the Intermittent Particle Objection. With regret, therefore, I shall reject Point Discretion.

3. Granulated Aether

Premise Four is almost trivial, merely asserting the existence of some region that is a globule. It might, for instance, be represented by a convex set of coordinate quadruples. I shall argue, that Premises Five (Arbitrarily Fine Coverings) and Premise Six (Arbitrarily Thin Boundaries),

---

66 Unknown because the number of points per unit of hypervolume (cm^4) is not known, even though we may suppose it is of roughly the order of magnitude 10^{120}.

however, are not especially firm because they depend on extrapolation from the larger to the smaller. Therefore we should not dismiss the *granule* hypotheses as too improbable. On these hypotheses some (small) extended regions have positive hypervolume but are not the sum of two regions of positive hypervolume – these regions may be called *granules.* Granule hypotheses may be illustrated by the case in which the granules are represented by tesseracts (4 dimensional analogs of cubes) each touching 80 others, or, for sake of visualisation, the case in which there are only 2 dimensions and the granules are square 'tiles', so each touches 8 others. Another example is the simplicial granulation in which all the granules are represented by pentatopes (also known as 5-simplices), the 4 dimensional analogs of tetrahedra.

It is natural to think of these granules as having extended locations in a *continuous* Space-time. That may be a reason for rejecting a granule hypothesis for Space-time itself. If so, this gives us a further reason not to identify the aether with Space-time, but no reason to reject granules.

Something should be said about the *shape* of the granules if Extended Simples is correct. Strictly speaking because simples have no proper parts they are shapeless, for our ordinary concept of shape can be analysed in terms of spatial relations between proper parts. Granules, however, have shape, in an analogical sense, specified by the adjacency relations between granules. Here *adjacency* is a primitive symmetric relation that is represented as follows:

> If granules x and y are represented by sets of coordinate quadruples X and Y, then x and y are adjacent if and only if the closures of X and Y intersect.[67]

Informally, the granules x and y are adjacent just in case either x = y or the representing regions overlap or they touch.

Now the shape of things in the ordinary sense explains how they fit together. Hence what I am proposing is that individual granules have shape in an analogical sense, namely a capacity to have various adjacen-

---

[67] A different range of hypotheses is obtained if we modify this, to obtain: If granules x and y are represented by sets of coordinate quadruples X and Y then x and y are adjacent if and only if the closures of X and Y intersect in a 3 dimensional set.

cy relations.[68] That this is a *capacity* may be shown by considering a coherent but implausible structure for the aether in which it consists of simple granules represented by unit tesseracts whose centres are quadruples of integers, but where there is a granule just in case not all these integers are even. So we have missing granules. In this case some granules have unrealised capacity for adjacency.

An interesting variant on Extended Simples is *Skeletal* Granulated Aether, in which all aether atoms are of 1 dimension. The extended simple is then replaced by the sum of its edges, namely its one-dimensional parts. This would have considerable advantage if we needed to consider a (connected piece of) aether of varying dimension for then the same simples could make up cells of differing dimensions. As far as I know, however, no current physical speculations are of this kind.

On the other hand Skeletal Granulated Aether has a problem with the characterisation of the hypervolume of the cells. Because it is a granule theory it is intuitive that there should be a way of assigning hypervolume to regions. But the only available measures of quantity (either the number of simples or the sum of their lengths) fail to distinguish between a granule and a region of zero (or infinitesimal) hypervolume. For instance in the 2 dimensional case a square would have quantity equal to 4 times the length of the side, but so would a straight line segment made up of 4 simples. Therefore we should favour the versions of Granulated Aether in which the granules are of maximal dimension.

One way in which granule hypotheses differ is by whether they postulate any lower dimensional regions than whole granules, and in particular points. The most straightforward, Extended Simples, is that in which all regions are sums of the granules, which are therefore aether simples. On this hypothesis there are no lower dimensional regions. If we suppose, as we should, that granules are globules this violates Premise Six (Arbitrarily Thin Boundaries) as well as Premise Five. It seems to me, however, that since both these premises are based on the same zoom-in extrapolation, violating just one of them is no great advantage.

---

68 Analogy in Aristotle's *pros hen* sense, according to which the typical causes – as well as effects – of health, that is, certain kinds of diet, certain kinds of exercise regime, etc are themselves said to be healthy.

If, however, there are any lower dimensional regions then it would be ad hoc to deny that there are ones of all lower dimensions including points. In this way we obtain another granule hypothesis, Pseudo-set Granules. This is based on a non-classical complete Heyting mereology in which some regions are *pseudo-singletons*, where I say that v is the pseudo-singleton {u}* if it has a greatest proper part u, that is, all proper parts of v are parts of u. According to this, the only simples are points – the 'vertices' of the granules. But in some cases the sum, x + y, of two points x and y, has the pseudo-singleton {x +y}*, which is an edge, namely the edge joining x to y. Likewise in some cases the sum u + v + w of three edges u, v, w etc has a pseudo- singleton {u + v + w}*, which is a 2 dimensional face; in some cases four or more faces have a sum which has a pseudo-singleton, which is a hyperface; and five or more hyperfaces have a sum which has an pseudo-singleton, which is a 4 dimensional granule. The granule, although not a simple is an *atom*, in the sense that it is not the sum of proper parts.

One advantage of Pseudo-set Granules over Extended Simples is that we might hanker after a non-dispositional non-analogous *shape* for the granules. This is provided by the relations, in this case mereological, between its parts. Another is that adjacency just is overlap, for adjacent granules share a boundary. Pseudo-set Granules does not violate Premise Six but it does violates Premise Eight. (Consider the sum of two adjacent granules x and y. Because they share a boundary, we cannot remove x from x + y without removing the whole of y.) It does not, however, violate the (even) more intuitive Premise Eight*

Another hypothesis, which I call Hybrid Granules, is that in which there are not just simple points but also simples of higher dimension such as the edges excluding the two points at their ends, faces excluding their vertices and edges, and so on. Then the highest dimensional simple is the granule excluding all its hyperfaces, faces, edges and vertices. This would be represented by an open set of coordinate quadruples, while the cell consisting of a granule with its hyperfaces etc is represented by the closed set. Like Pseudo-set Granules, it promises to give us a non-dispositional account of the *shape* of the granules: the adjacent lower dimensional regions providing a sort of external skeleton that determines its shape. Moreover, the only premise it violates is Premise Five.

For all but the most straightforward, Extended Simples, species of Granulated Aether, when two granules meet, the points, edges etc could be doubled up so that those adjacent to one simple four-dimensional region are not adjacent to any other. The difference between these variants is just one of fine structure, ignoring which I have described three granule hypotheses. Although there are others these are enough to illustrate the range.

My aim is not to reject just any of the eleven premises but to find one that can be partially undermined. I have already noted in Chapter Two the way Premises Five and Six seem to depend on zoom-in, the extrapolation from the larger to the smaller. I now consider in more detail the consequent undermining of Premise Five (Arbitrarily Fine Covering). It may be illustrated using the Left Half/Right Half intuition that a granule should, like any other extended region, be the sum of a right half and a left half. We could express this as the intuition that in every dimension there is a dichotomy of any region extended in that dimension.

Intuitive assertions may be analysed, in which case they cease to be intuitive, but become conclusions of inferences that may then be criticised. Accordingly, I analyse the intuitive basis of Premise Five, and in particular the Left Half/Right Half intuition, as well as Premise Six, as a consequence of the zoom-in extrapolation, namely that we can just zoom-in on ever smaller regions treating them as if they were macroscopic ones. Zooming-in has sometimes been defeated in the history of science, notably in the transition from classical to quantum mechanics, but maybe it is robust enough to remain in cases where it has not been defeated. I am no sceptic about the probability of induction so I grant that zoom-in holds, *other things being equal*. What I am sceptical about is that inferences by induction in the narrow sense are robust enough to defeat other probabilistic considerations. I conclude, therefore, that Premises Five and Six have some support but are not robust. Hence rejecting it is a sensible way of resolving the problem stated in the previous chapter.

I have just diagnosed the Left Half/Right Half intuition as a consequence of a zoom-in extrapolation, but Peter Simons has offered an alternative explanation, namely the initial plausibility of what he calls the Geometric Correspondence Principle (Simons 2004: 372). This principle states that any extended object has parts corresponding to the parts of the region it occupies. As he notes in his defence of extended simples, reflection on the case of Space composed of points might make this seem less plausible. In addition, I have the following two objections to that principle. First, there is some conceptual novelty in my recommendation to distinguish the aether from the Space-time that depends for its existence on the aether and on relations between regions (of the aether). Hence the aether and this dependent Space-time are two explications of our more intuitive idea of Space-time. Therefore, I am entitled to explicate the Geometric Correspondence Principle as:

> Any extended object has parts corresponding to the parts of the aether that constitute it.

That principle is no threat to Extended Simples.

My second objection is that the Geometric Correspondence Principle itself presupposes the absolute theory of Space-time, namely that it is does not depend on the relations between the things we think of as located in Space-time. As indicated in the Introduction the case for absolute Space-time only holds if it is identified with the aether. Hence either the Geometric Correspondence principle holds even for Extended Simples or its presupposition of absolute Space-time is false.

Assuming we can partially undermine Premise Five by diagnosing the intuition as a case of zooming-in, a similar argument should, as I have said, undermine Premise Six and so give a slight provisional advantage to Extended Simples over Pseudo-set Granules, which violates Premise Eight, but not, the firmer Premise Eight*. In addition, Extended Simples has only one kind of atom, while Pseudo-set Granules has five (or generally one more than the number of dimensions.) Hybrid Granules only violates Premise Five, which is good, but like Pseudo-set Granules has 5 kinds of atom, indeed 5 kinds of simple. Provisionally, then the order of preference is: Extended Simples, followed by Hybrid Granules, with Pseudo-set Granules in third place.

According to Extended Simples, a granule has shape only in the analogical sense of its capacity to touch other extended simples. If readers dislike this, then they should prefer one of the alternative granule hypotheses. But I shall assume that such an account of shape is adequate.

There is the further question of whether the adjacency relation between granules is a quantitative one. In the 2 dimensional case we might propose that the length of the common boundary between the representatives of the 'tiles' as the degree to which they touch, with area replacing length in the 3 dimensional and volume replacing length in the 4 dimensional cases.[69] Initially we might think that one promising physical theory, Loop Quantum Gravity, supports some granulated aether hypothesis. Because that theory assigns areas and volumes it would seem to provide a case for degrees of touching. In Chapter Six, however, I argue that Loop Quantum Gravity does not support a discrete aether hypothesis, and that the most likely theory that does is Causal Set theory, which does not require degrees of touching. A straightforward appeal to economy, therefore, shows that we should reject degrees of touching.

Extended Simples is provisionally the best granule hypothesis because it is the only version with just one kind of simple and that also satisfies the intuitive principle of Hypervolume Supplementation. In fact Extended Simples retains Weak Supplementation because it is a classical mereological theory. Pseudo-set Granules obeys only the weaker and less elegant axioms of complete Heyting mereology.

There is, however, a well-known objection to all granule hypotheses, namely the Weyl Tile Problem (Weyl 1949: 43, Van Bendegem 1987, Forrest 1995, McDaniel, 2007b). This concerns the intrinsic metric of a 2 dimensional tiled Space, that in which the distance between tiles x and y is the smallest number of steps from tile to adjacent tile required to get from x to y. If the tiles are represented by congruent regular polygonal regions in a Euclidean plane (squares, triangles or hexagons), then the intrinsic metric fails to approximate the standard Euclidean distance.

A related problem (Van Bendegem 2009: 3.1) is the lack of isotropy. The pattern of representations of the tiles will be a regular one of equilateral triangular, square or hexagonal regions, and so the lines between the representations of the tiles will be distinguished directions: at

---

69 In the 4 dimensional case we could assign infinitesimal hypervolumes to 3 dimensional regions in the representation, infinitesimal squared hypervolumes to 2 dimensional ones etc.

60° in the case of triangles, at 90° in the case of squares, and at 120° in the case of hexagons.

In 2 dimensions we may solve both the metric and isotropy problems using the Pinwheel tiling, made up cells of the same intrinsic shape and size but with some the mirror images of others. (See Diagram One, based on Radin 1995) The resulting metric approximates the Euclidean metric and the tiling has no privileged directions[70]. Presumably, something similar holds in higher dimensions.

This is not the end of the problem, though. A regular pattern of tiles will result in lack of isotropy, contrary to the appeal of symmetry. But in the absence of a regular arrangement of granules it would be an astounding coincidence if in fact the granules align themselves so that the resulting aether is flat (or of some other highly symmetric shape). This is, of course, no problem unless we have reason to take the aether to be flat or highly symmetric. I obtain, however, the corollary that if the aether is symmetric then it is not made of granules. In Chapter Six, I shall argue that either it is symmetric or made of granules. Then the corollary shows the disjunction is exclusive.

The coincidence of the granules aligning to result in a flat aether would not be astounding if we thought of Space-time as independently existing. For in that case it would be a priori quite probable that Space-time was flat and the aether granules would have to fit into it. But because, I have argued, Space-time depends for its existence on the aether, we would have to think of the granules as having adjacency relations that, out of all the conceivable arrangements, happen to be result in flat aether.

---

70 I am indebted to Charles Radin for his correspondence on this topic. The Penrose tiling is not isotropic for there are only 10 orientations for the tiles (Radin 1995: 27). The pinwheel tiling is, however, isotropic in the sense that given any angle θ and any positive angle δ, however small, the proportion (i.e. limiting relative frequency as the distance from some given point increases) of tiles with orientations between θ and $\theta + \delta$ is $\delta/2\pi$. (Radin, 1995: 27).

Note on Space-time

I have asserted that, on continuous theories of the aether, Space-time is not the aether but a construct that depends for its existence on the aether. In the case of granulated aether Space-time would not be a construct but rather a fictitious space. First consider a single granule, and its associated Space-time. We could perhaps think of each point location *p* as the property that an object has if *p* is its centre of mass, a property that could belong to a sum of granules if the individual granules had masses. Whether or not we do so, we can associate a topological space associated with each granule. For instance, suppose the granules are represented by pentatopes (5-simplices). In that case, the corresponding space is topologically equivalent to Pent = $\{ \langle t, x, y, z\rangle : 0 \leq t \leq x+y+z, 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1\}$. The whole of Space-time is then the result of identifying parts of the boundaries of the various topological spaces corresponding to granules, where the identified boundaries correspond to the hyperfaces between adjacent granules.[71] This Space-time could be given a metric structure but that metric seems to be of no relevance since the aether has a spatio-temporal structure not a purely spatial one.[72] More important, in 6 or fewer dimensions there is an associated differentiable manifold structure specified up to diffeomorphism. (See Chapter Five.) I consider this ‘construction’ to result in a fiction because the occurrence of centres of mass requires a metric and the choice of metric depends on just how we represent the pentatopes. If the metric were of physical significance then it would be appropriate to give a structuralist account of the centres of gravity. As it is, I judge the centres of gravity to be fictitious properties.

---

71 I chose the example of 5-simplices rather than tesseracts here because it is the practice of algebraic topologists to consider simplices in this kind of context.

72 Within the Space-time occupied by a granule we could use the metric for a regular 5-simplex in the space of coordinate quadruples. Then the distance between points that are not in the same 5-simplex is the minimum that is in accordance with the triangle inequality.

4. The Orthodoxy and some variants

The hypotheses I have considered thus far are the discrete ones. I now turn to continuous theories of the aether and, because of its familiarity, begin with the Orthodoxy. The Orthodoxy is that regions are represented by, and stand in a one to one correspondence with, all the non-empty sets of coordinate quadruples. A region has hypervolume if and only if the representing set has a measure, in which case the Lebesgue measure of the representing set approximates the hypervolume (say within 5%). Moreover, every region is the sum of its point parts, represented by singletons, and classical mereology holds.

That the measure of the representing set approximates the hypervolume is motivated by the restriction to a region of slight curvature. If there is no curvature at all then we have exact equality. Approximate equality implies that a region has zero or infinite hypervolume if and only if its representing set has zero or infinite measure, respectively. So an point has zero hypervolume.

One of the variations on the Orthodoxy is to suppose that some points are not simple, but I have already assumed we are here identifying equivalent regions so this variant may be ignored unless it is required later.

Because all points have zero hypervolume, the Orthodoxy and its variants are incompatible with the conjunction of Premises Two, Ten and Eleven, as noted in the previous chapter. The Orthodoxy itself is incompatible with the conjunction of Premises Ten and. Eleven. Premise Ten may be partially undermined by an appeal to humility: should we trust our intuitions in the uncountable case? Fair enough, but there is still a whole heap of trouble for the Orthodoxy and its variants, based upon the Banach Tarski theorem, to prove which we only required one application of the Axiom of Choice, namely the *Cross Section Principle*:

> If R is an equivalence relation on the Ws there are some Ws, the Zs, such that every W is related to just one Z by R.

Moreover, as discussed in the previous chapter the application of this is to an equivalence relation that is natural and easy to think about. We are considering two rotations of a sphere, and points are equivalent just in case some finite sequence of these rotations in some order takes one to the other.

This application of the Banach Tarski theorem to the Orthodoxy is rightly called the Banach Tarski *paradox*. Intuitively there should not be such regions. The paradox immediately implies that there are regions to which no hypervolume can be assigned, since hypervolume must be finitely additive on any regions with zero hypervolume overlaps, and hypervolume depends only on shape and size. But I have an intuition that the quantity question 'How much?' must have an answer when asked of any region, and that the answer is in this case the hypervolume.[73] Moreover the equation 1 + 1 = 1 is not even approximately correct, so curvature does not remove the paradox.

While I respect tradition, the historical fact that the Orthodoxy acquired its status before the discovery of the Banach Tarski theorem should make us suspect that it would never have acquired that status if the paradox had been known to Archimedes or to Newton. In addition I am prepared to concede that Space-time is as the Orthodoxy states, while maintaining that the aether on which it depends is not. In Chapter One I argued against any defence of the Orthodoxy by rejecting the Axiom of Choice.

One interesting family of nearly orthodox hypotheses is obtained by rejecting Premise Two, *Universal Summation,* which is implied by the Orthodoxy. So we obtain hypotheses in which we suppose that a class of points, the Xs have a sum if and only if they are represented by sets of coordinate quadruples of some designated kind. One such variant of the Orthodoxy is *Lebesgue Continuum*: all and only Lebesgue measurable non-empty sets represent regions.

Other things being equal we should restrict general principles as little as possible. This would support Lebesgue Continuum. There is, however, a serious objection to this hypothesis. Consider a non-measurable set K of coordinate triples <x, y, z>, say one of those used in the Banach Tarski theorem. Then we form the product I × K, where I = {t: 0 ≤ t ≤ 1}, the unit interval. So I × K = {<t, x, y, z>: t ∈ I, <, x, y, z> ∈ K}. According to the Orthodoxy I × K represents a non-measurable

---

73 The alternative explication as 'How many?' cannot be correct in this case, because it gives the same answer in all cases, namely $2^{\aleph_0}$.

region existing for one unit of Time, but according to Lebesgue Continuum there is no region represented by I × K. And that is what we expect intuitively. But now consider J, an uncountable subset of I of measure zero, say the Cantor set. Then, J × K has zero measure and hence, according to Lebesgue Continuum, does represent a region, b It is strange, counter-intuitive even, to allow a weirder set J × K to represent a region when the less weird I × K does not.

A more incisive objection is obtained by noting that our intuitions about spatio-temporal things such as the aether are based on two different ways of extrapolating from the 3 dimensional case. We may consider the 4 dimensional analogs of the 3 dimensional intuitions, but we also retain those intuitions for the 3 dimensional time-slices of the 4 dimensional things. Hence the Banach Tarski paradox should show us that the Orthodoxy is false even when restricted to time-slices. But given any set of coordinate triples X the set of quadruples J × X would represent a region according to Lebesgue Continuum and so X represents a time-slice of a region.

To avoid these pathological yet measurable regions, I deny that summation is the only way in which more complicated regions can depend on simpler ones. In this way I obtain a variant on the Orthodoxy, which I call *Borel Continuum*, the thesis that the regions correspond to the results of starting with regions we are happy to consider fundamental, the globules, and then considering the result of both countable summation and the taking of complements. These are represented by the non-empty Borel sets of quadruples. Any Lebesgue measurable set of quadruples will differ from a Borel set by a set of measure zero.

Borel Continuum violates universal summation. In addition, the way I described it illustrates a defect. Intuitively, more complicated regions depend on fundamental ones such as the globules, by being their sum. But not all parts of Borel Continuum are sums of globules. That is because there are two ways in which more complicated regions depend on less complicated ones, summation and taking complements. In some cases, we are forced to say that the region u – v depends on u and v. Or, equally counter-intuitive, we abandon the thesis that the more complicated regions depend ontologically on the less complicated.

A cross section of a Borel subset of $\Re 4$ is always a Borel subset of $\Re 4$, so the objection to Lebesgue Continuum does not, however, arise.[74]

5. Point-free theories

Point Discretion and Borel Contfinuum are two point-based hypotheses about the aether. I now consider some hypotheses in which every region has positive hypervolume and every region is the sum of regions of less hypervolume. We may call these *gunk* hypotheses provided we I ignore fine structure.

These gunk hypotheses have an advantage over point-based continuous hypotheses such as Borel Continuum, namely that the measure of quantity, the hyper-volume, is *faithful*. That is, if region u is a proper part of region w and if u is of finite hypervolume then w is of greater hyper-volume. Or, to avoid the restriction to the finite case, we have a strengthening of Weak Supplementation:

> If u is a proper part of w then there is some proper part v of w disjoint from u, and v has positive hyper-volume.

Although not a firm intuition, I have an a priori preference for the position that hyper-volume should be faithful.

The first gunk hypothesis I have called *Arntzenius Continuum* after Frank Arntzenius who has proposed it (2008). On it the regions are represented in a one to many way by sets of coordinate quadruples of positive Lebesgue measure. That is, the region is represented by an equivalence class of sets. In this case sets X and Y are equivalent if the differences X - Y and Y - X are both of measure zero. Then region x is part of

---

74 Without loss of generality, we may take the cross section to be obtained by putting $x = 0$. Define a *cylinder-set* to be a set W such that <x, y, z, t> ∈ W iff <0,x,y, z, t> ∈ W. Then every cross section of an open set is open and so a cross section of an open cylinder set and a fortiori of a Borel cylinder set. Every union of cross sections of Borel cylinder sets, the Xs, is a cross section of a Borel cylinder set, namely the union of the Xs, and every complement of the cross section of a cylinder set Y is the cross section of a cylinder set, namely the complement of Y. Hence, by mathematical induction on the number of operations (union and taking complements) required to form the Borel set from open sets, every cross section of a Borel set is the cross section of a Borel cylinder set and so itself Borel.

region y if for any sets X and Y in the equivalence classes representing x and y, respectively, the difference X – Y is of zero measure. Because every Lebesgue measurable set is equivalent to a Borel set, we may likewise obtain it by identifying equivalent Borel sets of positive measure.

Arntzenius Continuum has some attractive mathematical features.

1. Not only does it satisfy classical mereology, but also the sum of a class Z of regions is the sum of some countable sub-class.
2. All the eleven Premises except Premise One (Connected Parts) hold.
3. The topology may be characterised by means of a family of open regions, where an open region is one that is represented by an equivalence class containing a non-empty open set of quadruples.
4. The union of all the open sets in an equivalence class is itself in that class, and therefore may be used to represent the open region. I call such a set U a maximal open set meaning that if U is a proper subset of an open set V then V - U is of positive measure.
5. The open regions generate all regions in the sense that the class of all the regions is the smallest class of regions containing the open ones, and closed under countable summation and the complement operation.

Because of (5) we do not need to worry about which equivalence classes of Lebesgue measurable sets represent regions. Every open region is the sum of globules so once we include globules and deny that there are two regions differing by one of zero hypervolume the result must be Arntzenius Continuum unless we reject classical mereology.

Arntzenius Continuum violates Premise One (Connected Parts), because the complement of a supersponge contains no connected parts larger than a point and Arntzenius Continuum is point-free. This complement of a supersponge is totally disconnected in the sense that it has no connected parts. Not only is that seriously weird but it is open to an objection based on the intuition that every region depends ontologically on fundamental ones that are not too complicated. Totally disconnected regions are clearly too complicated. For ease of exposition I assume that the fundamental regions are globules. It follows that the totally disconnected regions must depend on globules. But they are not the *sum* of globules for globules are connected. So they would have to depend in some other way on the globules. There remains the intuition that every region depends on globules. This holds provided we allow more than one operation by which some regions ground others, namely summation and difference.

In this respect Arntzenius Continuum is no worse off that Borel Continuum, and because of its other advantages I rank it higher. This comparison will survive all the further considerations that I shall discuss, so my disjunction 'Granulated Aether or symmetric continuous aether' is to be understood with the rider that continuous aether is more likely to be Arntzenius Continuum than Borel Continuum. It is worth noting, though, that Borel Continuum would be advantageous in one respect if the aether's absolute (i.e. frame-independent) temporal ordering did not result in the familiar light cone structure.[75] The reason for this is that the pointy nature of light cones provides us with surrogates for the points that Arntzenius Continuum lacks. In the absence of these point-surrogates the characterisation of the symmetric geometry of the aether is somewhat more complicated.. (See Chapter Seven for further details.) I am not sure how seriously we should take this one advantageous respect.

If we ignore all considerations of the hypervolume of a region, then we should take seriously Tarski Continuum, namely the hypothesis that the regions correspond to the non-empty *regular open* sets of quadruples, where a regular open, also called perfectly open, set is one that is the interior of its closure (Tarski 1956). This is a classical mereology in which every region is the sum of countably many globules.

If the only defect was that it violated Premise Ten (Limiting Hypervolumes) then we might have another fit of humility in the face of the infinite. Unfortunately, there is not even a finitely additive hypervolume on the Tarski Continuum. Moreover unlike the case of the Orthodoxy, to go against the intuition that every region has a quantity

---

75 If the aether has a Minkowski Space-time structure, or some other highly symmetric shape, such as de Sitter Space-time, then there is an absolute (i.e. frame-independent) temporal ordering, that in which x precedes y if x ≠ y but the future light cone from x contains y.

could not even be defended on the (question-begging) grounds that the regions correspond to non-measurable classes. Compared to the rather minor failing of Arntzenius Continuum I find this quite damning.

That there is not a hypervolume assignable to every region of Tarski Continuum follows, appropriately enough, from an analog of the Banach Tarski theorem for regular open sets, and one that does not require the Axiom of Choice (Dougherty and Foreman 1992). There are disjoint regular open sets U_1 to U_5 of total Lebesgue measure no greater than $4\pi/3$ (the hypervolume of a unit ball persisting for unit time), such that V, the smallest regular open set containing U_1 to U_5, has measure equal to 8n/3 the sum of two balls of unit radius persisting for unit time). We obtain a paradox by taking the regular open sets U_1 to U_5 and V to represent regions u_1 to u_5 and v.

Tarski Continuum is just one version of a family of hypotheses according to which the regions are represented by some or all of the non-empty open sets of quadruples. In Chapter Two I noted some others, the Open Gunk hypotheses such as Locale Continuum.. Suppose we restrict our fits of humility to the uncountable case and so we grant Countable Subadditivity for hypervolumes (i.e. the principle that the hypervolume of the sum of countably many regions cannot exceed the sum of the hypervolumes). Then these open region variants violate Premise Eight. For if all regions are represented by open sets so are the globules, and there will be enough globules for us to find a countable class of them, K, whose hypervolumes add up to ½, where I assume that hvol(æ) ≥ 1, and whose representatives are dense in Œ, the set of quadruples representing æ.[76] Now consider the countable sum of the K, a region u. It is a part of œ but there is no part of œ disjoint from u, for if there were it would be represented by an open non-empty set disjoint from U the representative of u, contradicting density. So if Weak Supplementation holds, u = æ, but in any case by Hypervolume Supplementation (Prem-

---

[76] There are countably many quadruples of rational numbers. Those in Œ may be arranged as a sequence $\langle q_1, q_2, \ldots \rangle$. K consists of globules of hypervolume $(½)^{m+1}$ represented by an open set containing q_m, for all positive integers m.

ise Eight), vol(u) = vol(æ) ≥ 1. Hence the sum of the K has hypervolume greater than the sum of the hypervolumes of the K, contradicting Countable Subadditivity, which follows from Premises Nine and Ten. (See the previous Chapter.)

There are difficulties even if we reject Countable Subadditivity. For we may consider again the variant on Banach Tarski used to refute Tarski Continuum. There are disjoint regular open sets U1 to U5 of total Lebesgue measure no greater than 4π/3 (the hypervolume of a unit ball persisting for unit time), such that V, the smallest regular open set containing their U1 to U5, has measure equal to the sum of two balls of unit radius persisting for unit time, which we may take to be a region v of hypervolume 8π/3. So V is the interior of the union of the closures of U1 to U5. To preserve the finite additivity of hypervolume we would have to suppose that U1 to U5 are all subsets of an open set W representing a region w of hypervolume less than v. By Hypervolume Supplementation, v has a part x disjoint from w. So x is represented by a non-empty open set X disjoint from U1 to U5 but nonetheless a subset of V, contradicting the supposition that V is the interior of the union of the closures of U1 to U5.

Nor can we rescue Open Gunk theories by retreating from Premise Eight to the weaker, but more firmly intuitive, Premise Eight*. For if all regions are represented by open sets the two principles coincide. For two open sets intersecting in a non-empty set must intersect in a non-empty open set and hence one of positive measure. The retreat from Premise Eight to Premise Eight* was only viable in the case of Pseudo-set Granules.

6. Should we reject countable summation?

Thus far I have supposed that even if Premise Two (Universal Summation) fails any countable infinity of regions have a sum, but some might reject even that intuition. This leads to a hypothesis according to which if we adjoin the fictitious empty region ø then the lattice of regions forms a Boolean algebra but we do not require countable additivity of pairwise disjoint regions. (This is similar to Achille Varzi's definition of a *closed* mereology in Varzi 2000.) In that case we may restrict attention to comparatively few regions, by taking the algebra associated with the mereology to be the Boolean algebra generated by the *globules*, as in Sparse Continuum in which the globules are represented by non-empty convex open sets.

Such hypotheses are open to the criticism I made of Borel Continuum, namely that regions are not constituted by the summation of more fundamental ones (the globules). In addition, problems arise because, although we do not assume countable summation, there are still cases in which one region is in fact the sum of countably many others and in some such cases Countable Subadditivity fails. Consider a region u represented by a non-empty convex open set U of coordinate quadruples. For convenience take u to have hypervolume 1 unit. Then arrange the quadruples of rationals in U as a sequence χ_1, χ_2, χ_3, ... There for each j is a region v_j whose representing set is $U \cap W_j$ where W_j is the ball center χ_j radius $(½)^{j+1}$. Then u is in fact the sum of the v_j, but the sum of the hypervolumes of the v_j is less than 1 unit. What this example shows is that the attempt to restrict the regions to ones that can be described easily without assuming infinite summation results in infinite sums that are counter-intuitively large. Hence I reject Sparse Continuum and its variants, unless a strong case can be made for them on other grounds.

7. Countably many points or granules

A fairly natural suggestion is that the region of finite diameter œ is the sum of a countable infinity of points. I call this the Aleph Null Hypothesis. In that case we represent regions as non-empty subsets of some countable set of quadruples dense in Œ, the set representing œ. If all points have zero hypervolume, it then follows by Countable Subadditivity that every region of positive hypervolume has zero hypervolume, which is a contradiction. But if some region has positive hypervolume then we have a violation of Premise Three (the Diameter/Hypervolume nexus). In the case of Point Discretion we were able to undermine that intuition by replacing the, ‘How much?’ question by, ‘How many?’ No such undermining is available for the case of the Aleph Null Hypothesis. For all hyperballs would be assigned the same quantity, $\aleph_0$

Even if we ignored this problem we would still have the difficulty that there would have to be an infinite range of hypervolumes assigned to points. Otherwise œ contains only finitely many regions of non-zero hypervolume and so we have a discrete theory with some quite unnecessary parts of zero hypervolume added on. On the usual grounds of theoretical simplicity we should reject a theory that posits an infinite range of hypervolumes to its fundamental regions without further explanation.

The Aleph Null Hypothesis might be rescued by denying countable subadditivity. In that case every region that is the sum of finitely many points has zero hypervolume but the sum of countably many of them might have a non-zero hypervolume sum. The problem with this is that there is then no satisfactory way in which the shape and size of the region fixes its hypervolume. The countably many points should have representatives that form a dense set, for instance, those with rational coordinates. Hence we would expect the hypervolume of a region to equal the Lebesgue measure of the topological closure of the representing set. But that permits in a cheap version of the Banach Tarski Paradox. We can find disjoint regions u and v both parts of œ such that u, v and œ have the same finite non-zero hypervolume. And as usual we can ensure connectedness by allowing, in place of their being disjoint, the regions u and v to overlap in a region of very small hypervolume.

To meet this, the hypothesis would have to be modified further so that not only countable subadditivity is restricted but so is countable summation. The most straightforward response to this is to note that it is worse to abandon two intuitions than just one, unless there is some way of undermining them. So a hypothesis in which both countable subadditivity and countable summation are rejected is less likely than the four exemplars.

I also note the variant on Extended Simples, in which the granules differ in size, and a region of finite hypervolume may therefore contain countably many of them. As far as I can se this variant has no a priori advantage over Extended Simples itself that would compensate for the extra complexity.

8. Infinite dimensional variants

We may consider infinite dimensional aether, with regions represented by sets of infinite sequences of real numbers. Then, the distance between $<x_1, x_2, \ldots>$ and $<y_1, y_2, \ldots>$ is $\sqrt{((x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots)}$. If we concentrate on a region œ of finite extent, represented by a set of infinite sequences Œ, then by taking the sequence of zeros $<0, 0, \ldots>$ to be a member of Œ, it follows that for every sequence $<x_1, x_2, \ldots>$ in Œ the sum $x_1^2 + x_2^2 + \ldots$ is finite. In particular if we took all the x_n to be integers then all but finitely many of them must be zero. In any case the sequence $<x_1, x_2, \ldots>$ converges to 0 as n increases. To be sure, I have been considering diameters but the same consideration holds if there is a 'metric' on the aether that is approximately Minkowskian. In that case, the 'distance' between $<s, x_1, x_2, \ldots>$ and $<t, y_1, y_2, \ldots> = \sqrt{(s - t)^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - \ldots}$.

Infinite dimensional aether is, I claim, improbable. For the measure of extent (or its relativistic analog) coheres poorly with the measure of quantity. Consider a hypercube one corner of which is represented by a sequence of 0s, with the opposite corner represented by the sequence, $<x_1, x_2, \ldots>$. In order that the distance between them to be finite we require that the sequence, $<x_1, x_2, \ldots>$ converges to 0 as n increases. But the hypervolume is the infinite product of all the real numbers $x_1, x_2, \ldots$ and for that to be non-zero we require that the sequence converges to 1 as n increases.

If we had some empirical evidence for infinite dimensionality then it might be worth some ad hoc way of resolving the problem that the sequence, $<x_1, x_2, \ldots>$ converges to both 0 and 1 as n increases. Now we know what the empirical evidence would be like: a succession of physical theories each an improvement on the previous one that require at smaller and smaller scales the resort to more and more dimensions. To be sure even then we might debate whether to draw the conclusion that there are infinitely many dimensions. In any case, all we have so far are the rather speculative String and other Supergravity theories, taking the first step. 'One swallow does not make a summer.'

9. Aiming for completion

Ideally I would have provided a systematic way of surveying all the hypotheses that are not wildly implausible, but I do not know how to start. I do know that I have not listed all such hypotheses that I can think of. Notably absent are those in which regions are represented by the non-empty topologically closed sets, or by countable unions of such sets. The best I can do is to go through the four basic kinds of hypothesis, point-based discrete, point-free discrete, point-based continuous and point-free continuous and consider why I can reasonably hope that the best hypothesis is either the exemplar or some variant that I have considered. I should preface this exercise by reminding readers of my stipulation that fine structure is being ignored, by noting that just possibly we might have some good reason to give a different hypothesis for Space and for Time, and by asserting that, other things being equal, hybrid theories are less plausible than their parents.

Point-based discrete aether

Qualitative differences between points are conceivable, because of fine structure, but positing them is extravagant. So I assume all points are exactly similar. If there are finitely many in a given region of finite diameter the resulting hypothesis is Point Discretion. Otherwise point-based discrete aether hypotheses require an infinite number of points in a unit hypervolume of continuous Space-time. Unless we represent the points by a set of coordinate quadruples œ dense in Œ (i.e. Œ is the closure of the set of coordinates representing œ.) there is a further problem that œ has holes, represented by the quadruples in Œ but not in the closure of œ. Unless required by physics it is extravagant to countenance such holes. This leaves as a not too implausible the hypothesis in which points are represented by a dense set of quadruples, say using rational coordinates only. Strictly speaking this is a continuous hypothesis. The problem with it is in assigning a measure of quantity. The natural suggestion is that the hypervolume of a region should in this case be the Lebesgue measure of the closure of the representing set of coordinate quadruples. But, as already noted, that leads to a Banach Tarski type absurdity. Given a countable sum of points in a region œ of hypervolume 1 unit, represented by coordinate quadruples whose closure is Œ, œ is the sum of two disjoint parts u and v represented by sets U and V where U, V and the closure of Œ all have the same Lebesgue measure. So hvol(u), hvol(v) and hvol(œ) are approximately the same, which is absurd. Hence, the only plausible point-based discrete hypothesis is Point Discretion itself.

Point-free discrete aether
Any macroscopic region of finite hypervolume is approximately divisible into two. By that I mean that u = v + w where hvol(v) ≈ hvol(w) and either v and w are disjoint or hvol(v∧w) = 0, or, if we are being scrupulous, hvol(v∧w) is very small compared to hvol(v). In the point-free discrete case, such division eventually fails. Because we are ignoring hybrid hypotheses we shall reach such failure regardless of the hyperplane of approximate division and so we may assume that the aether is the sum of regions that which cannot be further divided into two, in the above fashion. Call these regions *granuloids.* Theoretical simplicity favours uniformity, so we may take the granuloids to be the same shape and size. They may overlap, in which case we can note in passing the suggestion that the hypervolumes of the overlaps is positive but much smaller than the overlapping granuloids Rejecting that extravagance, the granuloids are granules and we are left with the range of hypotheses considered previously.

Point-Based Continuous Hypotheses
There are very many hypotheses we could consider based on the choice of representing sets. Here there are two choices: that of the representations of points; and that of which sets of points have sums, which will then be represented by the corresponding sets. Regarding the first choice, I have already considered, and rejected, hypotheses according to which the number of aether points in a region of finite diameter is countable. This leaves very many hypotheses with uncountably many points. But there are two considerations in favour of representing a local region œ by a set Œ of coordinate quadruples that is, say, a convex set – unless, that is, we posit higher dimensions for the sake of the physics, in which case n-tuples replace quadruples for some n > 4. One is an appeal to topological simplicity: excising some quadruples from the representing set

Œ results in more complicated structure. The other is a characterisation problem. It will be hard enough to characterise the structure of the aether in some way that does not rely upon the representation, but I cannot even see how to begin that task in the point-based case unless Œ is a suitable set of coordinate quadruples, such as a convex one.

Granted that Œ is a suitable set of coordinate quadruples, Universal Summation then implies the Orthodoxy, which is only tenable if we reject the Axiom of Choice, as discussed previously. Curmudgeons have a principled way of restricting Universal Summation, but the problems with the Orthodoxy are not solved by the curmudgeon variant that requires regions to be connected. Another principled restriction is to exclude precisely the summations that result in regions to which no hypervolume can be assigned. This would lead to Lebesgue Continuum, but it should be noted that if we combine Lebesgue Continuum with the initially plausible thesis that all regions depend ontologically on points, then the principle of Dependent Quantity is violated, just as it was by the Orthodoxy combined with Solovay's Axiom .

I conclude that continuous point-based hypotheses are not point-*based* in the strict sense, because not all regions depend ontologically on points. That raises the question of just what the fundamental regions are, if they are not just the points. The most plausible answer is that they are the globules instead of, or as well as, the points. But in that case we no longer require uncountable summation and so may rely upon the principled restriction of Universal Summation to the countable case. In that case, the chief difference between hypotheses concerns the ways in which more complicated regions are grounded in the less complicated, and ultimately in globules. Countable summation is highly plausible. But if the sum of countably many regions is represented by something larger in Lebesgue measure than the union, then countable subadditivity will be violated. Hence we may assume that the countable sum of regions represented by sets U_j of coordinate quadruples is represented either by their union or a set differing by zero measure from their union. If the only way in which less complicated regions ground more complicated ones is (finite or countably infinite) summation then every region would be represented by one differing from an open one by a set of measure zero. This will be subject to the same objection as raised to Open Gunk .

We require, then, that there be an additional way of grounding the more in the less complicated. The only two that come to mind are (1) countable meets and (2) taking of complements. (1) would give us Gδ Continuum, with regions represented by the non-empty countable intersection of open sets. In Chapter One, I argued that Gδ Continuum was inferior to Borel Continuum, because there was no motivation for the thesis that the meet of regions depends ontologically on the regions in question, but, relying on realism about absences, the dependence of the differences x - y on x and y can be motivated.

Point-free Continuous Aether.

If the aether is point-free we may construct the point(-locations) of Space-time as ultrafilters with respect to the topological relation of interior parthood. Hence we may compare point-free theories by representing regions as sets of points in Space-time. If we ignore fine structure, no two regions are represented by the same points, so the point-free theories may be surveyed by first asking, 'What sort of non-empty sets do the representing?' and then 'When are two representing sets equivalent in that they represent the same regions?' Some answers to the first question might be: (1) regular open sets, (2) maximal open sets (X is maximal open if it is not a proper subset of an open set of the same Lebesgue measure), (3) all open sets, (4) countable intersections of open sets, (5) Borel sets, (6) Lebesgue measurable sets. One obvious answer to the second question is (a) that equivalence is identity. Other answers are that equivalent sets differ by: (b) at most finitely many points; (c) at most a countable infinity of points; (d) a set of at most 3 dimensions (with variants in which the number of dimensions is 1 or 2 not 3); (e) a set whose closure has empty interior; (f) a *meagre* set (that is a countable union of sets whose closures have empty interior; or (g) a set of zero Lebesgue measure. Not all the combinations of answers give different results. In particular if we consider Arntzenius Continuum and its variants based on answer (g), then (2) and (3) coincide as do (4), (5) and (6).

This still leaves a great variety, but Arntzenius Continuum leads all the rest. First note that representations by open sets are confounded in one way or another by the occurrence of open dense sets in Œ with Lebesgue measure less than that of Œ itself. Such a set is a countable union of open balls, which must represent globules. So either these balls fail to satisfy the subadditivity of hypervolume or their sum is a part of œ of less hypervolume than œ, which fails to satisfy Premise Eight*.

The next consideration is that if we abandon points we should retain the feature that the measure of quantity is *faithful* in the sense that if x is a proper part of y and hvol(y) is not infinite then hvol(x) < hvol(y). Such fidelity is directly intuitive only if we allow for infinitesimal differences, but a straightforward appeal to conceptual economy should make us prefer hypotheses without these infinitesimal differences. In that case fidelity follows from the intuitive idea that the part has less stuff making it up than the whole. Only answer (g) ensures that such fidelity. But given that we have excluded representations by open sets, answer (g) ensures that all the suggestions coincide with Arntzenius Continuum.

Provisional Conclusions

I have listed four hypotheses, the exemplars, together with variants, that are worth further investigation. They are: my a priori preferred hypothesis, Point Discretion, then Extended Simples and Arntzenius Continuum, followed by Borel Continuum. Pseudo-set Granules will, however, turn out to have some advantages compared to Extended Simples, of which it is a variant.

We might perhaps distinguish the spatial from the temporal aspects of the aether, giving one theory for the spatial and one for the temporal. Thus the aether might conceivably be temporally discrete but spatially continuous. I shall consider and reject this proposal in Chapter Six.

What other hypotheses have I ignored? Plenty but as far as I know none of them are a priori as likely as the exemplars. Now, in the context of further theory some of my exemplars might be rejected and some that had been dismissed might be reconsidered, especially those rejected only on grounds of simplicity. If further theory exhibits additional problems with all the hypotheses considered, then a more thorough survey of hypotheses would be required.

I have reached two provisional conclusions concerning the mereology, namely that all the theories worth considering are Boolean once the fictitious empty region is adjoined, but that we have no strong reason to assume classical mereology. Borel Continuum for instance is not consistent with Universal Summation.

I have also made two provisional comparisons.

(1) Point Discretion is the provisional winner. (It will, however, be rejected later.)

(2) Of the other three hypotheses it is hard to rank Extended Simples against Arntzenius Continuum, which, however, is superior to Borel Continuum.

4. Hypervolume and Topology

The previous chapters were incomplete in two ways. The first, which is the chief topic for the rest of this book, is that the ease or difficulty of adjoining extra structure might affect the ranking of the hypotheses. I begin that investigation in this chapter by considering topological structure. Because it is easier to characterise, this provides a foil for the problem of characterising differentiable structure, discussed in the next chapter.

I also discuss another respect in which the first three chapters were incomplete: although I was considering 4-dimensional aether the intuitions relied upon were based on those about objects extended in 3 dimensions. Yet Special and General Relativity are widely, and I think correctly, taken to imply that there is no such property as *the* diameter of something extended in Space and Time, for diameters are frame-relative.[77] Because I have relied upon intuitions about diameter I need to show that those intuitions survive transition to frame-relativity.

First, though, I reply to the objection that hypervolume and diameter are not fundamental quantities. That is, they depend ontologically on something else. I shall concentrate on hypervolume. Similar considerations apply to diameter.

1. Hypervolume and Ockham's Razor

One of the advantages of Point Discretion was that we replace hypervolume by another measure of quantity, the number of points, which is unproblematic. On the other hypotheses about the structure of the aether, it is tempting to posit a relation between two regions u and v and a number x, namely that the u and v have *hypervolumes in the proportion whose value is* x. This requires relations between things of different categories, namely regions and numbers, and such cross category relations should, I

77 Although we can define a frame-independent measure of extent it is not a diameter, because it is not an analog of the familiar, but frame-relative, 3 dimensional diameter. For it fails to satisfy the principle that if two regions overlap, then the diameter of their sum is no greater than the sum of their diameters.

say, be avoided if we can.[78] Moreover, I follow Newton in holding that numbers just are proportions, in which case the phrase 'whose value is' is redundant.[79] So I say that hypervolumes are properties of regions and that these properties themselves stand in various relations of *proportion,* which we then identify with non-negative numbers.[80] Considering volumes in three dimensions in place of hypervolume for simplicity of exposition, we may say that the volume 1 cubic metre and the volume 1 cubic foot stand in a proportion relation which is identical to some number (approximately 35.3).

Ockham's Razor is sharp when it comes to fundamental properties and relations, and I think we should posit a fairly small number of these. On my preferred account of hypervolumes a trans-categorical triadic relation is replaced by an infinity of distinct properties. So by Ockham's Razor these should not be fundamental. Yet, the objection would go, I have not explained what hypervolume depends on. Moreover, because hypervolume is not fundamental our intuitions about it are vulnerable to defeat by explication. That is, we should investigate to what extent our intuitions depend on ones concerning the more fundamental structure, and if they do, check that our intuitions survive this explication.

Except in the case of Point Discretion, hypervolume depends on the single dyadic *less-than-or-equal quantity* relation, $x \leq_q y$. There is a derived *equal quantity* relation, $x =_q y$, which holds if $x \leq_q y$ and $y \leq_q x$, and a derived less quantity relation, $x <_q y$, which holds if $x \leq_q y$ but not $x \leq_q y$. The theory becomes more complicated if, contrary to our intuitions, there are some regions to which no hypervolume is assigned. In

---

78 One reason is that if the Xs are related to the Ys, then in addition to the categories of relations between Xs and relations between Ys, there must be an additional category of relations between Xs and Ys, which is uneconomic.

79 'By number we understand not so much a multitude of unities, as the abstracted ratio of any quantity to another quantity of the same kind, which we take for unity' (Newton, 1728:2). See also (Forrest and Armstrong 1987).

80 To be sure, we may sometimes say that the hypervolume of u and the hypervolume of v stand in the ratio ∞ to 1, but I paraphrase this as saying that the hypervolume of v and the hypervolume of u stand in the ratio 0 to 1.

that case *having hypervolume* would have to be taken as a primitive property of regions.

For regions that have hypervolume – hopefully all regions – the $\leq_q$ relation is transitive reflexive and *linear*. Linearity is here the principle that for any regions x and y that have hypervolume either $x \leq_q y$ or $y \leq_q x$.

First let is consider the discrete case. If Point Discretion is correct, then, as already mentioned, we do not need the $\leq_q$ relation. For the quantity of a region is just the number of points. In the case of Extended Simples and its variants, all we need is the relation of same-quantity which we may either take to be a fundamental equivalence relation or characterise it thus: $x =_q y$ if $x \leq_q y$ and $y \leq_q x$. For then the plausible hypothesis that all extended simples have the same hypervolume enables us to use the number of simple parts in a region as a measure of its hypervolume, showing how hypervolume depends on the $\leq_q$ relation.[81]

In the continuous case we may use the standard methods of measurement theory due to Otto Hölder to derive hypervolume from the $\leq_q$ relation (See Suppes and Zinnes 1963). We start with the region œ that we may take to have unit hypervolume.[82] We may then characterise a region of *positive hypervolume* as a region u such that for some finite

---

81 If we were to posit extended simples that are not all of equal hypervolume, then we could still take hypervolume as dependent on the $\leq_q$ relation. In that case I would assume that for any two simple regions u and v there are integers m and n and some disjoint extended simples, the Ws, all of the equal hypervolume, such that u is of equal hypervolume to the sum of m of W and v of equal hypervolume to the sum of n of them. There are, however perverse examples in which the purported hypervolume could not be derived from the $\leq_q$ relation. For instance, suppose there are only finitely many regions in the whole universe, all of which are sums of extended simples, some of which have hypervolume 1 unit and some hypervolume $\sqrt{2}$ units. In such cases we could not ground the hypervolumes in the $\leq_q$ relation. But we have no reason to posit such perverse structures.

82 To avoid ending up with a system of infinite hypervolumes or infinitesimal hypervolumes we need an anthropocentric scale, taking œ to be, say, the region occupied by the standard metre during the Nineteenth Century.

number of regions x_1 to x_n, $x_j \leq_q u$, $x_j <_q$ œ, and $x_1 + \ldots + x_n =$ œ. Restricting attention to regions that have positive hypervolume, we may propose the following principles:

(1)If x is a part of y then $x \leq_q y$.

(2)If $x \leq_q y$ and $y \leq_q z$ then $x \leq_q z$.

(3)Either $x \leq_q y$ or $y \leq_q x$

(4)Given any x and y there is some z such that $x =_q z$ and y and z are disjoint.

(5)If $x \leq_q y$ and z is disjoint from x and y then $(x + z) \leq_q (y + z)$.

(6)If $x <_q y$ then there is some z disjoint from x such that $(x + z) =_q y$.

(7)If $x \leq_q y$ there is an integer m and pairwise disjoint regions u_1 to u_m such that $x =_q u_j$ for all j, and $y \leq_q (u_1 + \ldots + u_m)$.

From these principles it follows that there is a measure of quantity, hypervolume, taking values from 0 to ∞, such that (a) $x \leq_q y$ if and only if hvol (x) ≤ hvol(y), and (b) if x and y are disjoint, hvol(x + y) = hvol(x) + hvol(y).[83]

Premise Ten may be modified to obtain *Premise Ten**:

> Consider some totally ordered regions the Xs. Suppose there is some region w such that, for any x that is X, $x \leq_q w$. Then, if the Xs have a v, $v \leq_q w$.

Premise Ten* together with (1) to (7) show that hypervolume is countably subadditive, but in any case all we required in Chapter Two was that the sum of regions of zero hypervolume is of zero hypervolume, where a region x is of zero volume if for all y, $x + y \leq_q y$, for which Premise Ten* directly suffices.

Of the eleven Premises of Chapter Two those that are required in addition to Countable Subadditivity are already of a purely qualitative

---

83 Consider the equivalence classes *u, v* etc under $=_q$. Define R by *u*R*v* iff for some $x \in u$, $y \in v$, $x \leq_q y$. Define *u*S*v* as the equivalence class of any x + y such that $x \in u$, $y \in v$ and x and y are disjoint. Then it is easy to see that the system consisting of the equivalence classes, and the relations R and S satisfy the axioms of an *extensive system* (Suppes and Zinnes 1963: 42). Therefore, their representation theorem (Suppes and Zinnes 1963: 43) holds.

nature and can be expressed easily enough in terms of the *less-than-or-equal quantity* relation. Thus we may paraphrase 'hvol(u) < 1/n' by 'There are n regions of equal quantity x_1 to x_n, such that $u \leq_q x_j$ and such that $x_1 + \ldots + x_n =$ œ.'

2. The threat of Relativity

Rejecting as I do the neo-Lorentzian theories, championed by Craig (2001), I assume there is no necessarily privileged frame of reference. But diameter is frame-relative, so there is no such thing as *the* diameter for us to have intuitions about. Intuitions should, I submit, be stated in terms of genuine properties and relations, even if these are not fundamental, so this is a threat to some of premises used in Chapter Two.

It should be noted that a similar threat does not arise for hypervolume, which is independent of the choice of frame of reference, even in General Relativity, where it depends on the gravitational field but not the choice of coordinates.

I begin by noting that the Premises of Chapter Two concern only the following properties defined in terms of diameters, and their negations: a region *having zero diameter*; a region *having finite but non-zero diameter* and; a sequence of regions *having diameters tending to zero in the limit.* This last occurs whenever we consider arbitrarily small diameters. These properties, and the premises stated in terms of them, are frame-invariant. Furthermore we can explicate them using Alexandrov intervals, namely non-empty meets of future and past light cones. A region is of *finite diameter* if it is part of some Alexandrov interval. A region u is of *zero diameter* if: for any region z, there is an Alexandrov interval y such that $u \leq y$ and $y \leq_q z$. Assuming œ has finite volume, the sequence of regions u_j, j = 1, 2, . . . *has diameters tending to zero in the limit* if:

> For every positive integer m, there is some positive integer n, and an Alexandrov interval v_n such that $u_j \leq v_n$ if $j \geq n$, and there are m pairwise disjoint parts of œ, w_k, k = 1, 2, . . . , m for which $v_n \leq_q w_k$, k = 1. . . m.

This enables us to state the eleven Premises of Chapter Two without requiring diameters.

3. Mereotopology

Even if the topological structure is not fundamental, it is part of the necessary structure of the aether and so it is of interest that it can be described simply. In Mathematics, topology only applies to a point-based theory of the aether, but for our purpose, we need to extend the idea of topology to cover point-free theories. What we need is provided by the flourishing discipline of *mereotopology.*[84]

The fundamental mereotopological concept could be taken as *touching* or *adjacency*. It sounds a little odd, however, to say that every region touches itself, or even that any two overlapping regions touch. So instead I follow the standard usage that the dyadic relation of *connection* holds between adjacent or overlapping regions. There is then a harmless ambiguity between this relation and topological property of being connected that holds of a region unless it is the sum of two parts that are not related by the connection relation.

Sometimes is more convenient to consider *separation,* the negation of connection, namely neither overlapping nor touching. Separation is definable in terms of diameters as follows. The regions u and v are *separated* unless there is a sequence of regions x_n overlapping both u and v and the diameters of the x_n tend to zero.

Interior parthood, «, may be defined in terms of separation thus:

x « y if x is separated from every region disjoint from y.

Assuming Boolean mereology we expect the following to hold:

1. If u « v then u < v.
2. If w ≤ x « y ≤ z then w « z.
3. If u « x and v « y then u∨v « x∨y.
4. If u « x and v « y then u∧v « x∧y.
5. If u« x then ¬x « ¬u.

[84] For a comprehensive introduction to mereotopology, see (Casati and Varzi 1999).

Clearly if 5 holds we need only one of 3 or 4.[85] In addition, if we have a hypervolume measure we expect that:

6. If x « y and x is of finite hypervolume then hvol(x) < hvol(y).

Now we could define separation in terms of interior parthood thus: x is separated from y if x is an interior part of the complement of y. But that makes it a mystery why separation is symmetric, something that should be intuitively obvious. Hence I take separation or connection as more fundamental than interior parthood, even though the latter is more convenient theoretically.[86]

The mereotopology may therefore be characterised in terms of any of the easily inter-definable relations of connection, separation and interior parthood, and I hold that either connection or separation as fundamental. But there are alternative candidates for the fundamental topological item, and I need to explain why I reject them. For instance, some or all regions have the special property of *being open*, where an open region is one that is the sum of its interior parts. If we adjoin a fictitious empty region, ø, that is treated as open too. In order to mimic point-set topology we would suppose the resulting lattice of open regions is a complete Heyting one, as in the theory of *locales* (Johnstone 1982).

The argument for not taking the property *being open* as fundamental is that I take it that, other things being equal, our conceptual analysis is a guide to ontological dependence. I do not think that the concept of being open is conceptually primitive. Instead it is understood in terms of some other concept such as interior parthood, whereas separation is itself highly intuitive. This argument may be strengthened by noting that we have no way of visualising or imagining in a tactile fashion the difference between open and other regions, and hence nothing for our intuitions to get a grip on. By contrast, we can imagine separation. To be sure, in a point-based theory we can imagine the difference between an

---

85 Consider the hypothesis that regions are represented by the maximal open sets of quadruples. This would be an example of a non-Boolean system in which 1, 2, 3, 4 and 6 (below) held.

86 Because points of Space-time may be characterised as ultrafilters with respect to interior parthood (Forrest 2010).

open region and a closed one by considering two objects moving closer but incapable of overlapping. If they always occupy open regions they can touch but if they always occupy closed regions they must remain separated. But this way of getting a partial intuitive grasp of open regions only serves to support the claim that separation is a more fundamental relation.

But is topology fundamental? In the case of Minkowski Space-time and other plausible symmetric structures for the aether, such as de Sitter Space-time, we may derive the topological structure from the light-cone structure, and hence from the ordering of regions with respect to *absolute priority.*[87] (Region u is absolutely prior to region v if with respect to every frame of reference every part of u is earlier than every part of v.) On the standard, covariance interpretation of General Relativity, Space-time is assumed to have the structure of a differentiable manifold, which, to be sure, entails that it has a topological structure. The differentiable manifold structure turns out to be problematic and it is to be hoped that it could be replaced by the topological structure, as it can be on the hypothesis of Granulated Aether (See the next chapter). It may turn out, then, that topological structure is fundamental after all.

There is another way in which topology might turn out not to be fundamental. That is if the topological structure can be replaced by mereological structure, as on the Pseudo-set Granules hypothesis discussed below, or on Sparse Continuum or any other continuous Boolean mereology in which all the regions are open. For then we construct points, and the open sets are (in addition to Ø) the unions of sets representing open regions.

Topology, whether fundamental or not, is important for an additional reason. It provides some additional assurance that we have not overlooked any plausible hypothesis for the necessary structure of the aether. For given a suitable topology we may construct point-locations and hence adjoin points if necessary, obtained a point-based Space-time in which the aether is located. Space-time is then a point-set topology, which is a well-researched mathematical theory.

---

87 See (Moschella 2005) for an introduction to de Sitter and anti de-Sitter Space-time.

There are two ways of introducing these Space-time points, here considered as fictitious aether-points. If the open regions plus ø form a complete Heyting lattice then we may consider maximal proper open regions as in one to one correspondence with the points, because if there were points then the set of all points except point p would be a maximal proper open region. The further principle that if u and v overlap the same open regions then u = v, then shows that every region is the sum of real or fictitious, points. If we started with a point-based aether with a topological structure, then this procedure gives us the topology we started off with. (See Johnstone 1982.)

The second method relies upon a mereotopology characterised using connection, separation or interior parthood. We characterise the points of Space-time so as to represent regions by sets of points. Then connected regions are to be represented by sets whose closures overlap[88]. The method, pioneered by Peter Roeper (1997), is to take the points to be *ultrafilters* with respect to interior parthood. This is an explication of Whitehead's Russian doll construction of points as a sequence of ever smaller regions.[89]

A noteworthy feature of the ultrafilter construction is that in the case of point-based aether with points corresponding to all quadruples of real numbers the ultrafilter construction would adjoin a fictitious point at infinity. Likewise if we started with a gunk theory in which regions were represented by suitable sets of quadruples the ultrafilter construction would adjoin a point at infinity. In both cases the point at infinity corresponds to the ultrafilter that has as its members all the complements of regions of finite diameter.

---

88 This is a modification of Marshall Stone's method of representing Boolean algebras topologically (Stone, 1936). For more details on constructing points as a ultrafilters with respect to interior parthood see (Forrest 2010). With care we may avoid assuming the Axiom of Choice.

89 A *filter* is a set of regions X such that: (1) If $x \in X$ and $x \ll y$, then $y \in X$; and (2) If $x \in X$ and $y \in X$ then for some w, $w \in X$, $w \ll x$ and $w \ll y$. (Note: we do not treat ø as a region.) A *proper* filter is one that does not have every region as a member. An *ultrafilter* is a filter that is maximal among proper filters.

Set-theoretic constructions are not the right sort of thing to be considered points. So if aether is point free and Space-time is dependent on the aether and made up of points, then I propose that the points of Space-time are suitable *properties* of regions. I anticipate the retort that properties are no more suited to be points than are sets. Strictly speaking that is correct, and I should replace points by point locations. To a fictitious point p there corresponds an ultrafilter of regions X. Therefore the property of *being at p* may be identified with the conjunction of the relational properties: *overlapping u*, for all regions u in X. The property *being located at the point at infinity* is then just the property of *being of infinite extent*.

4. Topology in the discrete case

If Point Discretion holds, then connection collapses to overlapping, separation collapses to being disjoint, interior parthood collapses to parthood, and all regions are open. So the topology is trivial. Moreover it is compatible with any metric. This is, however, something more to be said. For there might well be a minimum non-zero value that diameters take, and in the simplest case all diameters will be a whole number times that minimum. Then we may describe points that are the minimum non-zero distance apart as *neighbours*, and call the sum of a point p and its neighbours the *neighbourhood* of p. Many concepts such as an analog of the property of connectedness can be defined in terms of neighbourhoods. That would be great if the regions had frame-invariant diameters. But they do not. Instead we may posit a light cone structure analogous to that for Minkowski Space-time. In that case there is a relation of *absolute priority* between any points u and v, ($u \prec v$), where point u is absolutely prior to point v just in case $u \neq v$ and u is part of every past light cone containing v. Then x is the *immediate predecessor* of z ($x \prec^* z$), if $x \prec z$, but there is no y such that $x \prec y \prec z$. We may then say that points x and z are neighbours if $x \neq z$ and either $x \prec^* z$ or $z \prec^* x$, or if for some points u and v, such that $u \prec^* x \prec^* v$ and $u \prec^* z \prec^* v$. We should, however, be suspicious of this topology-analog. For suppose we represent points as quadruples of integers. Then the points represented by <0, 0, 0, 0> and by <m, n, p, q> can be immediate neighbours even though m, n, p and q are all large positive numbers. This peculiar feature forms the basis of the objection to Point Discretion presented in Chapter Seven.

If the aether is granulated, the characterisation of topology in terms of diameter fails. The reason is that Arbitrary Fine Division does not hold and hence the distance between extended simple regions may not be characterised as the greatest lower bound of the diameters of overlapping regions. Fortunately in these theories the idea of two regions touching is still highly intuitive. Formally it is like the topology-analog for Point Discretion with *adjacency* replacing *being neighbours*. Moreover, the topology of extended simples is frame-invariant. For change of frame will change the shape of the sets representing regions, but not the adjacency. Likewise, the shape of the granule itself is frame-invariant, being characterised in terms of adjacency.

The Pseudo-set Granules hypothesis, although more complicated than Extended Simples, does have one great advantage, namely that we can characterise the topology in mereological terms, with connection defined as overlapping. The reason is that two granules that would be non-overlapping but adjacent on the Extended Simples hypothesis correspond, on the Pseudo-set Granules hypothesis, to atoms that have a common part, namely a shared facet, face, edge or vertex. Because the mereology is non-standard we obtain a non-trivial topology by identifying overlap with connection. In particular, the sum of two overlapping granules is a connected region (because of the failure of Weak Supplementation).

5. The characterisation problem

In the previous chapters I presented a variety of hypotheses about the aether by means of their coordinate representations. I require, however, a frame-independent characterisation, and one that does not involve set-theoretic constructions. That is straightforward in the case of Point Discretion – ignoring fine structure we may characterise that hypothesis by saying that a region of finite diameter is the sum of finitely many parts, each of which has zero (or infinitesimal) diameter.

The case of granulated aether is almost as straightforward. A granule may be defined as a region that cannot be decomposed as the sum of two regions of lesser but positive quantity. So we begin by hypothesising that any region of finite diameter is part of a finite sum of granules. In the case of Extended Simples every granule is simple and any region of finite diameter is the sum of disjoint granules. All that remains is to note their shape based on the primitive relation of adjacency. Variants such as Pseudo-set Granules are more complicated to describe, but the description provided in Chapter One meets the requirements of a characterisation independent of set-theoretic construction or coordinate representation.

In the case of a continuous point-based hypothesis it suffices to characterise the topological space these points form, together with some way of distinguishing regions of positive measure from those of zero measure. For then the description of the hypotheses in these terms proceeds as in the previous chapters.

Although topological spaces occur in bewildering variety those that current physics might take to be continuous Space-time are *manifolds* of 4 (or some other specified finite number N) dimensions.[90] Manifolds are topologically connected spaces that are locally represented by coordinate quadruples (or more generally N-tuples) of real numbers in the right way. We may think of the aether as the sum of overlapping globules, the J, each represented by coordinate quadruples. How the globules thus represented fit together is considered in the sub-discipline of algebraic topology, where the, now proven Poincare´ Conjecture, solves the characterisation problem, *provided* we can characterise the (mereo)topology of the globules.[91] Here I ignore the hypervolume, not for any principled reason, but merely because it neither helps nor hinders the solution of the characterisation problem.

The most straightforward case, in which the manifold lacks a boundary, is that in which for any j in J the topological space of point parts of j is homeomorphic (i.e. topologically equivalent) to the set of

---

90 If the whole physical universe is the sum of disconnected universes the aether could, of course, be a topological space that is the sum of disjoint manifolds. I grant that possibility, but for sake of exposition I concentrate on one universe.

91 Grigori Perelman proved the case of three dimensions. The cases of more than three, and in particular four, dimensions had already been proven.

quadruples of real numbers endowed with the usual topology. That is an extrinsic characterisation, setting up the problem of finding an intrinsic one, or at least one that does not resort to coordinates. The only way to do this that I know of is to rely on the existence of symmetries (automorphisms), in this case homeomorphisms from the topological space onto itself.[92] Let T be the topological space whose points are the point parts of j. Then I require that there be a set F of homeomorphisms of T onto T such that (0) to (5) hold:

0. If f and g are in F and if, for every point part x of j, f(x) = g(x), then f = g.

Because homeomorphisms are mappings and mappings are purely formal relations, the condition (0) is strictly redundant.

1. If f is in F, its inverse f^{-1} is also in F.
2. If f and g are in F, then the composite $f \circ g$ is in F.

If (1) and (2) hold, then F is a *group,* and we may derive

3. The identity automorphism, Id, which maps each point part of j to itself, is in F.

F is a *commutative* group if:

4. For any f and g in F, $f \circ g = g \circ f$.

Such a group of automorphisms F is said to be *transitive* if:

5. For any point parts u and v of j, there is an f in F, such that f(u) = v.

It follows from (0) to (5) that, for any point part of j, u, f(u) = u only if f = Id.

If there is a set of automorphisms, F, satisfying (0) to (5), then for any point part u of j there is a one to one correspondence between the members of F and the point parts of j, where f corresponds to f(u). We may therefore characterise the topological space T made up of the point-parts of j as homeomorphic to the quadruples of real numbers by providing conditions on commutative group F that suffice to show the group is

---

92 We should also require that they preserve hypervolume, but for simplicity of exposition I ignore this.

isomorphic to the group of quadruples of reals with addition as composition.[93]

Is this a satisfactory characterisation? In the Introduction I noted that reliance on a group of symmetries to characterise the structure of the aether might be unacceptable to some nominalists. There is a further difficulty in this case, however, because there are many such groups. Unless the aether has additional structure we cannot talk of *the* group of symmetries. Intuitively that detracts from the suitability of this way of characterising a manifold. If there were not much more serious problems characterising differentiable structure, discussed in the next chapter, this would be worth arguing about.

The characterisation of the structure of point-free aether is similar, but, I regret to say, even more complicated. The automorphisms are now mappings from the set of the parts of j onto itself that preserve both the mereology and the topology. The word 'part' replaces the phrase 'point part' in (0), and in the derived result (3). We might try replacing (5) by the following transitivity analog:

5*. For any parts u and v of j, there is some member of F, f, such that f(u) overlaps v.

In the point-based case (5*) implies (5) because points are regions. But if there are no points then (5*) is a weaker condition. For suppose we construct point locations in such a way that for each region u there is a set U of such point locations. (We exclude the 'point at infinity'.) Then F acts as a group of homeomorphisms from U onto itself. The conditions

---

93 By requiring the one to one correspondence between T and the group F to be a homeomorphism we may consider F to be a *topological* group. That is, it is itself a topological space and the operations of composition and taking inverses are continuous. Because F is a topological group the characterisation problem is then easily solved. The following conditions, for instance, are jointly sufficient for F to be isomorphic as a topological group to the N-tuples of real numbers with addition as the group operation.

1. F is a commutative topological group.
2. F is a locally compact, non-compact, connected space of topological dimension N, with a countable dense set.
3. Apart from the trivial subgroup {Id}, F has no subgroup with compact closure.

(0) to (4) and (5*) are not enough to ensure transitivity. For given a point location x, F(x) = {f(x): f is an F} is not necessarily the whole of U. Instead all we may assume is that F(x) is dense, that is cl(F(x)) = cl(U).

Because of the replacement of (5) by (5*), I shall require the aether to be the sum of globules, the Ks, that are interior parts of other globules, the Js. Consider k « j. Then the point locations for j ∈ J form a set V, those for k ∈ K a set U and cl(U) ⊆. V. In place of (5*) I require:

5**. For every k in K there is some compact set W of Fs, such that:
(a) for any parts u and v of k, there is some f in W, such that f(u) overlaps v, and (b) there is no proper subgroup of F that contains W.[94]

This shows that the associated action of F on the point locations is transitive, as required.

Conclusions about hypervolume and topology

1. Except in the case of Point Discretion where we do not need it, hypervolume may be taken to depend on a linear ordering, namely being of *no-less-quantity-than.*
2. Special and General Relativity cast doubt on the occurrence of frame-independent diameters. This does not threaten our intuitions about diameter because we may explicate the intuitions using the frame-independent Alexandrov intervals.
3. If topological structure is required then I note that Pseudo-set Granules have a non-standard mereology upon from which the topology could be derived. Absent some other structure, such as absolute priority on which topology could be grounded that is an advantage for those that hypothesis.
4. If the aether has no structure in addition to the measure of quantity and the topology, then the characterisation problem is solved in the

---

94 Because of the conditions stated in the previous footnote, (b) is redundant. Also note that (a) could not hold unless the set of all point locations for k in K has compact closure. By the Heine-Borel theorem the fact that we will eventually have a representation using coordinate N-tuples shows that this will be the case provided no k in K has the property of being of infinite extent.

case of granulated aether hypotheses. But in the continuous case, it lacks any entirely satisfactory solution. Ideally we would characterise the aether by noting characteristics of the whole group of symmetries (automorphisms) rather than considering some transitive commutative subgroup. But that requires more structure. In the point-free case the characterisation is rather more complicated. I shall later be arguing that a tenable hypothesis of continuous aether requires the 'reactionary' supposition that the aether is highly symmetric. The characterisation problem for the continuous aether will not therefore have to be solved in the context of topology and quantity alone. Hence, the not entirely satisfactory solutions of the last section will no longer be required.

5. The Problem with Differentiable Manifolds

The theory of differentiable manifolds has great mathematical beauty, but it is a metaphysical abomination. In this chapter I explain what the problem is, and argue that there are only two satisfactory solutions. One requires symmetry, the other granulated aether. In the next chapter I shall argue that granulation is less probable than symmetry relative to the current state of physics, contrary to its a priori appeal. In Chapter Seven I consider symmetric aether, and I shall argue the case for Arntzenius Continuum given symmetry. In the course of doing so, I shall revisit and (again) reject Point Discretion.

1. The characterisation problem

To illustrate the problem, I am supposing that Space-time has the structure of a smooth differentiable manifold, reflecting some – yet to be specified – structure of the aether itself.[95] For ease of exposition, initially I suppose that the aether is point-based and continuous. So it too is supposed to be a smooth manifold.

Smooth manifolds may be defined in either of two equivalent ways. The more elementary is to provide coordinates. The characterisation problem may be illustrated by the special case in which the smooth manifold is point-based with the points in one to one correspondence with the quadruples of real numbers, $\Re^4$.[96]

---

95 For a *smooth* manifold we restrict attention to infinitely differentiable functions. Considering other cases, such a functions with continuous derivatives, does not help solve the problem.

96 More generally, we suppose the manifold has a point-set topology specified by points and a distinguished family of open sets of points, and we suppose it is locally compact and has a countable dense set. Then we consider a finite, or at most countably infinite, collection of open sets $\{U_j\}$ whose union is the whole manifold. We also require that for each U_j there is 1 to 1 onto map f_j: $U_j \to \Re^4$. If two of the open sets U_j and U_k overlap consider the composite

A more elegant characterisation of a 4 dimensional smooth differentiable manifold is to require that there be a ring *Smoo* of continuous real-valued functions on the manifold, such that at each point the *derivative* operators form a 4 dimension real vector space.[97] A derivative operator X_u at point u must satisfy the following:

1. Linearity: $X_u(af + bg) = aX_u(f) + bX_u(g)$, for any f and g in *Smoo* and any real numbers a and b.
2. Leibniz Rule: $X_u(f \times g) = g(u) \times X(f) + f(u) \times X(g)$.

The characterisation problem is that the smooth manifold structure of the aether is being described in terms of something less fundamental than the aether itself. A more precise formulation of the problem is obtained by means of a Euthphroid: are the coordinate functions differentiable because the manifold has a certain kind of structure or vice versa? Mathematicians do not have to answer but, as David Armstrong would say, it is a compulsory question in the metaphysics exam. The problem is a challenge to show how the coordinates or the scalar functions could be more fundamental – or at least as fundamental – as the aether. For coordinates the task is hopeless, because there are infinitely many choices of coordinate systems. Cian Dorr has, however, suggested that the ring of smooth scalar functions might be metaphysically fundamental (Dorr, 2011).

Before discussing some attempts at solving the characterisation problem for smooth manifolds, I compare it with the problem for topological manifolds, discussed in the previous chapter. In that case there was no difficulty in characterising what it is to be a topological space. So I was able to consider the group of automorphisms, that is, the mappings from the space onto itself that preserve the topological structure. The characterisation problem was then soluble provided we could char-

---

mapping $f_k \circ f_j^{-1}: \Re^4 \rightarrow \Re^4$. It will be specified by 4 functions each of 4 variables. Then for a smooth manifold it is required that these functions are infinitely differentiable.

97 *Smoo* is required to be an algebra over $\Re$. Therefore, if f and g are in *Smoo* so is the product $f \times g$ and if a and b are any real numbers, the linear combination $af + bg$ is in *Smoo*.

acterise the group of automorphisms. At that stage, in the point-based case, I required that there be *some* commutative transitive subgroup of automorphisms of a certain kind. Maybe that is a fudge, but I claim the problem is far worse for differentiable manifolds. For by analogy we would consider the group of automorphisms of the ring *Smoo* of smooth functions (i.e.1 to 1 onto mappings preserving addition and multiplication in *Smoo.*). This would be fine if the smooth scalar functions could be characterised independently of differentiable structure. But how?

2 Exotic differential structures – a red herring

Even in one dimension, if we are given the differentiable manifold, there are topological automorphisms (i.e. 1 to 1 mappings of the manifold onto itself such that both the mapping and its inverse preserve the *topological* structure) that do not preserve differential structure. They can, nonetheless be considered diffeomorphisms that map one differentiable manifold onto another, where the two manifolds are the same topological space but endowed with two different differential structures.

A simple example to illustrate the way in which a topological automorphism may fail to preserve a given differentiable structure is obtained by transforming the t coordinate to $t' = \frac{1}{3}t^3$. This transformation is smooth with respect to the old structure but not with respect to the new. For $dt/dt' = t^{-2}$ has a singularity at $t = t' = 0$.

The *exotic* differentiable structures proven to exist by Simon Donaldson (1983) are also cases of the same topological manifold endowed with two different differentiable structures – but ones that are not diffeomorphic. Such exotic structures only exist in 4 or more dimensions. So it might seem we could solve the characterisation problem by supposing the aether is *foliated* and so has for each time, a 3 dimensional spatial structure, avoiding the embarrassment of exotic manifolds.

That solution fails, regardless of whether or not there is a suitable foliation. For, even though General Relativity is invariant under the automophisms of the smooth manifold,, it is, I argue, not sufficient to characterise the manifold up to diffeomorphism, that is up to equivalence of smooth manifolds. Hence we do not need exotic differentiable structures to make the point that the topology does not fix the differential structure.

The reason it is not sufficient to characterise the manifold up to an equivalence is that there is a fact of the matter whether there are singularities and if so where they are. Here I am considering singularities to be the points where the scalar field is not smooth, and in particular considering the case where the derivative is infinite. Consider again the example of the transformation $t' = \frac{1}{3}t^3$. Now consider a scalar field, specified by the function $f(t) = t^2$, or by $f^*(t') = (3t')^{2/3}$. With respect to one of the smooth structures there is a singularity at $t = t' = 0$, with respect to the other there is no singularity

So even though the two differential structures on the same topological manifold are diffeomorphic, one of them implies a spurious singularity. In this case the singularity is restricted to the hyperplane, $t = 0$, which might prompt the response that we should consider the ring of scalar fields that are smooth except on a set of points of zero measure. But that escape is blocked, because we could consider a homeomorphism from new to old coordinate quadruples that is nowhere differentiable, and yet – just to be on the safe side – preserves the Lebesgue measure that represents hypervolume.[98]

3.Scalar fields

The differentiable structure of a manifold may be specified using its topology together with a specification of which real-valued continuous functions are *smooth* (i.e. infinitely differentiable).[99]

The problem with this way of specifying the differentiable structure is that of saying just what the ring of smooth functions, *Smoo*, corresponds to. There are two different ways of trying to solve this problem, the radical and moderate. First consider the radical. We may take the purely algebraic structure of *Smoo*, with their operations of addition and multiplication and including constants. Then *Smoo* is a commutative algebra in the technical sense of mathematicians. The differentiation oper-

---

98 This result holds for n-tuples with $n \geq 2$. See (Panti 2011).

99 See any good introduction to differentiable geometry, such as (Alekseevskij, D. V., Vinograd, A.M. and Lychagin, V. V. 1988).

ations (also known as vector fields) are characterised algebraically using the Leibniz Rule:[100]

If X is a differentiation and f and g any members of *Smoo* then $X(f\times g) = f\times X(g) + g\times X(f)$.

Then the 'points' may be constructed as homomorphisms from the algebra to the real numbers, that is mappings such that $\varphi(f + g) = \varphi(f) + \varphi(g)$ and $\varphi(f\times g) = \varphi(f) \times \varphi(g)$. This interpretation is mathematically beautiful but undermines all our intuitions about both the aether and Space-time, replacing them with things-we-know-not-what, with uninterpreted operations of addition and multiplication (SAD again).

The moderate attempt at a solution to the problem is to grant that the aether has a topological structure, so it is something of which we have an intuitive grasp, and to find an interpretation for the algebra of real valued functions *Smoo*. Are they the actual scalar fields? No, for there might not be enough *actual* fields – the universe might have evolved in a perfectly symmetrical fashion. Indeed the wonder is why it did not if it had a perfectly symmetrical initial state. Say, for instance, that there was a perfect rotational symmetry about some axis. Then the actual fields would all be symmetrical (invariant) under rotations about that axis and so there would not be enough actual scalar fields to be all the smooth real valued functions. On this approach, then, we should acknowledge the lack of enough actual fields and be realists about merely possible scalar fields. Either the algebra *Smoo* then consists of just some out of all possible scalar fields or all of them. Unless we have already characterised the differentiable structure of the aether and so in a position to characterise *Smoo* as the *smooth* scalar fields, the feature that distinguishes members of *Smoo* from other possible scalar fields would be quite mysterious. So we should assume that *Smoo* consists of all the possible fields. That implies that the impossibility of a scalar field having a singularity at a point at which it is continuous but not differentia-

100 Previously I was considering a derivative at a point u, which was a vector X_u 'at u'. I am now considering the vector field X.

ble.[101] Now, I say, necessities are not to be multiplied more than is necessary, so we should be reluctant to restrict what is possible. A further difficulty is that if the occurrence of a certain ring of possible scalar fields is the ultimate physical explanation of the differential structure then there must exist some fundamental scalar fields. Now there are many scalar fields derived from vector and tensor fields but it is controversial whether there are any fundamental ones. The scalar Higgs field is a famous example but it is not fundamental, because, as I understand it, it is the result of symmetry-breaking in the early non-scalar Higgs field.

None of these difficulties is conclusive but taken together they make the prospect of characterising the differential structure in terms of the ring *Smoo* less plausible than it might initially seem.

Does quantum theory help? Suppose we adopt a many worlds interpretation in which the quantum state is specified by a probability distribution over the many 'worlds'. And suppose there is a scalar field, say, the Higgs field. Using the scalar fields in the various worlds we could form the ring of smooth scalar fields *Smoo*, without having to consider merely possible fields. That, however, requires us to compare the values of fields f and g in different worlds but *at the same point*. But this is perplexing. Not only is it counter-intuitive to think of points as individuals distinguished from each other independently of relations to other points, but it is incompatible with the thesis, defended in the Introduction, that the aether is the only fundamental (physical) substance (in the sense of property-bearer). For, if the very same points occur in different 'worlds', these 'worlds' are made up of the same aether having different properties. But we must be realists about these worlds if we are to use them to obtain the ring *Smoo* of smooth scalar fields. And the same stuff cannot have incompatible properties. It is far better to hypothesise that the aether is itself the sum of 'fibres' one for each 'world.[102]' In that case there

---

[101] I stipulate that if the derivative is infinite then the function is not differentiable at the point in question.

[102] Strictly speaking, we no more need such 'fibres' than we need points. The structure of the aether can be gunky across 'worlds' with fields varying smoothly across the worlds just as they do within worlds.

is no way of correlating the points in different ‘worlds’ except in terms of overall similarity of the distribution of properties near the points. But such a correlation will not distinguish the product, say, of fields g and h from the product of g and h∘F where h∘F is the composite of h with a diffeomorphism F.

The above difficulties may be avoided if the scalar fields are interpreted in such a way that the same region of the aether can have many different fields. I now turn to two such speculations, which, however, run into difficulties with the interpretation of multiplication.

4. Two extravagant proposals and a general difficulty

There are two somewhat uneconomic hypotheses that I would take seriously but for a general problem with treating the ring *Smoo* as fundamental, namely interpreting the operation of multiplication.

Proportions of aether

One way in which the aether might turn out to be quite unlike Space-time is if there could be half the aether in a given region, not meaning all the aether in half the region of Space-time, but half the aether everywhere in the region. Hence the aether would have a richer mereological structure than Space-time. In that case, for every part u of the aether we may consider the *regular* part u*, the sum of all parts x disjoint from everything disjoint from u. Then not merely are the reasons for assuming that the regular parts form a classical mereology every regular part may be assumed to consist of all of the aether co-located with u and none in any complement of u.

The proportion of aether will then be described by a function of points of Space-time taking values from 0 to 1. If we consider all the aether in a given region and none outside it that would be described by a function taking the value 1 in a given region and 0 outside, which would not be continuous. We may now consider a primitive property that some portions of aether have, namely being not merely continuous but smooth. Although not all smooth functions describe such portions of aether, those between 0 and 1 do, and these are enough to characterise the differentiable manifold.

The operation of differentiation does not map a representation of a smooth portion of aether to a smooth portion of aether, but on a point-based theory of the aether we can restrict attention to differentiation at a point, which assigns to the smooth portion of the aether a rate of increase of the proportion at the point and the different differentiations at a given point then correspond to the tangent vectors. (Such rates of increase in a direction specify the relativistic analog of an aether current, something that is not at odds with the Michelson and Morley experiment) To talk of the *current* of a portion of aether might seem odd but it is, I concede, a minor extension of the intuitive idea of a current of the whole of the aether. I know, however, of no other application of the idea of a proportion of the aether at a point so this is ad hoc.

Fuzzy regions[103]

Another way of characterising the differentiable structure of the aether is to abandon an intuition that is so basic I have not even listed it, namely that regions are precise. (To be sure a fuzzy region is a precise something or other so I am not positing any vagueness or indeterminacy in reality.) A fuzzy region assigns to each point of Space-time a value between 0 and 1, namely the degree to which the point belongs to the location of the region. So formally the fuzzy regions solution is like the previous one. One difference is that we might well suppose that all the fuzzy regions are smooth. In that case the point p of Space-time is a fiction corresponding to a function ψ from fuzzy regions to real numbers such that $\psi(x\times y) = \psi(x)\times p(y)$ and $\psi(\neg x)) = 1 - \psi(y)$.

A general problem with relying on the ring of smooth scalar functions

If we use the ring *Smoo* to characterise differentiable structure then we should be able to interpret the operations of addition and multiplication in some natural way. As a concession I grant that addition (and multiplication by a constant) are natural operations on any quantity, but what about the multiplication of two scalar fields? If a scalar functions f and g are interpreted as the proportions of aether, what is the product $f\times g$?

---

103 Compare Dorr (2011): 154-155.

There is a natural interpretation of the minimum of f and g, namely the function representing the intersection of the portions of aether. But that is not the same as the product.

This problem of interpretation is fairly general and given the ad hoc character of the two interpretations of *Smoo* considered in this section it eliminates them as serious contenders. If there is a fundamental scalar field required for the physics, then this problem might turn out to be *partially* solved by the physics. For instance, if some fundamental equation required an f^2 term for all scalar fields f, then we may characterise multiplication thus: $f \times g = ½ ((f + g)^2 - f^2 - g^2)$. The operation represented by squaring would then have some unexplained features, such as why the result is never negative, that would be explained if we had a more fundamental interpretation of multiplication. That is why I describe this as a partial solution.

5. Smooth regions

I have contrasted the ease of describing topological structure with the difficulty of describing the differentiable structure. The reason for this is that there are topological primitives such as *separation* that are intuitive, natural spatial concepts. The best way of describing the differentiable structure would be if there was some intuitive concept we could use in an analogous fashion. That mathematicians have not developed the theory in this way should make us pessimistic about the success of such a proposal. In this section I show what goes wrong.

We might try, then, taking the smoothness of a hypersurface, surface or curve as a primitive unanalysable but intuitively graspable property. Or, restricting attention to 4 dimensional regions we might consider *having a smooth boundary* as a primitive property. Or it might be preferable to consider the relation of *smoothly touching* that holds sometimes between two non-overlapping adjoining regions. In this context, I note that on both Borel and Arntzenius Continuum we could take the smoothly bounded regions diffeomorphic to a hyperball to be the globules on which all other regions depend ontologically. So they might already have a privileged position. I shall now argue, however, against this Smooth Region characterisation of the differentiable structure.

If there are points then a necessary condition for function f to be differentiable is that: (1) for every real number *k* the sum of all points x such that f(x) < *k* has a smooth boundary, and (2) so does the sum of all the points such that f(x) > *k*. Let us call this the *smooth boundary condition*. And there is a similar condition for the point-free case. The problem with the Smooth Region characterisation is that it only establishes the smooth boundary condition, which is not a sufficient condition for differentiability. For suppose the manifold has a differentiable structure given by coordinates t, x, y and z, and suppose that g is strictly increasing continuous but non-differentiable real valued function of real numbers. Consider the function h such that if point p has coordinates, t, x, y and z then h(p) = g(t). Then h satisfies the smooth boundary condition but is not smooth.

6 Granulated Aether and differentiable structure

Provided the aether has no more than 6 dimensions the granulated aether specifies up to diffeomorphism a differentiable manifold.[104] This solves any problem posed by exotic differential structures, but not the characterisation problem, which arises even if we have specified the manifold up to diffeomorphism.

Fortunately, granulated aether provides an alternative solution.[105] If we suppose the aether has the structure of a simplicial complex, then we may exploit the analogy between the theory of simplicial complexes and the theory of differentiable manifolds. The scalar, vector and tensor fields on the manifold may be characterized as approximations to suitable properties of, or relations between, granules. For this purpose I shall

---

104 More precisely, given a granule hypothesis in which, for definiteness I assume the granules are simplices (analogs of tetrahedra) we may consider a simplicial complex N that represents the aether in the sense that each cell is represented by a simplex. If there is a differentiable manifold M that is piecewise diffeomorphic to N and if the number of dimensions does not exceed 6 then M is unique up to diffeomorphism. My source for this is (Lurie, 2009)

105 This is one of those pieces of mathematics that are fairly well known but not easy to attribute to any one mathematician.

assume there is fictitious manifold used to represent the granules and on which various smooth scalar, vector and tensor-valued functions are defined. I assume that hypervolume is determined by the gravitational field. The granules may well be represented by regions of the same, small, hypervolume, but we need not assume this.

The granules are represented by sets of the points in a differentiable manifold, and each granule x has a hyper-volume equal to the hypervolume of X the set representing x. Then given any scalar function, f, there is an associated scalar field F that assigns to the granule a scalar property equal to the integral of f over X. Therefore, if we consider any sum of granules, z, the sum of the field values for all the granules in z equals the integral of f over Z the set of points representing z.

Given a (contravariant) vector function **g** on the representing set of coordinate quadruples, we can assign a scalar to any two granules x and y. If X and Y, which represent x and y, have a hyperface in common then the assigned scalar is the integral of **g** over that hyperface. Otherwise it is zero. Intuitively we let the hyperface sweep out a four dimensional region by moving it in the direction of the vector **g** by its length. If the vector makes a very small angle with the surface then the hypervolume swept out is small, but if it is near to 90° then it is greater for the same magnitude of the vector. So the vector function approximates a scalar valued relation between non-identical granules. If the relation between granules x and y has the value G(x,y) then G(y,x) = - G(x,y). For a surface has an orientation and the orientation going from X to Y is the reverse of going from Y to X.

A tensor-valued function **h** may be taken to represent a 3-place scalar relation H(x, y, z) between distinct granules x, y and z. Perhaps the easiest way to describe H is to note that (locally) **h** is the finite sum of tensor products of vector valued functions: $\mathbf{h} = \mathbf{f}_1 \otimes \mathbf{g}_1 + \mathbf{f}_2 \otimes \mathbf{g}_2 + \ldots$ Then H(x,y,z) =

$$\tfrac{1}{3}((F_1(x,y)\times G_1(x,z) + F_1(y,z)\times G_1(y,x) + F_1(z,x)\times G_1(z,y)) + (F_2(x,y)\times G_2(x,z) + F_2(y,z)\times G_2(y,x) + F_2(z,x)\times G_2(z,y)) + \ldots).$$

In this way granule hypotheses replace the puzzling vector and tensor properties by scalar relations.

7. The symmetry solution

Consider again the way I characterized a topological manifold using a suitable group of automorphisms of the topological structure. The reason that this cannot be generalized to provide a characterization of a differentiable manifold is that we must already have characterized differential structure in order to specify the relevant automorphisms. We may avoid that circularity problem by supposing some structure such that the transformations that preserve it will turn out to preserve the differentiable structure as well. For instance if there is a metric structure in the strict sense of a metric (i.e. one satisfying the Triangle Inequality) then we may consider isometries, transformations that preserve that metric. I am, however, concerned with a light cone structure.[106] So I consider the group of symmetries *Symm* of all 1 to 1 mappings of the set of parts of œ onto the set of parts of œ such that every g in *Symm* preserves: the mereological structure (i.e. parthood), hypervolume, the topological structure (i.e. separation) and the light cones.

The group of symmetries, *Symm,* is a topological group, with the topology such that a sequence g_1, g_2 . . . of members of *Symm* converges to the Identity mapping Id if:

> For any two separated regions u and v, there is some integer N such that, if n > N, u is disjoint from g_n(v).[107]

Provided *Symm* is not too large, we may then characterise the differential structure on *Symm*, turning it into a Lie group. For instance it suffices that there is some open neighbourhood of Id that has compact closure, is connected, and contains no subgroup of *Symm*. That last constraint (‘no small subgroups’) is plausible only if there is no fine structure, or we choose to ignore it. That there are open neighbourhoods with compact closure is, roughly speaking, the requirement that *Symm* be finite dimensional.[108] The immediately relevant feature of a Lie group is

---

106 By a cone I mean a solid cone, not just its boundary.

107 For the theory of topological groups see (Pontrjagin, 1939).

108 That these constraints suffice to show that *Symm* is a Lie Group follows from Hidehiko Yamabe’s (1950) solution to Hilbert’s Fifth Problem, based on earlier work by Andrew Gleason, and by Deane Montgomery and Leo Zippin.

that it has a unique differentiable manifold structure such that both the composition of symmetries and the mapping of each symmetry to its inverse are diffeomorphisms. Provided *Symm* acts transitively on Space-time this specifies its differentiable manifold structure. (To say it acts transitively is to say that for any points of Space-time x and y there is some symmetry g such that g(x) = y. This presupposes that Space-time is point-based even if the aether itself is not.)

For instance, suppose in fact Space-time has the structure of Minkowski Space-time. Then: (1) the group of all symmetries *Symm* is isomorphic to the Poincare´ group P; and (2) the group of those symmetries leaving a given point p fixed, *Stab*(p), is isomorphic to the Lorentz group L. Then Space-time may be represented by the *quotient space* P/L.[109] The future light cones turn out to be in one to one correspondence with the members of P/L. If the aether is itself point-based that is because there is just one future light cone with vertex at a given point p. Otherwise we may construct point locations in such a way that there is just one future light cone with vertex a given point location.[110] In this case the stabiliser is a normal subgroup of *Symm* and so the quotient is itself a group and P/L is isomorphic to the additive group of quadruples of real numbers.[111]

In this special case we could have characterised the differentiable manifold structure more directly by noting that there is a unique commu-

---

109 There is an equivalence relation on P, with g and k equivalent if $g \circ k^{-1} \in L$. Then the members of P/L are the equivalence classes.

110 For every future light cone x there is a corresponding past light cone x^-. So there is a property that belongs to anything whose location includes some region y such that if z is the (or more generally any) complement of y, $x \wedge z$ is separated from $x^- \wedge z$. This is the property of being located at the (fictitious) common vertex of x and x^-.

111 Because H is a subgroup of G, H is closed under the operations of multiplication and taking inverses. In addition because H is a *normal* subgroup, for every $g \in G$ and $h \in H$, $g \circ h \circ g^{-1} \in H$. Therefore, $g \circ k^{-1} \in L$ iff $k \circ g^{-1} \in L$ and we may define the product of equivalence classes as the equivalence class of products.

tative transitive subgroup *Comm* of *Symm*. Then the point locations are in one to one correspondence with the members of *Comm*. In the case being discussed, *Comm* is isomorphic to the additive group of quadruples of real numbers and so the differentiable structure of Space-time may be specified as that which *Comm* has as a Lie group.

Provided there are neither too many nor too few symmetries this method is quite general. The case of a general topological manifold is one of too many symmetries, because the group of automorphisms is infinite dimensional. The case of General Relativity with the gravitational field interpreted as due to curvature is one of too few symmetries, that is, not enough to ensure transitivity, because in general the only symmetry is the identity map. That is because the details of the gravitational field depend on the detailed distribution of mass/energy, and so even if the gravitational field around the Sun is approximately the same as that around some other star the precise field depends on, among other things, the posture and gestures of every inhabitant of Earth, and so has a negligible probability of being replicated even on a star with an inhabited planet orbiting it.

I have found, then, only two ways solving the problems raised by the differentiable manifold structure. One requires granulated aether. The other requires a highly symmetric structure, for instance that of Minkowski Space-time. In Chapter Seven, I show how we may relax the symmetry requirement enough to allow for the rejection of realism about the future. But the symmetry constraint is still a somewhat strict one. In Chapter One, Section 3, I argued that granulated aether is rather unlikely to be highly symmetric, so the two disjuncts are excusive: granulation or symmetry but not both. I conclude this chapter with another argument for the same conclusion. It is an ad hominem in Locke's sense of being audience-specific. The audience in question consists of those who reject dynamic theories of Time.

8. The Hole Problem

The Hole Problem, which delayed General Relativity from 1913 to 1915, was re-introduced by John Stachel and subsequently presented as an argument against realism about Space-time by John Norton and John Earman (Norton 2008). Given my terminology it is an argument against the reality of the aether, one to which I need to reply. We assume, as is plausible, that if the aether is real then there is a fact of the matter as to whether a given macroscopic object, a star say, is in a given region. We next note the gauge freedom of General Relativity, namely that the states defined by the distribution of the energy-momentum and the gravitational field can neither be inferred from observation nor determined by Einstein's equations. This is quite general, but is most easily illustrated in the case in, which we expect Special Relativity to be an excellent approximation. This would often be described as the case of nearly empty Space-time with just a few small well-separated rocks in it, although of course I say that Space-time would not be nearly empty, but full of the aether. General Relativity permits a state described as just such an approximation to Special Relativity, in which the aether is almost flat everywhere. There are, however, infinitely many other state-descriptions permitted by the general relativistic equations, including those in which the aether is almost flat outside the 'hole' and highly curved inside it. (The problem arises because General Relativity implies six independent equations but the gravitational field requires ten scalar functions to specify it.) *Other things being equal* it is reasonable to conclude that two state-descriptions that are indistinguishable in this way are in fact two descriptions of the same state. But if we do draw this conclusion then it is said there is no fact of the matter as to whether a given event, say the collision of two rocks, is located in a given region, for such locational 'facts' depend on which description is used.

The most common response to the Hole Problem, goes back to Einstein, and echoes Leibniz, so I call it the Einstein-Leibniz solution. It is to deny that the different state-descriptions correspond to different ways the world could be (Norton 2008 §9). It is this solution that is incompatible with the reality of the aether. It permits us to re-describe a world in which light goes in almost straight lines so that it curves considerably in the 'hole', but we are told not to get upset, because the one event in the hole, namely the collision of the two rocks, will still *seem* to be at a definite place.

Now we always used to assume that light in a vacuum travelled in straight lines. To be sure we were later told that this was not quite right: light gets bent a little – and near black holes gets bent in extreme ways.

And I grant that we should be humble and stand corrected: our pre-scientific world-view is in all sorts of ways open to revision. In spite of that concession, I reject the above Einstein-Leibniz solution. That light travelled in straight lines was a presupposition of astronomy, and of optics more generally, for over two thousand years (and maybe for as long as modern humans have been around to think about the stars.) We may grant that this is only approximately correct, in much the way that we grant that Newton's theory of gravity is only approximately correct, but once we allow that describing reality with light going in curves is as valid as it going in almost straight lines an unacceptable scepticism follows. Think how it applies to the astronomy of the Solar System. Tycho Brahe, the story goes, was asked by young Kepler what the orbit of Mars was and Brahe shook him saying '*That* is the orbit of Mars'. Kepler went on to discover that planets moved in (near) elliptical orbits around the sum. But the Einstein-Leibniz solution tells us something quite different. Shake the orbits as violently as Brahe shook Kepler and that is as correct as the ellipse, being either just another hypothesis that is as likely or just another way of describing the same orbit.[112] For on the Einstein-Leibniz solution, it just a computational dodge to insist that light goes in almost straight lines.

A related objection to the Einstein-Leibniz solution is that Special Relativity is intended to approximate General Relativity locally. That was the motivation for giving each tangent space a metric just like that of Minkowski Space-time. But the Einstein-Leibniz solution tells us that in the case we intuitively expect to be almost flat, however small a region may be, it is as correct to say that in this region the aether deviates greatly from flatness as to say it is almost flat.

It has been suggested that the Hole Problem is just another case of gauge invariance like the addition of the gradient of an arbitrary scalar function to the 4-potential used in electromagnetic theory (See Norton 2008, §10.3). There are, however, two differences that undermine this 'companions in guilt' defence. One is that no amount of hanky panky with electromagnetic field theory offends common sense, which lacks

---

[112] I am not saying that all continuous orbits are permissible on this account; merely that it permits wiggly ones.

opinions on that topic. The other is that the gauge invariance of electromagnetism is local, and, if anything, serves to support point-free theories rather than undermine realism about the aether. For given any open region, however small, we may consider the electromagnetic field to be specified by a vector-valued function on it with equivalence up to the addition of the gradient of a scalar field on it. So the field may be analysed in terms of the electromagnetic properties of arbitrarily small regions. By contrast the gauge invarianceof Einstein's equations is non-local.

There is a pattern to the history of physics concerning relativistic implications. Theories are often initially formulated in a relativistic or perspectival way. The introduction of more sophisticated mathematics then removes the relativistic implication. That is what happened with Special Relativity and the same has happened to gauge theories such as electromagnetism. The Einstein-Leibniz solution to the Hole Problem cannot therefore find robust precedents in gauge theories.

A fairly straightforward solution to the Hole Problem is to restrict attention to solutions of Einstein's equations in which the curvature is spatial rather than spatio-temporal.[113] But that requires a foliation of the aether into a succession of Space-like hypersurfaces. This would cohere nicely with dynamic theories of Time, making sense of the idea that the rate of passage of Time is not uniformly the trivial second per second.[114] Otherwise it is open to the criticism that it is an ad hoc complication.[115] So those who reject dynamic theories should not adopt this solution.

---

113 We could take the manifold to be of the form $\Sigma \times \Re$, where Σ is 3 dimensional, and $\Re$ is the real line.

114 Somewhere, Greenwich presumably, time passes at a second per second, just as, in the old days, a standard metre bar in Paris was a metre long. But if the hypersurfaces of the foliation are not parallel then those that are 1 second apart at Greenwich could be more or less than 1 second apart elsewhere.

115 Unless one of the Einstein Aether theories turns out to be empirically verified. (Jacobsen 2008). These are theories that attach physical significance to a vector field specifying the foliation. I fail to see, though, why we should think such complications are warranted.

Granulated Aether provides an interesting variant on the foliation solution, and one that I judge to be less open to the criticism that it is ad hoc We might plausibly take the granules to be pentatopes, four-dimensional analogs of tetrahedra, with four vertices Space-like separated and forming a tetrahedral hyperface, with the fifth point separated from these four vertices in a null (light-ray) way. The fifth point is thus either future or past of the spatial tetrahedron, with both temporal orientations occurring. In that case, the metric structure of Space-time is specified by the lengths of the six sides of the tetrahedra, which are, therefore, not regular. When the granules are represented in a continuous Space-time then the tetrahedra will make up spatial hypersurfaces. The stretching of regular tetrahedra required to turn them into the irregular ones representing the spatial faces of the granules, specifies a tensor, and hence the metric field. Because we took four of the ten edges to be null there are only six degrees of freedom, solving the Hole Problem and establishing the granulated aether disjunct.

The other disjunct, symmetry, is based on the most straightforward solution to the Hole Problem. This is to hypothesise a flat, or otherwise highly symmetric, aether with an associated 'empty Space' gravitational field and require the deviation from this to be as little as compatible with General Relativity.[116] The idea is that the flat case requires no explanation and General Relativity explains deviations from the flat case. On this interpretation, the gravitational field is like any other field except

---

116 If the deviation from the special relativistic metric is given by a symmetric tensor field, Δ, then Δ(p) may be thought of as a linear map from the tangent space at p to its dual. Hence if Δ*(p) is the dual of Δ(p), the composite Δ(p)*∘Δ(p) is a linear map from the tangent space at p to itself. The scalar field, trace(Δ*∘Δ) is thus a scalar field measuring deviation from the special relativistic case. To retain General Relativity as strictly accurate we first consider solutions to Einstein's equations and then take the one that minimises this scalar field to be correct. A more elegant theory would be obtained by adding a small multiple of trace(Δ*∘Δ) to the Lagrangian for General Relativity.

that it is a symmetric tensor field.[117] Therefore it characterises something with the same mathematical properties as a metric.[118] Because there is a background of flat Space-time, we may stipulate that the gravitational field differs from that background by as little as possible.

---

[117] Symmetric in the sense that **g**(x,y) = **g**(y,x) for all points x and y.

[118] The gravitational field proper when added to the metric for ‘empty Space-time’ provides the ‘metric’ of Einstein’s theory.

6. Contemporary Physics and Discrete Aether

In this chapter I consider some theories that are prima facie relevant to the structure of the aether: String Theory, and more generally Supergravity theories; Loop Quantum Gravity; Causal Set theory; and Dynamical Triangulation.

String Theory, along with other Supergravity theories predicts particles that are yet to be discovered. Absent empirical confirmation, these theories, generated by the Consistent Renormalisation research program, as I call it, do not support a continuous aether theory. Rather they presuppose such a theory. On the other hand, Causal Set theory presupposes discrete aether, and its significance is largely that as usually formulated it presupposes Point Discretion. I shall argue that it becomes more plausible if adapted to Granulated Aether, and so has the same ontology as Dynamical Triangulation. As for Loop Quantum Gravity, the bad news is that in spite of appearances it offers no support for discrete theories – nor does it disconfirm them.

There is not much joy, then, regarding the empirical evidence for or against discrete aether. I shall, however, obtain three results. The first, with which I begin the chapter, is that contemporary physics undermines an a priori argument against discrete theories of the aether, which I call the Argument from Scale Invariance. The second result is that discrete theories of the aether are threatened with a serious non-locality problem, and that Granulated Aether is superior to Point Discretion because of the way it deals with this threat. The third result shows that discrete aether is a high risk, high reward hypothesis, with symmetric continuous aether being low risk and low reward. That is because, I argue, discrete aether is committed to geometrodynamics, the thesis that all four fundamental forces are to be understood as a consequence of the detailed shape of the aether.[119] By contrast, symmetric continuous aether requires the treatment of gravity as somewhat similar to other forces. The differences between the two ways of unifying physics are:

---

[119] Maybe research on quantum gravity will lead to discrete geometrodynamics. An alternative approach is that of Stephen Wolfram (2002).

1. The combination of Pseudo-set Granules, as the preferred version of Granulated Aether, with geometrodynamics promises a theory that is fundamentally mereological, with all other structure depending upon the mereology. There is also the hope of simple laws expressible in terms of positive integers. The result would be a triumph for simplicity. (That is the high reward). It would require a great deal of mathematics to relate this underlying structure to physics as we now have it, and it would be rash to assume this mathematics even exists (whether or not mathematicians discover it).
2. Treating gravity much like any other force is straightforward in principle, and there are interpretations of quantum field theory that ensure consistency. I anticipate that the resulting unification will require several parameters to be adjusted to fit the empirical data, and so, if it is successful, it will not be as strikingly confirmed as discrete geometrodynamics would be.

Readers may not agree with my overall assessment that discrete geometrodynamics is less probable precisely because it is bolder.[120] So my chief conclusion is disjunctive: The aether is either granulated, in which case probably Pseudo-set Granules and geometrodynamics are correct and simplicity triumphs, or the aether is symmetric Arntzenius Continuum, the hypothesis to be developed in the next chapter.

1. Undermining the Argument from Scale Invariance

Theoretical physicists rely heavily on the metaphysical principle that necessities should be highly symmetric. Thus the thesis that the shape of Space-time might have asymmetric geometry – Einstein likened it to the surface of a potato – is tied to the thesis that the general relativistic metric is (nomologically) contingent, and that the necessary structure of Space-time is merely its structure as a differentiable manifold, from which is derived the requirement that the laws of nature be covariant,

---

120 Karl Popper praised bold conjectures. He was also an avowed sceptic about ordinary induction. Those of us who are not sceptics could argue that bold conjectures are confirmed surprisingly often and so are not as improbable as we would judge a priori.

that is statable in terms of the differentiable structure. This is a strong symmetry condition. Alternatively, if we require a necessary geometry such as that of Minkowski or de Sitter Space-time, then the requirements of isotropy (no privileged direction) and homogeneity (no privileged position) amount to symmetry conditions on the physically necessary.

One candidate symmetry is *scale invariance*, the thesis that, if all spatio-temporal distances were multiplied by a scale factor, this would leave the necessary structure of the universe unchanged. Taking the gravitational constant to be fixed, the empirical equivalence of our universe to one like ours but uniformly scaled up by a factor μ may be achieved only if there is also an accompanying multiplication of mass by μ. So when I consider re-scaling I assume spatio-temporal distance and mass are both multiplied by some positive μ.

The case for such scale invariance is that asymmetries are not to be multiplied more than is necessary. This is not an especially strong case, so it is worth noting the spurious empiricist argument for scale invariance that might have seemed to buttress it. If everything, including we ourselves, was scaled up or down, who could notice the difference?[121] This rhetorical question presupposes that the aether is not discrete. For if it is discrete we may use the number of aether atoms as an absolute measure of quantity and scaling observers down so that they are of the scale of an aether-atom would be noticeable if, per impossibile, such scaled-down observers could exist. Therefore, it should not be used to argue against discrete aether.

Granulated Aether and Point Discretion do not violate scale invariance but rather make it vacuously true. For, absent unnecessary complications, they both imply an absolute measure of quantity and hence prohibit any change of scale. Such trivialisation is, however, contrary to the a priori appeal of scale invariance. I conclude that, unless undermined, the Argument from Scale Invariance provides a case against discrete theories, although not an especially strong one.

The way to undermine the argument is to show that scale invariance should be abandoned even on continuous aether hypotheses. For

---

[121] The topic was discussed by, for instance by Schlesinger (1964) and Grünbaum (1964, 1967).

instance the cosmological constant, Λ, would be affected by a change of scale unless it is exactly zero, contrary to the current consensus. The problem with this way of undermining the Argument from Scale Invariance is that Λ might be considered nomologically contingent, and the consequence of the conditions early on in our part of the universe. There are many other examples of important physical constants that would have been different given a change of scale. Among these are the coupling constants that compare the strengths of various fields, for instance the fine-structure constant, α. It is to be hoped that all these will be derived from a 'theory of everything'. This forces us to speculate about whether an as yet undiscovered theory will be scale-invariant. What we should expect depends on how firm is our intuition that symmetry should be maximised. Given a firm intuition we should expect a scale-invariant theory of everything, but if the intuition is less firm we should be agnostic on that topic. Hence the argument from scale invariance against discrete theories of the aether merely supports my disjunctive conclusion, namely that if you back symmetry over simplicity you should hold the aether to be continuous. Otherwise, there is no case for scale invariance, and hence no obstacle to discrete theories of the aether.

2. The need to derive Planck scale discrete aether

If Granulated Aether is correct we might expect the hypervolume of a granule to be about 1 Planck unit, in which case their density in Planck units will be some constant ρ, which is of the order of magnitude 1 Planck unit, that is about 10^{173} granules per sec^4. If Point Discretion holds then the aether atoms have zero hypervolume but we might again expect a density ρ of about 10^{173} per sec^4.

We should accept this estimate for the density of aether atoms given a discrete theory. For, I say, we should minimize the number of fundamental, that is, non-derived, physical constants – the stuff of fine tuning. To be sure, current physics is plagued with a proliferation of physical constants but the hope is that all or most of these are not genuine constants and vary from domain to domain of the universe. Moreover, humble metaphysicians that we are, we will be guided by physicists and take G, c and h to be non-contingent. Hence the density of aether atoms ρ would be a physical constant. It is not plausible that ρ vary from domain to domain. In fact so intuitive is it that the hypervolume of a granule is a natural unit that if ρ's expression in Planck units varied, then we would take that as implying that one of c, G or h varied, something we are here denying. Hence the only way to avoid ρ being an unwanted fundamental physical constant is to derive it from the overall theory. An example of such a derivation is provided by the discrete interpretation of Loop Quantum Gravity, discussed below which, suitably interpreted, implies that the minimum area of a loop is $4\log_e 3$ (approximately 4.4) Planck units. (Baez 2003). How we interpret that minimum area will depend on both the shape of the aether granules and the local curvature of Space-time, but it is plausible that it is of the order of magnitude of a Planck unit, and, more important, determined by the structure of the aether.

So why should we be reluctant to admit non-derived physical constants? Because theoretical understanding is a matter of explaining the more complicated in terms of the less complicated. The number of non-derived constants is a partial measure of the complexity of a hypothesis and so, other things being equal, should be minimized.

Could we use the fine-tuning required for sentient life to undermine the case against the non-derived density of aether atoms? [122] Not if we reject, as I have done, the variability of ρ in a given universe. For if ρ has a precise value, that it has this value rather than some other, almost equal, value (also suited to life) should be derived from the rest of the theory.

I conclude, then, that if the aether is discrete, ρ, the density of atoms per Planck hypervolume, should be predicted by the theory.

---

122 There is an extensive literature on fine-tuning in the philosophy of religion. (See Ratzsch 2012). As a case for the existence of God I prefer Coarse Tuning (Forrest 2007).

3. A challenge for discrete theories

In the previous section I argued that a satisfactory quantum gravity based on discrete aether should be able to derive the density of atoms. There are two extreme cases. In the first we take a theory that does not support the discrete character of the aether, such as String Theory, and we nonetheless hypothesise that the aether is discrete. In that case, we are adding ρ to the list of non-derived constants, which is a disadvantage compared to the case of continuous aether in which: (1) the most natural units are Planck units, so none of G, c and h count as non-derived physical constants; and (2) the density of atoms ρ is the only value it could have for a continuous theory, namely infinity.

There would be an even more serious disadvantage if the theory required structure at a scale much less than the Planck units. For instance, if String Theory required the strings to be of the order of a Planck unit in length, then to be strings and to vibrate we might hypothesise sub-Planck scale structure. And that would prevent the identification of the natural units with the Planck units. String Theory is, however, quite compatible with the strings typically being about 100 Planck units in length.[123] We might then take the strings and branes to be one Planck unit thick.[124] I conclude that if String Theory turns out to be correct the case against discrete aether, although of some weight, is far from conclusive.

The other extreme case is that in which taking the aether to be discrete we can derive an upper bound for ρ. For in that case, the discrete character of the aether explains why $h \neq 0$ and hence why classical theories are incorrect, something presupposed but not explained by the use of

---

123 At one stage it was assumed the strings would be of Planck scale length but later it was realised they would be a couple of orders of magnitude longer than that.

124 Even if the strings were a Planck unit in length there might be a way out, namely take the strings, branes etc to be specified by patterns of aether atoms rather than consisting of them. Consider, for instance aether atoms laid in a helical pattern around an axis in Space-time. Then the axis is distinguished without being made up of aether.

Planck units. This would be a significant advantage for discrete theories of the aether.

4. Against String Theory and Super-gravity.

The strings and branes of String Theory are thin, but as far as I know there is no way of deriving their thickness or even of showing that they have non-zero thickness. Hence the argument of the previous section shows that theory to be somewhat improbable given discrete aether. The same holds for other Supergravity theories. So we have a case for not combining these theories with discrete aether. Maybe the disjunction of String theory and other Supergravity theories will be confirmed by the discovery of the novel particles they predict, such as gravitinos. That would, therefore, provide a case for continuous and hence symmetric aether, and also for Arntzenius Continuum. I now argue, however, that absent such empirical confirmation, String Theory and Supergravity are improbable – even relative to the hypothesis of continuous aether. My argument is that they are based on a research program the core of which *presupposes* (1) that the aether is continuous and (2) that there are no particles that are extended in all dimensions (i.e. of the same number of spatial dimensions as Space itself) Given the complexity of String Theory and of Supergravity more generally we should, I now argue, abandon this core, *unless some empirical predictions are confirmed.*

String Theory and other Supergravity theories belong to what I call the Consistent Renormalisation research program, which splendidly illustrates Imre Lakatos' methodology (Lakatos 1970). The *core* is the requirement that the standard approach to quantum field theory can be developed without divergent (i.e. infinite) integrals. This is in response to the unsatisfactory character of arbitrary and ad hoc cut-offs by which the energy and momenta of particles are bounded. (For instance to avoid the 'ultraviolet catastrophe' the cut-off ensures that the frequencies are less than some arbitrarily chosen very high value.) String Theory started with the brilliant idea that we might use strings to provide a theory of quarks and the associated strong nuclear force without cut-offs. It turned out that strings were not needed for that purpose, but meanwhile an anomaly was discovered, namely that String Theory predicted that some *closed* strings (i.e. ones closed into a loop and so without free ends) would have as excitations massless spin 2 particles, something not required for quark theory. This anomaly was then inverted into a ‘prediction’, by noting that gravitons (quantized gravitational waves) would be just such massless spin 2 particles. Consequently String Theory re-established itself as a theory of everything, that is, a unification of the theories of the electromagnetic, weak strong and gravitational forces. Another anomaly resulted from the discovery that there were five different string theories that could be converted into each other using transformations, known as dualities. (This is reminiscent of the notorious wave/particle duality, said to have been resolved by quantum field theory.) To resolve the duality, that is remove the anomaly, it was proposed that there must be a more basic underlying theory, which is known as M-theory and which posits *branes*, higher dimensional sheets to which the open strings are attached. This M-theory has now produced another anomaly, namely the proliferation of *landscapes*, that is models for the ‘vacuum’, that is, the matter-free universe. There are estimated to be about 10^{500} of these landscapes, in part because of the proliferation of the 6 dimensional manifolds describing the dimensions additional to the experienced 4 dimensions of Space-time.[125] Some hope that of these landscapes only one will be suited to life.

As is well known, String Theory requires 6 extra dimensions. It is also a Super-gravity theory, requiring super-symmetry, a consequence of which is that the familiar particles (bosons and fermions) have super-partners (fermions and bosons respectively). For instance not only are gravitons implied by the theory but also *gravitinos*, fermions of spin 3/2. Some of these super-partners are predicted to be discovered using the Large Hadron Accelerator. If they are not found, then this will be a further anomaly and it remains to see whether the research program could invert it into some further startling prediction.

---

125 These are vacuum solutions of the higher dimensional analog of General Relativity. Being massless they are ‘Ricci flat’. They must also be compact. This forces on them a unique (differential) topology, that of a torus in 6 dimensions, but there is still an enormous proliferation of general relativistic metrics. See (Douglas and Kachru, 2007).

The obvious criticism is that String Theory and Supergravity are just too complicated to be a fundamental theory, which as a 'theory of everything' it would have to be. This criticism may be developed into what I call the Modeling Objection, namely that a sufficiently rich and interesting mathematical structure has the resources to model an enormous range of fairly simple theories, but its success in so doing does not support the hypothesis that these are more than models. For instance a Quinean philosopher might claim that the fundamental structure of physics is that of set theory. And to be sure we can almost certainly model the correct physics set-theoretically if we can model it using mathematics. But that is a cheat, precisely because set theory is an almost universal model for mathematics. Likewise my complaint that String Theory is complicated may be developed into the argument that ten dimensional geometry is just so rich we expect it to model an enormous variety of theories, but that this too is cheating. In Lakatosian terms, I say that the research program degenerated as soon as the excessively rich 10 dimensional geometry was invoked, although the researchers were not aware of this richness at the time.

Here I distinguish two rather different appeals to simplicity. Even theories that are not fundamental should be relatively simple, that is as simple as we can make them given the empirical constraints, but for a fundamental theory we expect an absolutely simple theory. Or at least that is my metaphysical intuition, which I hope readers share. Chemistry in the year 1900, for example, seemed as simple as the evidence allowed but, because of the proliferation of chemical atoms, not simple enough to treated as fundamental, even at the time.

I have appealed to a metaphysical intuition concerning simplicity, but intuitions are defeasible and, I take it, proponents of String Theory consider that the intuition is defeated by the sheer difficulty of providing a consistent quantum field theory that would count as a 'theory of everything'. The history of the research program is one of force majeure: twists and turns required for consistency. When I say 'required' for consistency, I mean 'required to avoid other complexities'. For, as Jack Smart is fond of saying, any theory can be rendered consistent with enough ad hoc qualifications.

The underlying principle is correct: if respect for the mathematics and the empirical constraints force complexity upon us then we are wrong in assuming the theory of everything must be simple. But this principle requires an extraordinary thoroughness if it is to be applied correctly. For it may happen that because of the accumulated complexities we should reconsider one of the earlier stages in the program.

There is a 10^{-33} cm elephant in the room – or is it a 20,000 nanogram gorilla? Discrete theories of the aether effortlessly prevent the ultraviolet catastrophe by putting an upper bound on frequencies of 10^{45} sec^{-1}, the frequency of a gamma ray whose wavelength is 10^{-33} cm and whose energy is equivalent to 20,000 nanograms. Even if the aether is continuous, the ultraviolet catastrophe may be avoided by positing that the fundamental particles are extended regions of the aether with suitable properties – presumably their diameters will be of the order of magnitude of the Planck length. Their extended character forces upon us an upper bound to the particle density and hence the ultraviolet catastrophe is avoided.[126] If there is an infrared catastrophe, then we can avoid that by resort to a spatially finite Space-time. The Consistent Renormalisation research program turns out to be based on a core of conservatism, with a reliance upon point particles. Initially that was sound methodology, but the excessive richness of 10 dimensional geometry has shown it has degenerated.

Nonetheless, I wish it well as a fun-filled research program, and if the prediction of super-particles is confirmed then the fact that these are implied by String Theory and Super Gravity, but not rivals, would have confirmed them. A corollary is that if they turn out, improbably, to be thus confirmed, then we should not combine them with discrete aether because the implications of String Theory are based on the twists and turns of a research program that presupposes the aether is continuous.

---

126 Compare the ideal gas law PV = kT (Pressure times volume equals a constant k times temperature). Even if we ignore the forces between molecules we need to make a correction to allow for the molecules not being point particles obtaining P(V – b) = kT where b is the sum of the volumes of the molecules themselves.

For without continuity, the ultraviolet catastrophe would never have occurred in the first place.

5 Does Loop Quantum Gravity imply discrete aether?

Loop Quantum Gravity is based on the hypothesis that the correct quantum gravity will result from the (canonical) quantisation of General Relativity. It is cautious in that it is not intended as a 'theory of everything', and such caution increases its probability. Just how probable it is depends on the plausibility of canonical quantisation. The least kind judgement on canonical quantisation is that it is a recipe for coming up with one out of, for all we know, many possible theories that have the required classical limit in contexts in which Planck's constant may be ignored. To this can be added the admittedly weak support given by ordinary induction: canonical quantisation has worked fairly well in the cases of electrodynamics and chromodynamics (quark theory).

My interest in Loop Quantum Gravity derives from the way it implies a discrete *spectrum* for area and volume, a result that suggests the aether is discrete, even though the theory being quantised is not discrete. I say 'suggests' for just what a spectrum of an 'observable' is requires further, and contentious, interpretation. If we follow that suggestion we may argue by reductio ad absurdum. Within the scope of the supposition that the aether is continuous, Loop Quantum Gravity would arise, it is said, as a quantisation, which then would imply that the aether is not, after all, continuous. I shall argue, however, that the suggestion should be ignored.

Loop Quantum Gravity is motivated by the difficulty in quantising the way the aether curves. The difficulty is that quantum theory requires a volume-analog assignable to sets of the states in the *configuration space*, which in this case consists of all the physically possible ways the aether could be curved at a given time-coordinate t. (It is to be hoped that resulting quantum gravity does not depend on the choice of the t coordinate.)

In the case of a single particle the configuration space is Space itself and the volume-analog the volume itself. Then the pure states correspond to the Hilbert space of square-integrable complex functions on the configuration space. It is easy to find volume-analogs for a given finite number of particles and then by considering more and more particles consider the case of any finite number of particles. The resulting Hilbert space is Fock space, used for instance in quantum electrodynamics. But how can we find a volume-analog on the configuration space for General Relativity?

Loop Quantum Gravity derives from a way of characterising the gravitational field due to Abhay Ashtekar, Giorgio Immirzi and Fernando Barbero, which enables us to find an appropriate volume-analog (Thiemann 2003: I.22). The idea is that curvature may be characterised by the way in which a quadruple of independent directions varies as it is transported around a closed loop. This is analogous to noting what happens to a pair of directions at, say, the South Pole if it is transported around a loop that goes up longitude zero to the equator, around to longitude 90 and then down again to the South Pole. It ends up as a different pair of directions because of the Earth's curvature. The transformation of the directions on the Earth's surface is given by a rotation, In the case of General Relativity the transformation is specified by a Lorentz transformation.

If we consider a family of loops joined up to form a *graph* with a finite number of vertices, the way the axes transform as we go round the graph gives us more information about the curvature the more loops we add in a given region. By considering a graph with a countable infinity of vertices in a given region the transformation associated with loops connecting any finite number of them can specify the curvature completely. Suppose we now take a set of four coordinate axes for a trip around the loop. Because of the curvature the overall result is that the axes undergo a transformation. At the cost of some arbitrariness (gauge freedom) we may analyse these transformations resulting from going around loops using Lorentz transformations as we take the axes from vertex to vertex. The result is that we may specify the general relativistic metric using a graph with countably many vertices, and hence countably many edges, assigning to each edge a Lorentz transformation.[127] Hence

---

[127] More generally, if we allow a connection with torsion – needed for taking intrinsic angular momentum (spin) into account – then the Lorentz group is replaced by the larger Poincaré group. This is important because it has SU(3)

the metric for General Relativity is described by assigning a Lorentz transformation to each of countably many edges. As I understand it, Loop Quantum Gravity is the result of 'quantising' the system by assigning an irreducible representation of the Lorentz group to each edge.[128] This procedure is analogous to the quantisation of electrodynamics where the quantum system is described using a countable infinity of one particle states. The chief difference is that we replace particles by edges.

In the resulting theory both the area of a loop and its volume have discrete *spectra* with minimum values obtained at the Planck scale. The precise values depend on the details of the theory. This result is suggestive of discrete aether, but merely suggestive, for the spectrum is not defined as the set of possible values but rather the set of eigenvalues. To obtain more than a mere suggestion, we would need some extra premise, such as the following *Definite Range principle:*

> The value of an 'observable' is definitely in the range from the greatest lower bound to the least upper bound of the eigenvalues.

Whether that principle holds depends on which quantities are given the status of 'observables'. But the case of spin shows that we should not grant that status too readily. For consider the values S_x, S_y and S_z of the x, y and z component of the spin of a spin ½ particle. And suppose we grant observable status to their squares S_x^2, S_y^2, and S_z^2. These quantities have only one eigenvalue, $(h/4\pi)^2$, so by the definite range principle all would definitely have that value, showing that $S_x = \pm (h/4\pi)$, $S_y = \pm (h/4\pi)$, and $S_z = \pm (h/4\pi)$. So the spin-vector points from the centre of a cube of side $h/2\pi$ with edges parallel to the three axes, to one of the eight corners. This is absurd, because the choice of x, y and z coordinates was arbitrary. What has gone wrong? A quibble is that the spin of a spin ½

as a finite dimensional representation, and SU(3) is the symmetry group for unified strong/electro-weak theory.

128 The representation is by means of operators on the one-dimensional subspaces of a Hilbert space and hence corresponds to a representation by means of unitary operators of the covering group SU(2). These irreducible representations correspond to the spin states of particles of spin 0 (trivial), ½, 1 etc.

particle is not really a vector, but we could consider instead a transversely polarised spin 1 boson with mass (i.e. not the photon) and its two transverse polarisations. But it is convenient to consider the more familiar example of spin ½. Either we should reject the Definite Range principle, or we should not assign observable status to components of spin, or we should deny that the square of an observable is always an observable. A similar argument would hold if we considered instead of S_x^2 S_y^2, and S_z^2 the magnitudes of the spin components: $|S_x|$, $|S_y|$, and $|S_z|$. So we would also have to deny that the magnitudes are observables.

We may now apply these three ways of dealing with the problem raised by spin to the case of area and volume. First, if we just deny the Definite Range principle the argument that there is a minimum area (and volume) collapses. Next, if we deny that quantities as physically important as the components of spin are observables then what right have we to assume area and volume are observables? Finally, if we deny that squares or absolute values of observables are observables, then I note that the area of a small parallelogram, or the volume of a small parallelepiped, may sensibly be interpreted as the scalar magnitudes (i.e. lengths) of the exterior products of two, or three respectively, of the vectors with length and direction of the sides, or edges respectively. Even if these exterior products (also known as antisymmetric tensors) are genuine observables then we have no reason to expect their magnitudes to be observables also. The situation may be illustrated by case of the area of a surface in three dimensions, in which case the exterior product may be replaced by the more familiar vector product. If we take a small parallelogram on the surface whose sides are the vectors **ds** and **ds′** then their vector product **ds×ds′** is a vector perpendicular to the surface with magnitude equal to the area of the parallelogram. But **ds×ds′** = - **ds′×ds**. That is, they are equal and opposite vectors. To calculate the surface area we do not integrate **ds×ds′**, which for any finite surface without an edge would give us zero, because the vector products in one direction are balanced by those in the opposite direction. Instead, we have to integrate the magnitude of **ds×ds′**.

The combination of continuous aether with Loop Quantum Gravity can be defended, therefore, even if the failure of the Definite Range principle is restricted to the magnitudes of signed or vector quantities that are genuine observables.

Loop Quantum Gravity offers, then, little support for discrete aether. Moreover, there is an argument to show that if treated as a fundamental theory of the aether it requires continuity. For it tells us that if we go round a loop, then the effect of the curvature is specified by a Lorentz transformation. So there must be something with the right structure for the Lorentz transformation to transform. If the cells were simples, however, then the corresponding transformation would be just a permutation of its vertices. If the permutation preserves orientation then it would be a member of the group A5, with 60 members. Maybe Loop Quantum Gravity could be modified so that this finite group replaces the Lorentz group. But that would not be Loop Quantum Gravity itself. As it is, Loop Quantum Gravity is more appropriate for a hypothesis in which the aether is made up of *cells,* each of which is flat and has the structure of a convex region of Minkowski Space-time. For in that case travelling around a loop would result in a Lorentz transformation.

This flat cell interpretation would not affect the entropy calculation for a black hole, which is an important ‘prediction’ of Loop Quantum Gravity (Ashtekar, Baez, Corichi and Krasnov 1998). For the flat internal structure of the cells would contribute zero entropy. Nor, it should be noted, does this settle the simplicity versus symmetry issue. For if we treat gravity as a field on a symmetric aether such as one with the structure of Minkowski Space-time then such a cell structure corresponds to a field restricted to the boundaries of the cells. I doubt, though, that we need draw the conclusion that there really is a cell structure for we may well think that quantum states with precise eigenvalues are idealisations and that actual states are always somewhat fuzzy being superpositions of such states. This would blur the sharpness of the boundaries.

Although there is no conclusive objection to combining Loop Quantum Gravity with a theory of the aether as consisting of cells each of which has the structure of a convex portion of Minkowski Space-time, that hypothesis is not especially simple. For like granulation it requires some Planck level structure, but it does not have the advantage of discrete theories such as Pseudo-set Granules. So in the absence of greater empirical support we should reject it.

6. Causal Set theory

Initially my preferred hypothesis was Point Discretion. In the previous chapter I noted two advantages that Granulated Aether might have over continuous aether in the case in which the aether is not symmetric, namely that of characterising a differentiable manifold. Whether Point Discretion shares these advantages depends on whether a certain hoped for result, which Sorkin calls a *Hauptvermutung*, can be proven. In addition there is the (temporal) Non-Locality Problem, arising from the non-locality of the discrete analog of Minkowski Space-time. But first I shall sketch Causal Set theory to provide some context.

Causal Set theory, developed by Peter Szekeres (1995), Rafael Sorkin (2003) and others, presupposes Point Discretion. It is based on points with a single basic relation – one of the (interderivable) relations of *absolute priority* ($u \prec v$), being the *immediate predecessor* ($u \prec^* v$), or their converses.[129] There is a derived *metathety* relation of y being *between* two points x and z that holds if either $x \prec y \prec z$ or $z \prec y \prec x$. If $u \prec v$ then u is said to be an *ancestor* of v, if $u \prec^* v$ then u is said to be a *parent* of v. The structure is just that of a partial ordering and one important problem was that of explaining how a partial ordering can come to approximate a differentiable manifold with a general relativistic metric. Sorkin's solution is to note that in the context of quantum theory we require a probability distribution over the many ways the aether (he would call it that) could be. And there is a very natural way of associating a manifold equipped with a general relativistic metric with such a probability distribution, namely *sprinkling* – the result of a random selection of at most countably many points from the manifold.[130]

---

129 In this context 'absolute priority' is appropriate in place of 'frame-independent priority', because we are positing a fundamental relation between regions.

130 If $u \prec v$, then the absolute temporal distance between u and v is defined as the length of the longest *ordered chain* connecting u to v, i.e. the greatest integer n for which there are points $u_0, \ldots, u_n$, such that $u = u_0$, $u_j \prec^* u_{j+1}$ and $u_n = v$. If the absolute temporal distance between u and v is large, then

I want to emphasise that Causal Set Theory is just the sort of account of the aether that we should hope for a priori. It combines my initial favourite account of the aether with a single primitive relation. The title 'Causal Set theory' suggests that the relation is causation. But that is not essential. Instead I call the relation that of *absolute priority.*

Causal Set theory does not posit any group theoretic symmetry for the aether. One problem, therefore, is that the Einsteinian (i.e. general relativistic) manifold, which I think of as Space-time, must be derived from the structure of the aether points. Sorkin hypothesises that the probability distribution over the states of a causal set made up of points is given by *sprinkling,* that is a random choice of representing points from the general relativistic manifold (Sorkin 2003: 9-10). The problem is that of inverting this procedure: given a range of physically possible causal sets, find an Einsteinian manifold such that they are the likely result of sprinklings. Now Sorkin suggests two constraints on the dynamics of causal sets (Sorkin 2003: 13). The first is 'discrete general covariance', which in effect says that the order in which points are born, that is come into existence, is not relevant to the dynamics. Sorkin's second constraint is 'Bell causality', which 'is meant to capture the intuition that a birth taking place in one region of the cause cannot be influenced by other births that occur in regions space-like to the first region' (Sorkin, 2003: 13).

These constraints lead to a formula for the probability of the birth of a new point with specified numbers ϖ of ancestors and m of parents (Sorkin, 2003: 13-14). This probability is proportional to $\lambda(\varpi, m) = \sum t_k \times (\varpi - m)!/((k - m)! \times (\varpi - k)!)$, where the constants t_k have yet to be specified, and the sum is over all k from m to ϖ.

A different approach to the dynamics of causal sets is to specify the *action* associated with a given subset X of the causal set and rely on the Feynman's sum over histories method of associating probability distributions with actions. Thus in 4 dimensions, the Benincasa-Dowker action associated with X is the quantity: $S(X) = N - N_1 + 9N_2 - 16N_3 + 8N_4$. Here N is the number of elements in X, N_1 is the number of chains

sum of all the points z such that $u \prec z \prec v$ should be a very good approximation to an Alexandrov interval with vertices u and v.

of length 2 in X, N_2 the number of chains of length 3 in X, and so on. (A chain of length 2 is a pair of points u and v such that u $\prec^*$ v, a chain of length n + 1 is n + 1 points u, $w_1, \ldots, w_n$ such that u $\prec^*$ w_1 and $w_1, \ldots$ w_n is a chain of length n.)

Given this, or some other formula, the Feynman method for assigning probabilities given an action, will then specify a probability distribution for a range of causal sets, one that will given some boundary conditions hopefully specify a unique Einsteinian manifold from which that distribution could be arrived at by sprinkling (Surya, 2011: 14).[131]

Thus we may hope to provide the Hauptvermutung (Sorkin, 2003: 10), showing that sprinkling over different Einsteinian manifolds is likely to give different causal sets. In that case we could say that the Einsteinian manifold specified is Space-time, in which the points are located.

Given this Hauptvermutung, causal set theory would specify a differentiable manifold thus replicating one of the major advantages of granulated aether. Moreover if the dynamics really does show that the Einsteinian manifold is specified by the boundary conditions then the Hole Problem is solved.

There is, however, a further problem, noted by Sumati Surya. We would expect the theory of gravity to be compatible with nearly empty Space-time equipped with almost the special relativistic metric. In that, and many other cases, each point is likely to have an infinity of 'parents', in which case the formulae fail to make sense. Or so it might seem, but before judging this I shall state the non-locality problem for Point Discretion starting with the straightforward incompatibility of temporal locality with discrete Lorentz invariance.

7. The non-locality of discrete Special Relativity

I begin by considering the non-actual and perhaps impossible case of a discrete analog of Minkowski Space-time. This raises a serious problem

---

131 Boundary conditions presumably specify the whole of history prior to the points of interest. Contrast this with the continuous case where we only need specify conditions on a suitable hypersurface.

in the case of the possible and perhaps actual case of an approximation.

To obtain a discrete analog of Minkowski Space-time we represent points by quadruples of integers. Likewise there is a discrete analog of spatial three dimensional Euclidean space with points represented by triples of integers. In the discrete Euclidean case the angle of a rotation is restricted to multiples of right angle, with the axes of rotation restricted to one of the three coordinate axes. This is a version of Weyl's problem. It is noteworthy, therefore, that in the case of the 4 dimensional discrete (analog of) Minkowski Space-time there is a countable infinity of transformations that keep fixed some past light cone u. These transformations form the group of all the Lorentz transformations whose matrices have integer values, and whose inverses also have integer values. Moreover, there are enough such Lorentz transformations to establish approximate spatial isotropy, where the approximation may be as accurate as we please.[132]

---

[132] The point p represented by a quadruple of integers <t, x, y, z> may instead be represented by the 2 × 2 complex valued matrix M(p) = [t +z, y + ix]&[y – ix, t – z]. (I shall use the convention that n by n matrix is to be specified by listing the rows in order joined by ampersands.) Given that representation, a Lorentz transformation may be assigned to any 2×2 complex matrix A of determinant 1, where if the Lorentz transformation maps p to q, M(q) = AM(p)A*. If A is [a, b]&[c, d] then A* is [a*, c*] &[b*,d*], where z* is the complex conjugate of z. Notice that A and –A correspond to the same Lorentz transformation.

We now restrict attention to Lorentz transformations for which the all of a, b, c and d are Gaussian integers (i.e. complex numbers of the form m + in, where m and n are integers) such that ad - bc = 1, and such that aa* + bb* + cc* + dd* is even. Then the corresponding Lorentz transformation maps quadruples of integers to quadruples of integers. The special case in which a, b c and d are all real numbers is noted by Schwarz (1976). A family of such transformations is obtained by taking any two Gaussian integers f and g and putting a =$2f^2$, b =(2fg +1), c = (2fg -1), d = $2g^2$.

Now the spatial directions of the vectors represented by the quadruples correspond to points on the sphere, which may themselves be represented as ratios of complex number. The effect on the spatial directions of the transformation represented by [a,b]&[b,c] is then a transformation of the sphere that sends the ratio u:v to (au + bv): cu + dv. The mathematical theory of

We have, therefore, an initially satisfactory discrete (analog of) Minkowski Space-time with points represented by quadruples of integers. But it suffers from non-locality. Consider a given reference-frame and a given point of discrete Minkowski Space-time. Then not only does it have infinitely many ‘parents’, that is immediate predecessors, but these occur arbitrarily far in the past and arbitrarily far away in Space with respect to the frame being considered. If therefore we think of direct causal influences being transferred from the immediate predecessors of a point to that point then the direct causal influences are transferred across arbitrarily large gaps of Space and Time. This is not the non-locality that coherent quantum states display, which implies statistical correlations across Space, nor is it restricted to very small times and distances. It involves causation operating across billions of years and light years.

The problem posed by such non-locality may be illustrated by considering point particles travelling at near the speed of light. Suppose m and n are large positive integers with no common factor other than 1. And suppose that m^2 units of length is about a light year. So if the unit is the Planck length, suppose m is of the order of magnitude 10^{24}. Consider a particle with trajectory represented by the set of quadruples $\{<k(m^2 + n^2), k(m^2 - n^2), 2kmn, 0>: k \in Z\}$, where Z is the set of integers, and the coordinates are for the observer's frame of reference. Although this satisfies the criterion that the trajectory seems macroscopically connected with respect to some frame, relative to the observer’s frame the particle only exists about once a year. Or, to take a more extreme case, a particle could have only existed twice since the Big Bang. Moreover these

---

these *Möbius* transformations is well known and sometimes discussed when considering how the night sky would look to an astronaut passing Earth at a high relative speed. By taking ff* and gg* large enough there is a spatial direction ζ, such that the transformation sends all directions other than those close to the opposite of ζ, to some direction very close to ζ. The direction ζ, is specified by the ratio f*: g*, and so is easily seen to be arbitrarily close to any specified direction. In the case in which the direction is close to the opposite of ζ, we may of course rotate by 180° first. This establishes approximate spatial isotropy in the sense that given any direction we may, to as good an approximation as we please, map it to any other direction using a Lorentz transformation with integer values for its matrix.

would be particles of familiar kinds such as photons or neutrinos. I call this the Problem of Occasional Occurrence. This, together with the more general non-locality is a serious problem for Point Discretion.

One conclusion we might draw is that Point Discretion coheres poorly with the hypothesis that the aether is symmetric. We might, instead, deny that there are point particles. But the problem is quite resilient. For, as I now argue, it holds provided the aether approximates discrete Minkowski. Space-time locally and it holds even if there are no point particles.

The discrete structure is at the Planck scale; the events of familiar subatomic particle interaction and so forth interacting take much longer, say N Planck times. Then to say that direct causation is non-local might be interpreted as saying the gap between cause and effect is of the same order of magnitude, N Planck times. One rationale for that interpretation would be that such a gap allows some later event to prevent the effect in question, an effect that has already been guaranteed by the cause. For example, a gamma ray might be about to generate an electron/positron pair when it is hit by some other electron – out of left field, as it were, knocking the gamma ray sideways. It is too late for the electron/positron pair not to come into existence – they have been directly caused – but the gamma ray is now in the wrong place.[133] This problem is very like the problem faced by proposed examples of backwards causation. If the cause guarantees the effect but comes after the effect, something might prevent the cause even after the effect has occurred.

For this reason, or merely because it is intuitive, I say that there is problematic non-locality if the direct cause precedes the effect by a gap of at least the time the cause (or the effect) takes, namely N Planck times. I now argue as follows. Suppose the cause takes N Planck times and then, one Planck time later, the effect begins. Now consider the same kind of process but occurring in a frame that is moving at near the speed of light relative to ours. Then the gap of 1 unit can be stretched to N units, although the N units will then be stretched to N^2 units. With a

---

[133] By making the interfering particle travel fast enough it could intervene even if the gap is rather less than N Planck times, but for simplicity I ignore this and consider a gap of N Planck times.

gap of N units we have the problematic non-locality. The only way this argument can fail is if there is no region of Space-time of the required extent for which the Space-time is approximately Minkowskian. A Planck time is of the order of magnitude of 10^{-44} sec. We may take the quark scale to be of the order of 10^{-18} cm, so events at that scale should be of order of magnitude about 10^{-30} second, making N of the order of magnitude 10^{14}. But even if N was as great as 10^{25}, N^2 Planck times is about 10^6 sec, which is less than a year and it is plausible that there are regions of about a light year across lasting a year in which the General Relativistic Space-time is a good approximation to Minkowski Space-time. So a process occurring in a frame of reference moving with respect to ours with velocity very near that of the speed of light will leave a gap long enough for a process taking the same proper time but in our frame of reference to interfere in what is being proposed as direct causation.

Causal set theorists such as Sorkin recognise the non-locality and the consequent need to introduce a cut-off. What this amounts to is the denial that the discrete aether is fully Lorentz invariant because frames of reference with velocities close to the speed of light are prohibited. Note that this cut-off is more severe than the plausible exclusion of frames of reference with respect to which the Big Bang is less than a Planck time ago. For it also excludes the frames with respect to which the Big Bang is many orders of magnitude times 10 billion years ago.

This problem is quite general for Point Discretion, but it is even worse for Causal Set theory with a dynamics that counts immediate predecessors, where we obtain infinite quantities. Maybe this special problem for Causal Set theory can be solved by requiring that no point has infinitely many predecessors, thus excluding as impossible the case of almost flat discrete aether that has always existed. In that case we may hope that a suitable formula for the action will provide us with the Hauptvermutung. Nonetheless the counter-intuitive non-locality remains. So some more drastic solution is required.

I now ask why Granulated Aether is not itself beset by the non-locality problem, and the answer is that it would be if it were strictly frame-invariant. For consider a granule represented by a pentatope of coordinate quadruples whose base is a spatial tetrahedron. The granule’s shape is purely topological in that it can as faithfully be represented by any pentatope, whatever the shape. Hence with respect to the ‘wrong’ frame of reference its immediate predecessor would be years ago. The solution to the problem in this case is to note that the fictitious Space-time in which the granules are located, and whose coordinates may be used to represent the aether, may be taken to be one in which no granule is located in pentatopes that depart too far from regularity. This fictitious Space-time is not Lorentz invariant and has imprecise simultaneity, implying an imprecise cut-off to the velocities of frames of reference. That is, there might be many permitted frames of reference, namely those with respect to which the granules are not too far from regular pentatopes. None of the frames are ones with respect to which some immediate predecessor of a granule is a long time ago.

A necessary and precise absolute simultaneity is considered problematic because it makes it mysterious why the laws of nature are such that in the idealised case of Special Relativity they are independent of the choice of frame of reference. Neo-Lorentzians such as Craig (2001) accept this mystery as the price they think they must pay for dynamic theories of Time. Assuming they are wrong (Forrest 2008) we must avoid this mystery. To do so it suffices that there is a range of permissible frames of reference. If we fix one of them, F_0, then the range will be specified by a set S of Lorentz transformations as the set of frames mapped from F_0 by some $g \in S$. Provided S generates all Lorentz transformations (i.e. provided there is no proper subgroup of the Lorentz transformations containing S) any law of nature that has the same statement in every permissible frame must have the same statement in every frame, as required for Special Relativity.

It should be noted that because the Lorentz group is not commutative there is no principled way of picking out the centre of the range of permissible frames of reference. If there were it could be argued that the central frame was in fact that of absolute simultaneity, re-introducing the problem that I claim to have solved.

This solution to the non-locality problem arises in a natural way in the case of Granulated Aether, but would be the result of an arbitrary cut-off to frames of reference in the case of Point Discretion. This difference is the reason why I prefer Granulated Aether (and, in particular, Pseudo-set Granules) to the initially more attractive Point Discretion. I note that Causal Set theory may easily be re-interpreted as a theory of granulated aether. We just replace points by extended simples and take the immediate ‘causal’ priority of u over v, $u \prec^{*} v$, as one of three ways in which u and v can share a hyperface. A happy consequence of this re-interpretation of Causal Set Theory is that the number of immediate predecessors may be taken to be finite, as required, for instance, in the Benincasa-Dowker action.

Alternatively, we might resort to Dynamic Triangulation theory, which seeks to develop quantum gravity by using the ideas of classical statistical mechanics but applying them to the division of the aether into cells, which I take to be granules (Ambjørn, Carfora and Marzuoli, 1997). As I understand it, Dynamical Triangulation was initially stated in terms of the length of edges, but it is hoped that a purely topological, that is, qualitative, theory, can be developed. (Ambjørn, Carfora and Marzuoli, 1997: 297-8). The resulting theory will be similar to Causal Set theory.

8. From discrete aether to geometrodynamics

Geometrodynamics is the ambitious project of using quantum gravity to provide a ‘theory of everything’. In the continuous case it would require that fundamental particles (quarks and leptons) be suitable knots or other topological features in the aether. In the discrete case these particles could be patterns analysable in terms of adjacency, including its special case of immediate causal priority. I would like discrete geometrodynamics to be correct, and it would be a triumph for ontological simplicity, especially on the Pseudo-set Granules version.. But I judge it to be bold in the Popperian sense, and so unlikely prior to testing.[134] Readers may

---

[134] The noteworthy feature of a Popperian bold conjecture is that prior to testing it is judged improbable that the test will be in conformity with the conjecture. From that it follows both that the conjecture is improbable prior to testing and that if tested and not found wanted it will be significantly confirmed. I noted previously that bold conjectures have been confirmed more often than

disagree: in the words of an excellent metaphysician, 'I would think the less of reality if it were not so'.[135] So I do not rely on my judgement in this case. What I shall argue, though, is that discrete theories of the aether are committed to geometrodynamics.

To argue this, consider again the topic of quantum gravity. Some readers may resist the need to quantise gravity.[136] But in that case we have no reason to believe in discrete aether since, prior to quantisation, General Relativity and its variants are all continuous theories. So the thesis that aether is discrete is committed to the quantum gravity project. Next suppose we undertake this project by treating gravity as somewhat like the other forces and so to be quantised in the same way, using Feynman's sum over histories method or something similar. Initially that would seem to require the aether to be flat or to have some other symmetric shape. For the interpretation of General Relativity as showing that the aether is itself curved in an irregular way, seems to imply that gravity is quite different in origin from other forces. But the initial inference overlooks the possibility of treating all the forces as due to the curvature of the aether, as in the geometrodynamic program. There are, then, two currently feasible methods of treating gravity as like other forces: assimilate gravity to them or them to gravity. It is hard to assign a probability to the third way, namely something quite novel – neither curvature nor forces – that explains both of them. Only if neither of the currently feasible methods succeeds we would seek the novel third way. Given that, as I shall argue, it is far from bold to treat gravity as much like another force, I judge it fairly unlikely that some as yet un-thought-of third way will be required.

---

we would have expected a priori, and so we should be somewhat less reluctant to affirm them.

135 Said only half in jest by Keith Campbell in conversation.

136 Suppose, for instance, we interpret quantum theory in terms of many worlds in each of which there are particles interacting with each other. Then General Relativity without any quantisation may be interpreted as showing that the worlds are curved by the particles and that between interactions the particles follow geodesics. The case for quantising gravity has, however, been ably made by Thomas Thiemann (2003: 6-7).

We are left, then, with geometrodynamics on the one hand and gravity-as-a-force on the other. It remains to sketch gravity-as-a force, and to explain why it is unlikely given discrete theory. The sketch will enable me to comment on the differences between gravity and other forces and so reply to the objection that these differences are best understood by taking Space-time to be contingently curved.

Gravity is not *exactly* like other forces. The first, but perhaps least important difference is the curious combination of (1) the way gravity is characterised by the general relativistic 'metric', but (2) there is still a metric in the case in which we ignore gravity and Space-time is Minkowskian. We might perhaps take this to show that the gravitational field is the deviation from the 'metric' of Minkowski Space-time. But that would be to introduce a necessary structure consisting of the light cones or some equivalent way of characterising Minkowski Space-time, even though it is redundant because there is a contingent gravitational field that replaces the Minkowski structure in all but the matter-free case.

Instead I submit that the matter-free case is not gravity-free, and that the so-called absence of gravity is just the case of an especially straightforward 'metric', that of Minkowski Space-time. Additional support for this simplification may be derived from the probable need to introduce a cosmological constant, which also shows that the matter-free case is not free of all gravity-like features..

It is initially probable, therefore, that the necessary structure of the aether does not include the light cones or, equivalently, the partial ordering of absolute priority. This coheres with the case, made in the previous chapter for decoupling the metric and order aspects of Space-time, because neither the aether nor Space-time is ordered by anything necessary but rather by the contingent gravitational field. In the next chapter I shall assume that affine aether has no necessary light cone structure, even if necessarily it has some light cone structure. The alternative, in which light cones are necessary presents no difficulties and, in the point-free case, is less complicated, which should occasion some doubts as to whether the cone-free approach is correct. It will also turn out that in the case in which the structure of the aether is not affine but that of de Sitter Space-time the cosmological constant implies a light cone structure.

The most straightforward case, however, is that in which affine geometry holds of the aether. This is the geometry in which there are hyperplanes and an equivalence relation on hyperplanes of being *either identical or non-identical but parallel*, where parallel implies non-identical. This equivalence relation is such that such the analog of that Euclid's Fifth Postulate holds, namely parallel hyperplanes never intersect. Affine geometry may be characterised axiomatically using the primitive relation of *metathety*, namely the three-place relation between points p, q, and r that holds just in case p, q, and r lie on some line with q between p and r. (Coppel 1998). A less fundamental description is as the geometry in which Space-time is represented by quadruples of real numbers in such a way that hyperplanes correspond to linear equations, but without any metric structure or specification of perpendiculars. So, for instance, although hyper-ellipsoids may be characterised there are no special hyper-ellipsoids designated as hyperspheres.

The second difference between gravity and other forces may be illustrated by considering the way a field acts on a single particle – a special case that generalises to many particles and to the way the field acts on itself. The dynamics is constrained by the *action*, which, taking Planck's constant to be 2π, is $\int p_q(dq)$, where: (1) the integral is over the path of the particle, and (2) p_q is a real valued linear function on the vector space T of translations of Space-time (displacement vectors) [137]. The linear real valued functions on T form a vector space, T*, the dual of T. The vector p_q itself depends on the position q. and because we put Planck's constant equal to 2π, may be identified with the energy-momentum vector at q.[138] The energy-momentum is a property of the

[137] This generalises: (1) we usually consider either a field or matter that is spread out, in both cases the energy-momentum vector is replaced by a tensor; and (2) in the curved case T is the tangent space at q.

[138] Strictly speaking, the energy-momentum should be thought of as belonging to an affine not a vector space. The difference is that a vector space is an affine space with a distinguished member, the zero vector. In the case we are considering the choice of a zero vector introduces gauge invariance, which shows we should not treat p_q as a member of a vector space. The gauge in-

particle and so not defined for all locations but merely those on the path, and it will vary as the particle is affected by the field.

The field affects the motion of the particle by specifying how the energy-momentum varies along the path. Such a path-dependent correlation is an *affine connection* in the differential geometers' sense. For every point q of Space-time and every vector u in T the connection specifies a way of correlating energy-momenta at q with energy-momenta at a point q + u, displaced by u from q, for a particle travelling in a straight line. So the connection that characterises a field is a mapping assigning to each position q and each vector u in T a mapping F that correlates the energy-momenta at q with those at q + u. The mapping F preserves the affine structure and therefore given an arbitrary choice of zero energy-momentum at each point it may be analysed as the composite of two mappings, a translation and a linear transformation that leaves the zero vector fixed. We may then treat the translation as the non-gravitational aspect of the field and the linear transformation as the gravitational field itself. [139]

In a unified field theory we should expect there to be a single field but it could still have these two components, so in that sense gravity would be distinct from the rest of the field.[140]

---

variance in question is shown by considering a smooth function r assigning a member of T* to every location and whose exterior derivative is zero. If we replace p_q by p_q + r(q) that has no effect on the difference between action integrals joining two given points and hence no effect on the dynamics.

139 I understand that Einstein rejected this way of unifying fields. (He was considering only the electromagnetic and gravitational field.) His rejection seems to have been because he tried to unify the fields by generalising General Relativity, relying on the curvature to specify the evolution of the unified field.

140 In a unified theory theenergy- momentum space could be a larger space and perhaps not even affine. (The unified theory is even more unified if the symmetry group of the momentum space is a 'simple' group, that is, one without a normal subgroup.) In any case, we generalise from energy-momentum vectors to energy-momentum tensors. Then the *translation* adds to the angular momentum in a way that is said to be gravity acting on the in-

There is one additional feature of the dynamics that might or might not be fundamental, but is worth mentioning, to counter readers' worries that I have not interpreted General Relativity *as they know it*. Energy-momentum space is not merely affine, it has, as far as we know, a necessary Minkowski-space structure, with a constant μ, characteristic of the kind of particle, which is usually taken to be positive and to be the square of the rest mass. If the energy-momentum is represented by the quadruple $<e, p_1, p_2, p_3>$ then $\mu = e^2 - p_1^2 - p_2^2 - p_3^2$. Because the connection preserves not merely the affine structure of the momenta but μ is preserved, the connection assigns a member of the Poincare´ group (a combination of a translation and a Lorentz transformation) to each point q and vector u. [141] Strictly speaking it is a Minkowski connection, therefore.

This brings me to the apparent superiority of the curved Space-time interpretation of General Relativity, namely the way in which curvature enters into the equations. My response is that curvature does not need to be understood in terms of a 'metric', which is a natural way of thinking only if we are already thinking of a 'metric' in the flat case, as with Minkowski Space. Instead, therefore, I note that the idea of curvature arises whenever we have an affine connection, and a 'metric', if there is one, defines the curvature by defining an affine connection. The translation component of the connection does not affect the curvature, hence the link between curvature and gravity.[142]

I anticipate the objection that the connection makes the affine or other symmetric structure redundant and a differential manifold would do instead. My response is that the affine or other symmetric structure of the aether enables us to preserve Newton's First Law, with friendly

---

trinsic angular momentum, the spin. This appears to be the motivation for modification of General Relativity by Élie Cartan (1922.)

141 Because the standard case is one in which the underlying Space-time is curved it would be usual to describe the connection as an assignment of a member of the Lie algebra of the Poincaré group to q and a tangent vector.

142 The translation component is called the *torsion* because of the application to defects in crystals.

amendments. When the gravitational influence of other matter is negligible and when matter and the non-gravitational forces are also negligible the trajectory of a particle differs negligibly from a straight line (or analog).

Without in any way minimising the mathematical work required, I judge, therefore, that there is only one potential obstacle to a unification of forces on the assumption of a background symmetric Space-time, such as affine space. This is the notorious problem of the divergences. When we use the Feynman sum-over-histories method of quantising a system we find that all the processes we expect to occur have exactly similar duplicates in frames travelling with high velocities relative to us. Unfortunately the appropriate 'probability' measure on the set of frames of reference is the unique (up to a multiplicative constant) Lorentz invariant one, and this is not a probability measure, with total value 1, but has infinite total value. This is, I claim, the ultimate source of the divergences, which may be avoided by means of range of permissible frames of reference, as used in the case of Granulated Aether. That amounts to both an ultraviolet cut-off (excluding frames moving fast towards us) and an infrared cut-off (excluding frames moving fast away). We need to avoid ad hoc choices of the range of privileged frames, but fortunately there is an obvious source of cut-offs, the postulate that the some fundamental fermions of non-zero rest mass have extended locations. For the sake of being definite, suppose an electron has some (non-infinitesimal) positive diameter when it is in a rest frame. Then if it is travelling very close to the speed of light with respect to us its diameter at right angles to its velocity will too large to avoid other electrons that are travelling at low speeds. The postulate of extended electrons puts two constraints on the dynamics, the first is to put an upper limit on the number of them per unit volume and the second is to put an upper limit on their relative velocities. This supposes that no two electrons intersect, but that is plausible since they are fermions.

This way of avoiding the divergences is not ad hoc because there are two reasons why point particles are improbable. The first, which is so elementary it makes me blush, is that point particles moving in straight lines have an infinitesimal probability of collision, and so the attractive idea that particles interact by transferring energy-momentum (and spin) on contact requires them to be extended. The second is that I am taking the aether to be the one fundamental substance and hence particles are just regions of aether with suitable properties. Combining this with the superiority of point-free Arntzenius Continuum over an aether composed of points I conclude that there are no point particles.

It only remains to show that symmetric Space-time is improbable given granulated aether. That follows from the way the aether is more fundamental than Space-time. It is not as if we start with some intuition that Space-time is flat and then cut it up into pentatopic domains that we declare correspond to the granules. Instead we start with what is fundamental, the granules and then ask why they should fit together in just such away that the Space-time they are located in is flat, or otherwise symmetric. If we have no answer to that question then we should suppose it is improbable, and that the arrangement of granules is higgledy-piggledy but such that a fictitious Space-time in which they are located is flat enough at anthropocentric scales (neither too small nor too large).

This completes my case for the disjunction: either granulated aether without symmetric Space-time but with geometrodynamics, or symmetric Arntzenius Continuum, with gravity being treated much like the other forces. I shall conclude this chapter by considering some further arguments for discrete aether. Regrettably none of them succeed.

9. The case for discrete time

Why does the aether persist? Perhaps the most intuitive answer is that every part of the aether *endures* forever, that is it lacks any division into temporal parts. But I have two reasons for rejecting that thesis. The first is that it implies an aether current, something that Michelson and Morley looked for, and which we might expect to find some empirical evidence for if it occurred, but we do not. The second is Lewis' problem of temporary intrinsics (1986: 202-5): the very same portion of the aether would have to have different intrinsic properties at different times, contrary to the Indiscernibility of Identicals.[143]

---

[143] For a survey of solutions to the problem of temporary intrinsics see (Gallois 2011). One is to say that persisting entities have time-dependent properties, another that they time-dependently instantiate properties. If either of these

We require a different hypothesis to explain the universe's persistence. The most straightforward is that each stage is *directly* caused by some of the earlier stages (Hans Reichenbach's *genidentity*, 1957: 270-271). Here I say that u directly causes w if that causal relation does not depend on u's causing v that causes w, for some v.[144] If persistence is due to direct causation, then causation, at least as it applies to the persistence of the aether, is discrete, in the sense that indirect causation always depends on finite causal chains. To be sure, discrete causation does not entail a discrete theory of the aether, but I now argue that the combination of discrete causation with a continuous theory of the aether is awkward.

Suppose the aether is temporally discrete but spatially continuous. Then there cannot be any continuous processes, because the structure would have to change suddenly at the end of a region. So I infer that if the aether is spatially continuous it is also temporally continuous. Combining that with discrete causation requires that a layer of aether divisible into thinner layers be caused as a whole by an earlier layer similarly divisible, with divisions between the layers.

Might not we then have 1 unit thickness followed by ½ followed by ¼ etc Zeno-style, resulting in an explanation of persistence for only 2 units of time in all? To exclude this, we have to posit that layers of some fixed thickness, say the Planck time, directly cause later layers of fixed thickness. The objection to this is that unless Time is discrete the uniform thickness of the layers is an ad hoc hypothesis to save discrete causation, whereas discrete causation is implied by a discrete theory of the aether.

---

holds then there must be a Time distinct from the aether that fills Space-time. Time thus distinguished from Space-time is considered below, when discrete Time is discussed. Another solution is the presentist insistence that only the present is real. I fail to see how the persistence of the aether or any other physical thing can be explained by presentists without invoking either an enduring or timeless sustaining cause, something I also consider below.

144 In case of direct causation there *can* be an intermediate cause v. What is excluded is that u's causing w *depends* on the intermediate causation.

That was the argument and I now consider alternative explanations of persistence of the aether. Some would say, ‘It just does’. I am not impressed. For there are answers to the question and so not to answer it is to multiply mysteries.

Do laws of nature provide an explanation of the persistence of the aether? I have not, and shall not, assume that causation is primitive, so direct causation might well be itself explained by laws of nature – conservation laws maybe. But to provide an alternative we require laws that explain why the aether persists even if not all causation depends on direct causation, or perhaps even if there is no causation at all. First suppose the law states that if the aether exists at time t it exists at some later time. In that case, we have the Zeno-style problem that the aether persists for 1 second then ½ second then ¼ second etc and we have not explained how it persists for 2 seconds. Next suppose the law states if the aether exists at time t then it exists at time t + ε. Then rather than being an alternative to discrete causation it implies discrete causation. For if we consider the slice of aether that has existed for the last ε it directly causes the slice existing for the next ε. So it is not that easy to state how the law works. My suggestion is that it would have two parts. The first, a *metric completion* law, stating that if there is a sequence of time coordinates, t_1, t_2, etc at which the aether exists and this sequence converges to a limit, t*, then the aether exists at t*. The other would be the law that if the aether exists at time t it exists at *some* later time. Together these laws imply that the aether exists forever.

I have two objections to these or to any other laws that imply the continued existence of the aether. The first is that they work too well; implying that necessarily if the aether ever exists then it exists forever. I incline towards the position that the aether always has and always will exist but this does not seem to be necessary. For if the universe were expanding somewhat less rapidly it would collapse into a final crunch. My other objection is audience-specific and directed at those who assume General Relativity in a version that requires all laws to be covariant. It is that unless Time is discrete, the time coordinate may be re-scaled so that the infinite future becomes finite. So the law in question would not satisfy the general relativistic requirement of covariance. By contrast, the explanation of persistence in terms of discrete causation is covariant. If we re-scale Time we merely change the shape and size of the layers or cells of aether without affecting the causal description.

An alternative explanation that should be taken seriously is that there is an enduring entity, such as the God of classical theism, that sustains the aether. So the aether is not caused by earlier stages. Or, the divine persistence might be due to each divine stage directly causing the next, but this persisting God could then sustain the aether.

My challenge, then, is directed at those, whether theist or atheist, who reject the idea that the aether is sustained by God, to explain why it persists without relying upon discrete aether.

That was the challenge. I now take it up by distinguishing between the Space-time in which the aether is located and the temporal order. I do so in the context of the Growing Block theory, with the block in question being the aether. That is, I suppose the grammatical present is pre-tensed and if we want to restrict attention to what exists now we have to be explicit and use the word 'now' or something equivalent. So, as a Growing Block theorist I say that the aether that exists at some time in the past is a proper part of the aether that now exists. I shall also assume that Space-time is symmetric, because that is what I believe holds unless the aether is discrete, and I am here considering the case from discrete Time to discrete aether, so I can argue by reductio ad absurdum, supposing the aether is not discrete.

On this Growing Block theory the truth is what is now true, which is not the same as what was true a while ago. A. For some *a* and *b,* '*a* is the aether' is now true but '*b* is the aether' was true and *b* is a proper part of *a.* Space-time does not change: truths about it are true at all times. So the aether grows into Space-time. (I explain in the next chapter how the whole of Space-time can be constructed from the aether that does not fill it.) Space-time has a partial ordering of its points with respect to frame-independent priority. So it is not just a hyperspace, but on the Growing Block there is an additional temporal ordering of the past (and present) locations of the aether, with the locations of earlier ones being included in those of later ones. The passage of Time concerns the growth of the aether with respect to that temporal ordering. Rather than call this temporal ordering *hyper-time* it would be more appropriate, although pedantic, to re-name Space-time as Space-*hypotime*. In any case we may distinguish the metric aspects of Time, which are thoroughly unified with those of Space, from the ordinal aspects.[145] This temporal ordering I take to be a total not merely a partial ordering. This is not the place to defend the resulting theory of Time. Rather I have illustrated how we might distinguish the two aspects of Time. Given that we do, so we may then solve the problem of why the universe persists without invoking God as sustaining cause, by taking the temporal ordering to be discrete in the sense that for every x and y there are only finitely many z after x and before y. In that case the rate of passage of Time may be measured in seconds, not, note, seconds per second: with the rate being the thickness of the layer between two successive stages of the growth of the aether, which is analogous to tree rings. There is no reason to expect this growth either to be spatially or temporally uniform.

I conclude that unless the aether is sustained in existence by God or some other sustaining cause, Time is discrete, but by distinguishing Time from the hypotime that is the fourth dimension of Space-time, I resist the conclusion that the aether is itself discrete.

Supertask arguments, even if otherwise acceptable, likewise show that it is the temporal ordering that is discrete, rather than directly showing that the aether is. These arguments proceed by showing that if Time is continuous then there is an infinite sequence of events such that any finite number are jointly possible but the whole series is so counter-intuitive that we might judge it impossible. Then we have a contradiction, because if the whole series is impossible it must cease to be possible before it is completed and hence some finite number of the events are not jointly possible. There are many ingenious examples but the following variant on the Ross-Littlewood Paradox suffices (van Bendegem, 1994). At time 0, some kind of particle, a magnetic monopole say, is unique. Then some events result first in the creation of two new monopoles and then the destruction of the old. (By ‘then’ we mean at a later time with respect to some arbitrarily chosen frame of reference.) A short

---

[145] More generally the ordering is that some enduring and everlasting states of affairs come into existence before others. I am not claiming that all states of affairs are everlasting only the fundamental ones on which others depend.

time later three new magnetic monopoles are created and then the two already existing are destroyed. By chance this is repeated more and more often so that over a second interval there are more and more, but at the end of the process there are none, because every monopole that came into existence has ceased to exist. The whole process is of infinitesimal probability but seems nonetheless to be possible unless Time is discrete. My response is, 'So what?' Supertask enthusiasts might, however, consider it impossible and hence they have a reductio ad absurdum of the supposition that Time is continuous.

My conclusion, then, is that arguments for the discrete character of Time may well succeed, but if they do the advocate of continuous aether should take this as showing that Time is not the same as the fourth dimension of the Space-time in which continuous aether is located. To naturalists, the idea of Time distinct from Space-time might seem a metaphysical extravagance, but my guess is that the naturalists in question will not be impressed by the argument for discrete Time in the first place.

7. Symmetric Space-time

I have argued that unless the aether is discrete we should adopt a revisionary interpretation of General Relativity, according to which Space-time is not lumpy and bumpy like a potato, but is affine or de Sitterish, or of some other nice symmetric shape.

Is the aether itself symmetric? From its symmetry it would follow that the it pervades the whole of Space-time. On the Arntzenius Continuum hypothesis, they would differ only in that Space-time is made up of point locations, constructed from regions in the way described in Chapter Four. This would exclude both the Growing Block and Presentism, if these are interpreted as denying the reality of any future aether. For if there is no future aether then the aether is not invariant under those symmetries that map the past to the future.

There is another complication with the use of symmetries to characterize the aether. Realists about universals have at their disposal a group-theoretic approach based on realism about symmetries, which are dyadic relations between regions of aether. Nominalists might, however, reject symmetries, in which case they are committed to an alternative way of characterizing symmetric Space-time, using axioms. This is fairly straightforward for affine and Minkowski Space-time, but for other symmetric spaces such as de Sitter Space-time I do not know of axiomatic treatments.

One final issue concerns the characterization of Space-time if the aether is point-free. The axiomatic approach requires quite a serious complication; the group-theoretic one generalizes straightforwardly.

I shall begin by considering the axiomatic approach. I then suppose the aether is point-based and consider the group-theoretic approach for flat Space-time, before generalizing it to the point-based curved case, and the point-free cases.

1. The axiomatic approach

The point-based case
There are intuitive axioms (Coppel 1998) that ensure a point-based continuous aether is representable as flat. The primitive relation used is that of *metathety,* namely the relation that holds between points p, q and r in that order if they lie on a straight line and q is between p and r.

The requirement for a mapping F from the points to the quadruples to be a (faithful) *representation* is that:

> F is a 1 to 1 mapping and; if q is between p and r, then F(q) is a convex linear combination of F(p) and F(r), that is, for some real number x, $0 \leq x \leq 1$, and $F(q) = xF(p) + (1 - x)F(r)$.

Coppel's axioms are sufficient to ensure that the aether is representable as flat. Moreover, if the convex region in question is not the set of *all* quadruples of real numbers, then we may take Space-time point locations to correspond to all the quadruples, exhibiting the aether as occupying a convex portion of Space-time, which would be sufficient to accommodate the Growing Block or Presentism. (Here I am supposing the presentist believes in a present layer of the aether of, say, a Planck time thickness.)

Being representable as flat is not the same as being flat, however. So this approach, although basically sound, must be modified to include hypervolume (or, more austerely, the comparative quantity relation) as primitive in addition to metathety. Moreover, even if the aether is flat the representation could be misleading, because of fine structure in the form of extra dimensions that do not contribute to the quantity. Compare aether of three spatial and one temporal dimension, with a merely possible aether in which only the temporal and two of the three spatial dimensions contribute to the quantity. Ignoring Time for the sake of simplicity this would be a two dimensional aether except that it has an extra infinitesimal thickness in the third dimension. In both the actual and the merely possible cases, the axioms concerning metathety would be satisfied, and the number of dimensions characterized as three. But I consider that the difference between the two cases should be reflected in the axioms, and this can be done, in 3 dimensions, as follows. In the actual case, but not the other one, if we take any 4 points none of which is part of every convex region containing the other 3, then any convex region containing all 4 of them must have positive quantity.

A more serious misrepresentation of the aether would occur if in fact it is curved but *conformal* to flat. Consider Einsteinian manifolds, that is, 4-dimensional manifolds with a general relativistic ‘metric’. David Malament's result (1977) shows that their structure is specified by the light cones and the hypervolume, but it is not specified by the light cones alone. In particular, there can be an Einsteinian manifold E that is not flat and yet there is a 1 to 1 onto mapping H from E to a Minkowski Space-time, M, such that H and its inverse are smooth (infinitely differentiable) and that H maps light cones to light cones and geodesics to geodesics. The mapping H also preserves metathety, if we define it replacing ‘straight line’ by ‘geodesic’. Hence E satisfies the axioms for a flat Space-time. To avoid treating E as flat we need to distinguish E from M using the hypervolume.

The easiest way of introducing the hypervolume to solve this problem is requiring it to be translation-invariant, which is an immediate consequence of the symmetry-theoretic approach. This is not so straightforward using axioms, however. We may characterize parallelepipeds and then consider one of them, u, divided into two of equal hypervolume, v and w by a hyperplane parallel to one of u’s faces, and divided into another two of equal hypervolume, x and y by a hyperplane parallel to another of u’s faces. The requirement that the four parallelepipeds $v \wedge x$, $v \wedge y$, $w \wedge x$, and $w \wedge y$ are always of equal hypervolume should, I think, ensure the translation invariance of hypervolume.(See Diagram Two for the two dimensional analog.)

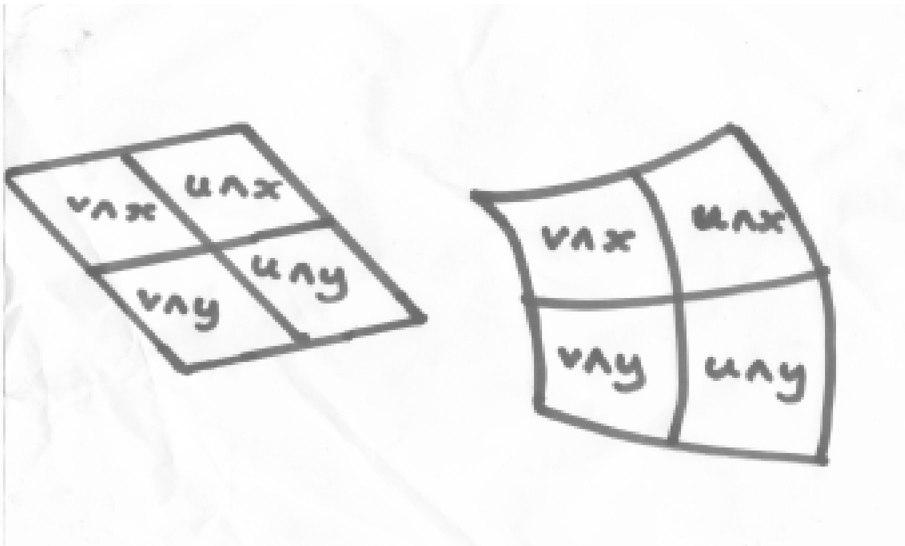

Diagram Two

I conclude that the axiomatic characterization of affine, or, more generally, flat aether succeeds in the point-based case, but at the cost of some complexity. In Chapter Five I explained why if we treat gravity as just another field we can, in the case of a flat Space-time, choose between a Minkowski structure or the, more general, affine structure.[146] Economy supports the latter, but I note that, as Alfred Robb (1914) showed, we can find axioms that characterize Minkowski Space-time.[147]

---

146 Bearing in mind that a flat Space-time is affine only if it extends infinitely in all dimensions. For ease of exposition I am here ignoring spatially finite flat Space-times.

147 Robb was the pioneer. His treatment is not merely rather complicated but it characterises as Minkowskian, Space-times that are merely conformably equivalent to Minkowskian. This is the problem that I considered above when discussing affine aether. In addition Robb assumes there are points.

I recommend an axiomatic system that takes the part/whole relation, frame-independent temporal priority, and comparative quantity as basic. Using these, we may characterise a metathety relation between past light cones that corresponds in the point-based case to metathety of the vertices of the cones. Then we use the axioms for affine spaces stated in terms of the metathety and hypervolume, mentioned above. Finally we require some constraint on the past light cones to ensure they have the right shape to be cones. This last step is quite easy, because we can characterise a two dimensional slice of

The point-free case

Even if the aether itself is point-free, I take Space-time to be made of points (point locations). A region of Space-time *w* is said to be *convex* if whenever two points p and r belong to w any point q between p and r also belongs to w. This idea of convexity is, I suggest, a natural concept and in the point-free case we may take it as a primitive property that some regions of the aether have and others lack. We may use it to provide an axiomatic characterization of point-free affine aether, without any mention of the constructed points.

We may then say that an aether region u has *flat boundary* if both u and some complement of u are convex. In the case of greatest interest, Arntzenius Continuum, every region has a unique complement so we may say a region has flat boundary if both it and its complement are convex. I shall restrict attention to that case.

The unordered pair {u,v} consisting of two convex regions each the complement of the other is a surrogate for a hyperplane, so in this case I denote {u,v} as H(u, v). We expect both u and v to be representable by open sets of quadruples separated by a hyperplanar set of quadruples.

Given two distinct hyperplane-surrogates H(x,y) and H(u,v) it might happen that one of u, v, x or y has another of u, v, x or y as a proper part. In the Arntzenius Continuum case we expect this to occur only if the two representing hyperplanar sets fail to intersect. We may therefore use this to characterize the relation of being parallel, where I stipulate that being parallel is anti-reflexive. It will then be an axiom that the relation of *being either identical or parallel* is transitive. From its definition it is symmetric so it is an equivalence relation. An axiomatic

---

a cone in N dimensions using the converse of Pascal's theorem for conic sections.

The first step can be broken down into two stages. First we may characterise metathety for 'points' (i.e. vertices of cones) that are neither space-like nor time-like separated but lie on a light ray. For a past light cone and a future light cone that touch but have no overlap of positive volume must define such a line. Then we use Pappus' Theorem to characterise other cases of metathety.

characterization of n dimensional affine aether, for n ≥ 3, may then be obtained by:

1. Requiring the equivalence classes of hyperplane-surrogates to satisfy the axioms for the (dual of) a projective space of n - 1 dimensions over the real numbers; and
2. Putting constraints on the hypervolume similar to those for the point-based case.

If we construct the points of Space-time using the ultrafilter construction we will thus obtain an affine structure for the Space-time in which the regions are located. This requires, however, the aether to occupy all of Space-time, whereas in the point-based case it sufficed that the aether occupied only a convex region. (We need all Space-time to ensure that the relation that we treat as being parallel is a genuine equivalence relation.) So this is not compatible with the strict presentist or Growing Block theories that deny the existence of future fundamental things such as future portions of the aether. To get around this restriction we could consider the one-cornered regions that are meets of 4 suitably arranged regions with flat boundaries. Such regions act as surrogate points, and we may characterize when three of them are lined up in such a way that the corner of one of them is between the other two corners and on a straight line. But that is a messy way of characterizing affine aether.

2. The group-theoretic characterisation of affine aether

Inspired by Felix Klein's Erlanger program, in which different geometries are characterised algebraically by their structure-preserving mappings, I shall show how in some cases the geometry of the aether may be characterised using a group of symmetries. I begin with the most straightforward cases, those in which the aether is point based and pervades the whole of Space-time, with which it may be identified. In this case the symmetries are global. Later I shall consider local symmetries.

A symmetry g is here taken to be a universal. As mentioned in the Introduction this supposition requires quasi-realism about universals, namely either realism or a nominalist theory that can mimic realism. If I want to stress that a symmetry g is a relation I write it as R_g. Because g maps regions of aether in a one to one onto fashion, if R_guv and R_gwx, then u = w if and only if v = x. As a symmetry, g preserves the fundamental properties of and relations between points. In the affine case I shall characterise the aether using four primitives: the part/whole relation of mereology, the comparative quantity relation, the property of convexity, and the idea of a symmetry itself. So I require that the symmetry preserve symmetries. This follows from the assumption that for any symmetry g, and any region x there are regions y and z such that $R_g xy$ and $R_g zx$. In that case, given the convention that identity is a symmetry, the symmetries form a *group.* That is, any two symmetries have a composite that is also a symmetry, composition is associative and any symmetry g has an inverse, g^{-1}, which when composed with g results in the identity map. The symmetry g then preserves symmetry in the sense that if h is any symmetry $g^{-1} \circ h \circ g$ is also a symmetry. The topology is derived from convexity and hypervolume so the relation of separation is also preserved.[148]

The inclusion of the concept of symmetry among the primitives is required, I hold, if we are to use symmetry to describe the nature of the aether. This is a problem because symmetry *supervenes* upon the other primitives, and so it would seem more economic to take it to *depend ontologically* on the other primitives. This is a known problem for those who take both universals and particulars to be fundamental. For instance, the existence of the universal *being spherical* supervenes on the spherical objects but neither the objects nor the universal are thought of as less fundamental than the others. One sort of nominalist seeks to solve the problem by taking the universal to depend for its existence on the particulars and their resemblance. On the other hand, a bundle theorist solves the problem by treating the particulars as just bundles of universals. My own preferred solution is platonist – the universals are necessary beings and so supervene on the particulars only in a trivial sense that does not

---

148 We may characterise the (not necessarily open) regions with flat boundary as those convex regions with convex complement. Two convex regions u and v are separated if they are parts of disjoint regions with flat boundary w and x such that the meet of the complements, $\neg w \wedge \neg x$ is of positive hypervolume. Then two regions y and z are separated if every convex part of y is separated from every convex part of z.

threaten their being taken as primitives. In any case, whatever solution we adopt generally for universals applies to the case of symmetries.

Initially I shall assume that no two symmetries are actually co-extensive: that is if g(u) = h(u) for every region u then g = h. If this holds then I say that the group of symmetries acts *faithfully*. In the next section I relax that requirement.

The group of symmetries G inherits a topological structure from the aether. Given any regions u and v, the set U(u, v) = {g ∈ G: g(u) « v} must be open, although it might well be Ø, and the open sets of symmetries are precisely the unions of some of the U(u, v). The theory of topological groups is well developed and, as noted in Chapter Five, under fairly weak constraints G has a unique structure as a Lie group, that is, a group with a natural differentiable manifold structure. Not only does this solve the problem of characterising differentiable structure, but it also enables us to use the classification of Lie groups to describe hypothetical geometric structures for the aether.

In the point-based case, affine aether may be characterised by requiring the following:

1. The group of all symmetries G has just one *transitive* commutative sub-group of symmetries, V, which we call the *vectors*. (To say that V is transitive is to say that given any two point parts u and v of the aether there is some h ∈ V such that h(u) = v.)[149]
2. V has no compact subgroup except the trivial one {Id}.
3. V is connected and locally compact. [150]

We may then treat V as a real vector space, writing g∘k as g + k, Id as 0, and the inverse g^{-1} of g as -g. Then the number of dimensions, M, is the

---

149 (1) implies that G acts faithfully. This is relevant, because it is conceivable that there be vectors and schwectors that are co-extensive. In that case the group of vectors could not be characterised as the only transitive faithful subgroup of all the symmetries.

150 Conditions (2) and (3) imply that V is a connected Lie group. All commutative connected Lie groups of dimension M have the same Lie algebra (with [x,y] = 0 for all x and y.) The one dimensional subgroups must be either circles or lines. Circles are excluded by (2), so V must be a real vector space of M dimensions.

smallest cardinality of any subset X of V that *generates* V. That is, V is the smallest group containing {λg: λ is a real number and g ∈ V}.

I require the convex regions to satisfy the following:

> If regions u and g(u) are both parts of a convex region w then so is h(u) where h = λg and 0 < λ < 1.

I have taken convexity as primitive and required the symmetries to preserve it. I could have chosen some other shape that must be preserved by translation, provided it is sufficiently restrictive to ensure that there is a unique transitive commutative group of symmetries. If we took as fundamental the light cone structure or the associated ordering with respect to absolute, that is, frame-independent, priority, then it would suffice that there is a commutative group that acts faithfully, preserves cones and is *transitive on cones*, in the sense that given any two past light cones u and v some member of the group maps u to v. To ensure that the regions labelled cones are the correct shape, it suffices that the group of all the symmetries leaving any given cone fixed is isomorphic to the group of those Lorentz transformations that do not reverse temporal orientation. This group may be characterised intrinsically.[151]

The requirement of a *unique* transitive commutative subgroup was relaxed in the analogous characterisation of a topological manifold, discussed previously, in which case the symmetries need only preserve the topological structure. The group of symmetries is then rather large but the requirement of being a patch for a topological manifold is that it contains *some* commutative subgroup acting transitively – it will in fact contain infinitely many.

3. Generalising to the non-affine case

In the point-based case we can generalise the symmetry-theoretic account rather easily. For the property of convexity does not require af-

---

151 It is generated by a connected group G and a symmetry j, not in G, that is its own inverse, i.e. j^2 = Id. G must be isomorphic to the connected component of the Lorentz group $SO^+(1,3)$. This in turn is the unique 6 dimensional simple, connected, Lie group.

fine structure.[152] We are to suppose G is the group of all symmetries and that G acts transitively on the aether. But we no longer need them to act faithfully. Then consider any point part **o**. We define an equivalence relation on G by saying that g and h are equivalent if they map **o** to the same point. Then the points are in one to one correspondence with the equivalence classes, the set of which is called the quotient space $G/Stab_{\mathbf{o}}$. If two symmetries act in the same way they must be equivalent. The correspondence correlates any point part v of the aether with the equivalence class of those g sending **o** to v. The point **o** plays the role of the origin in the correlation of an affine Space-time with a vector space. To be sure, a different choice of the origin, **o**′, would result in different equivalence classes, for in the non-commutative case we can have $g(\mathbf{o}) = h(\mathbf{o})$ but $g(\mathbf{o}') \neq h(\mathbf{o}')$. Nonetheless this procedure can be used to describe symmetric structure for the aether in terms of a group of symmetries G and any subgroup H, which may be taken to be $Stab_{\mathbf{o}}$, for the origin **o**.[153]

Two examples illustrate this. The first is that in which the aether is circular in the three spatial dimensions, but still flat, in that it is locally indistinguishable from affine Space-time. In that case the group is commutative and I shall replace multiplicative by additive notation as is conventional. Consider then the commutative group V of symmetries that we recognise as the 4 dimensional vector space. And fix some point as the origin **o**. Then if we represent V as the set of all coordinate quadruples, we may take $Stab_{\mathbf{o}}$ to be H, the set of all integer quadruples. H

---

152 Each member of the Lie group, other than Id, generates either a line or a circle. In the line case we require the same condition on convexity as for the affine case. In the circle case we require that if u and g(u) are parts of a convex region w, then so is h(u) for any h on the *lesser* of the two arcs connecting Id with g. By the lesser arc I mean the one that is mapped into the other arc by inversion. In the special case in which both the arcs are lesser by this definition the definition of convexity requires that both be taken as the lesser arc.

153 We define the equivalence relation by saying that g and g′ are equivalent if for some $h \in H$ $g = g' \circ h$. Then we take the points of the symmetric space to be represented by the equivalence classes. H is itself an equivalence class and represents the origin.

may be characterised up to group isomorphism as the torsion-free commutative group with three generators but not two.[154] The torsion-freedom is in fact redundant because H may be taken as any subgroup of V that has three generators but not two.

We would get the same result if we characterised a group of symmetries acting faithfully and with no fixed points (except for the identity symmetry which leaves every point fixed.)[155] I prefer, however, to consider the whole of V as the group of symmetries on the grounds that the very same universal, a vector with a given length and direction, relates pairs of points in the affine case and the case in which some of the dimensions are curved up in a circle.

Another example is de Sitter Space-time. (I shall ignore the case of anti de Sitter Space-time: interested readers will be able to adapt what I say.) It is the result of replacing affine Space-time by a hyperspherical Space of radius that increases with Time. It may be represented by the set of quintuples $\{<v, w, x, y, z>: w^2 + x^2 + y^2 + z^2 = v^2 + \alpha^2\}$, where α is a constant.[156] We may take the group of symmetries G, in this case acting faithfully on de Sitter Space-time, to be SO(1, 4), that is those 5 by 5 matrices that act on the quintuples $<v, w, x, y, z>$ preserving the quadratic expression $v^2 - (w^2 + x^2 + y^2 + z^2)$ and which preserve orientation. If u is a point represented by $<v, w, x, y, z>$, then $Stab_u$ is represented by those members of G that leave fixed $<v, w, x, y, z>$. This is isomorphic to SO(1, 3) the group of 4 by 4 matrices that preserve the quadratic expression $v^2 - (x^2 + y^2 + z^2)$ and that preserve orientation.

---

154 By *torsion-free* we mean that for any positive integer n if nx = 0 then x = 0. To say H has k *generators* is to say that there are k members not included in any proper subgroup of H.

155 The unique connected commutative Lie group of dimension 4 whose maximal compact subgroups are of dimension 3.

156 The de Sitter cosmology has the cosmological constant $\Lambda = 3/\alpha^2$. The expansion is exponential if time is measured as the proper time of a galactic cluster.

Because there is an independent way of describing SO(1, 4) and SO(1,3) then we have the required characterisation of de Sitter Space-time.[157]

If we compare the group of all symmetries for Sitter Space-time and for affine Space-time we see that the stabiliser of a given point in de Sitter Space-time is a smaller group than that for an affine Space-time of the same number of dimensions. So we may say that affine Space-time is more symmetric because it has an unrestricted spatio-temporal isotropy. In the de Sitter case the stabiliser is a Lorentz group and so isotropy is restricted in that we distinguish space-like, null, and time-like separations. As a consequence, matter-free de Sitter Space-time has a unique light cone structure, whereas affine Space-time has infinitely many.

Affine aether is initially somewhat more probable. First, absent other considerations it was preferable not to privilege the matter-free case as gravity free. (See the previous Chapter.) And second, I say, not merely is isotropy a priori probable, but the greater the isotropy the greater the probability. However, current cosmology supports de Sitter aether. This is because of the posited inflationary expansion soon after the Big Bang (Schmidt 1993). Such expansion is a consequence of de Sitter Space-time, in which case the subsequent deceleration of the expansion is explained as the result of matter. On the other hand if we suppose an affine aether we require something mysterious such as dark energy to explain the inflationary expansion.

---

157 SO(1, 4) has a connected normal subgroup $SO^+(1, 4)$, such that the quotient is the group with just two members. $SO^+(1, 4)$ is one of only three Lie groups of 10 dimensions that is connected and ‘simple’ in the group theoretic sense of having no normal subgroup. A normal subgroup N of a group G is defined as one such that for all $g \in G$ and $h \in N$, $g^{-1} \circ h \circ g \in N$.

These three groups, $SO^+(0, 5)$, $SO^+(1, 4)$, and $SO^+(2, 3)$ differ in the size of their maximal compact subgroup, being of 10 dimensions for $SO^+(0, 5)$, which is itself compact, 6 for $SO^+(1, 4)$ but only 4 for $SO^+(2, 3)$. Once we have picked out $SO^+(1, 4)$ as the unique Lie group of 10 dimensions that is connected, simple and has a compact subgroup of 6 dimensions but no more than 6, we may note that stabiliser of any point u is another 6 dimensional group whose maximal connected subgroup is the unique simple Lie group of 6 dimensions.

As a consequence, the case for the symmetry-theoretic continuous aether over granulated aether is, I regret to say, weakened a little by Bayes' Rule in probability theory, because I 'predicted' affine aether but 'discovered' de Sitter aether.

There is another, happier, consequence. Because the light cone structure is specified in de Sitter aether, we might as well use light cones rather than convexity as the structure that is to be preserved by the symmetries and this will turn out to save the point-free de Sitter case from an objection to be discussed below.

4. Spatio-temporally restricted aether

Those of us who hold a dynamic theory of Time should deny the reality of any future aether. Nonetheless, as I shall argue, we can believe in future Space-time. Maybe some of those who hold that the universe came into existence a finite time ago might likewise hold that Space-time predated the aether. Perhaps some presentists hold that the aether is nothing but a thin layer, a Planck time thick, say. They too might believe in Space-time that has an infinite past and future.

In all these cases we may use the group-theoretic approach to construct Space-time from a *partially symmetric* aether if three conditions hold. The first is that the aether is connected. The second is that it is the sum of open convex regions. (This excludes the case in which the physical universe is made up of point-particles, spatially one dimensional strings, or indeed branes of spatial dimension less than that of Space itself.)

The third condition is that of determinate spatial extent. Its most straightforward version is that in which the aether is infinite in all spatial dimensions. The less straightforward case is that in which there is an upper bound to how far the objects are away but that is because Space is toroidal so as we go farther away on a certain direction we eventually come back to where we start.[158] The condition is that all such questions should be settled. So it excludes the case of a finite universe that could equally well be described as occupying part of an infinite Space or a fi-

---

[158] There are other, less plausible, cases, namely those in which Space is finite in some directions, but infinite in others, like the surface of a cylinder.

nite, toroidal Space. It does not, however, exclude the presentist hypothesis that the aether is a slice of finite temporal thickness because I hold that circular Time is impossible.

We could try using the theory of partially symmetric stuff to provide an alternative to an aether pervading the whole universes. For suppose there are extended particles composed of some sort of stuff, which I shall call prime matter. Then we hypothesise that the prime matter occupies an open part of Space-time. And suppose the particles undergo fission and fusion. Then the prime matter constituting two particles that fuse makes up a connected open region. Likewise for two particles arrived at by fission. Given enough fission and fusion it is plausible that the sum of the prime matter would be spatio-temporally connected. And we may assume it is the sum of open convex regions. So in this case Space-time may be constructed out of the prime matter rather than an all-pervading aether. The same would hold if there was not fission and fusion but particles that collided with distortion so that they came into contact along surface rather than just at a point or just along an edge.[159] My reason for rejecting this way of avoiding the all-pervading aether is that the aether's existence is non-contingent and hence no doubt is cast upon it by a thoroughly contingent arrangement of particles.

A group of partial symmetries may be defined by relaxing some of the conditions required for a group of symmetries, considered as dyadic relations. The relations still correspond to one to one functions, that is, they are relations R_g such that (1) if R_gxy and R_gxz then y = z, and (2) If R_gwy and R_gxy then w = x. But we no longer require that for every point x there is some y such that R_gxy. The identity relation Id relates every region to itself and we still require that for any g in the group there is an inverse g^{-1} its converse. (So if S is the inverse of R, Sxy if and only if Ryx.)

---

159 As far as I can see, contact in just one or two dimensions would permit convexity-preserving symmetries that rotate one particle relative to another, which by enlarging the group of (partial) symmetries would interfere with the symmetry-theoretic characterisation.

I note that the composition of members of the group no longer corresponds to the product of relations. For if $R_g xy$ and $R_h yz$ then, to be sure, $R_{g \circ h} xz$, but the converse need not hold. I shall also assume that the group acts with *local fidelity*. That is if for some open region u g(x) and h(x) are 'defined' and identical for every part of u, then g = h.[160] Because the aether is open and connected this shows that the symmetries are maximal in that there are no two symmetries g and h such that h is the restriction of g to some proper sub-domain. That is, we cannot have g ≠ h such that $R_g xy$ only if $R_h xy$.

A group G of partial symmetries can be used to construct Space-time as follows. Given any point parts u and x we can consider those partial symmetries g such that g(u) = x, including the special case in which u = x and g = Id. In the previous section I considered the case of an all-pervading point-based aether. In that case we can, if we wish, identify Space-time with the aether. But I now consider how we may use the group of partial symmetries to construct the Space-time locations in such a way that every region of the aether has for its locations every member of the corresponding set of properties.

Given any point parts u and x we can consider those partial symmetries g such that g(u) = x. If g is such a partial symmetry, x has the relational property $L_{g;u}$ of having a part standing in the converse of R_g to u. If there is no point-part of the aether x such that this holds then the relational property is uninstantiated. I shall permit this but I still require u to exist for there to be a genuine property $L_{g;u}$. We may define an equivalence relation on the set L_G of those $L_{g;u}$ for which g is in G and u is a point-part of the aether by: $L_{g;u} \sim L_{k;v}$ just in case $h \circ g(u) = h \circ k(v)$ for some partial symmetry h. And we may take a location to be the conjunction of all the members of L_G in a given equivalence class. So every point-part of the aether instantiates a location and Space-time is composed of all the locations, instantiated or otherwise.

This way of constructing Space-time using uninstantiated properties requires that the aether characterises uniquely the group of partial

---

160 To say that g(x) is *defined* is to say that, for some y, $R_g xy$.

symmetries G. The theory of Lie groups shows that the way the group acts on any open region v already significantly constrains the group. Taken together with the conditions stated above, this show there is a unique connected Lie group that acts as partial symmetries on the aether. For example, in the affine case the fact that G has a subgroup H acting on v in a transitive commutative fashion shows that H is commutative. The existence of parts of the aether arbitrarily far away in each spatial direction together with the assumption that Time is non-circular shows that H is isomorphic to the a four dimensional vector space. Every member of L_G is equivalent to some member of L_H, so the set of locations have the required affine structure.

5. Symmetric point-free aether.

To provide a symmetry-theoretic account of point-free aether, such as the favoured hypothesis of Arntzenius Continuum, I adopt the following procedure: (1) Modify the point-based characterisation so that it no longer mentions points; and (2) show how a point-based Space-time may be constructed from the point-free aether. First I begin with a general method that relies on the aether having a light cone structure, or equivalently a suitable partial ordering of absolute (frame-independent) priority, that specifies gravity in the matter-free case. In the last chapter I submitted that it was preferable not to posit such a light cone structure, and in general it was fairly easy to avoid this, using convexity instead. However, de Sitter and anti de Sitter Space-time come equipped with light cones. So this method is appropriate for those structures. Currently it seems quite likely that mass-free Space-time is a de Sitter Space-time, but the a priori preferable case of affine aether should not be ignored. Therefore I also need to discuss an alternative characterisation for the flat case, including affine aether.

Because cones are pointy and because there is only one past light cone through each point, we may use the past light cones as surrogates for points. My reason for preferring past light cones over future ones is that if there is past aether but no future aether I would like the (constructed) locations of the most recent light cones to include the present. I could just as easily have considered double cones.

The symmetries must preserve the mereological structure, the light cones, and hypervolume, as well as preserving the symmetries themselves. The action of a group of symmetries is faithful if no two symmetries act on regions in the same way and transitive if for any two future light cones, u and v there is some symmetry g such that g(u) = v. So the first part of the procedure is easy.

The construction of Space-time is just as straightforward. The Alexandrov intervals are defined as the non-empty intersections of past and future light cones. If w is any past light cone, I say an Alexandrov interval x *fits* w if there is no Alexandrov interval y that is also part of w such that y is disjoint from some past light cone v that x is also part of. Then given any past light cone v and symmetry g we may consider the property $M_{g;v}$ of intersecting every Alexandrov interval that fits g(v). Let M_G be the set of all such $M_{g;v}$ and define an equivalence relation on M_G by $M_{g;u} \sim M_{k;v}$ if $h \circ g(u) = h \circ k(v)$ for some symmetry h. The conjunction of members of an equivalence class (instantiated or otherwise) is then a point location. Every part of the aether may be represented as the set of all locations belonging to it.

This way of characterising the shape of the aether, and of constructing points, requires that there be enough symmetries mapping past cones to past cones, and so is not available if there is a thin present layer of aether – or, on the, already rejected, hypothesis of extended particles made of prime matter. We can, however, modify the hypothesis by requiring not that past light cones be preserved but that Alexandrov intervals (and of course hypervolume) be. This will give us enough partial symmetries and we may then characterise transitivity by saying that given any two Alexandrov intervals of the same hypervolume some partial symmetry maps one to the other.

The cone-free, point-free affine case has one annoying complication, namely the difficulty of characterising a suitable analog for transitivity. The problem is that in the point-free case we require that for any regions u and v of the same shape and with the same size and situated in a parallel way, some translation (vector) maps u to v. A structure of light cones (and hence of Alexandrov intervals) provided us with a suitable conception of being the same shape, and all future light cones are point-ing in the same direction, but things are more complicated if convex regions are used in place of light cones. We can characterise a *half* as a convex region whose complement is convex. In the Arntzenius Continuum case all halves are the same shape but they are not all arranged in a parallel way. Hence we require rotations as well as translations to map any half to any half. We can characterise transitivity by requiring that given any halves u and v, if u and v have parallel boundaries, and so 'point' in the same direction, some symmetry maps u to v. If the aether extends infinitely in all directions, including the future, this parallel boundary condition may itself be analysed as one half being part of the other.

A problem arises, however, for those who hold that there is no future aether. For then two hyperplanes in Space-time that intersect in the future will turn out to bound halves of the aether one of which is part of the other. Initially, therefore, we might infer that the combination of belief in affine Space-time and no future aether excludes Arntzenius Continuum. Instead of drawing this conclusion we could adapt the characterisation of topological manifold in the point-free case, described in Chapter Four. But it gets messy. It is neater to treat *being parallel* as primitive, but that is uneconomic.

The point locations may be constructed using ultrafilters.[161] In the case in which the aether does not occupy the whole of Space-time, con-

---

161 We might also construct Space-time by noting that each point p of Space-time sets up a way of assigning subsets of the group V of vectors to regions, namely the vectors that map p to points in the region. In the case of all-pervading aether, we may use this to characterise a point as a map ψ from non-empty convex open sets of vectors to convex open regions such that:

1. If U is any non-empty set of non-empty open convex sets of vectors whose union u is convex, then $\psi(u)$ is the least upper bound of $\{\psi(x): x \in U\}$.
2. $\psi(X \cap Y) = \psi(X) \wedge \psi(Y)$ if $X \cap Y \neq \emptyset$; $\psi(X)$ and $\psi(Y)$ are disjoint if $X \cap Y = \emptyset$
3. If v is any vector, $\psi(\{v + x: x \in X\}) = v(\psi(X))$.

A mapping ψ may then be said to be a *location* of a region u if every open convex set of vectors that contains the identity (zero) vector is mapped by ψ into some region that overlaps u. This construction is, however, more com-sider the property $N_{g;v}$ that a region of aether has/would have if intersects/would intersect g(v). Let N_G be the set of all such $N_{g;v}$ and define a transitive anti-symmetric relation on N_G by $N_{g;u} \ll N_{k;v}$ if $h \circ g(u)$ is an interior part of $h \circ k(v)$ for some h. Then a point location is the conjunction of all the properties in N_G that belong to a given ultrafilter with respect to «. This is a large conjunction of fairly straightforward relational properties, which I judge to be a satisfactory structure for a non-fundamental property to have.

6. Extending symmetry to many 'worlds'

In the Introduction, I noted that we might hold that there are many worlds but that I would concentrate on just one of them. That is justified provided we are considering many spatio-temporally disconnected worlds, as in Lewis' theory. In that case it is strictly speaking false to say that the aether has, for instance, affine structure. But the required modification of hypotheses is obvious – each connected component has, in this case, affine structure.

Instead, we might well hold that that actual Universe is composed of many spatio-temporally connected universes. Either they are now connected (with respect to any frames of reference) or, if no longer connected, they form a branching structure, where in a branching Universe, given any two regions u and v there is some region that absolutely prior to them both (McCall 1994). If we hold some such 'many worlds' theory, perhaps as an interpretation of quantum theory, then we should examine the implications for the investigation of the structure of the aether.

The branching Universe is the sum of overlapping universes. Given any two such universes x and y their meet $x \wedge y$ is a like an non-branching universe up to some Space-like hypersurface. If the branching is ordered by the integers then this may be taken as the discrete Time mentioned in Chapter Six, which I argued was compatible with a continuous aether. If we are realists about both the future and the past aether then each of the universes may be characterised using symmetry. They might, for in-

---

plicated than resort to ultrafilters when interpreted in terms of an ontology of properties and relations.

stance, all have affine structure or all have de Sitter Space-time structure. The only implication seems to be that a branching Universe supports the distinction of Time from Space-time and so offers a reply to the objection, that some might find initially persuasive, that this distinction is an ad hoc way of saving continuous theories from arguments for discrete Time.

If the many universes are always connected, then they may be interpreted as being located in 4 dimensional *fibres* that an N + 1 dimensional Space-time is divided into, where N, the number of spatial dimensions, is greater than 3. (I use the term 'fibres' even though they have 3 spatial dimensions because this is standard mathematical terminology and the higher dimensional Space-time would be said to have a *fibration.* See Rowland, 2009) We, the observers, may then be thought of as having more than 3 spatial dimensions, with the apparent indeterminacy of the physical universe being due to the very slight differences between the states of the fibres we overlap. Given the case for Arntzenius Continuum the 4 dimensional universes are themselves constructs, like points because every region is of more than 4 dimensions. But that's fine.

The N + 1 dimensional aether could have the structure of a higher dimensional Minkowski Space-time even though the fibres are de Sitter Space-times. For simplicity consider N = 4. Then for any k > 0, the hypersurface $\{<v, w, x, y, z>: w^2 + x^2 + y^2 + z^2 - v^2 = k\}$ represents a de Sitter Space-time. By varying the k we can stack the hypersurfaces up neatly to obtain a fibration of the region of a 5 dimensional Minkowski aether represented by $\{<v, w, x, y, z>: w^2 + x^2 + y^2 + z^2 - v^2 > 0\}$.

The higher dimensional Minkowskian aether may easily be characterised using symmetries, using the pointy nature of the light cones. But there is a further implication: I do not know how to characterise de Sitter aether axiomatically, but the Minkowskian case is amenable to an axiomatic treatment.[162] This overcomes an obstacle in the way of those nom-

---

162 The difficulty with an axiomatic characterisation of de Sitter spaces is that spatially they are like spherical geometry but temporally they are like a hyperbolic geometry. So nominalist might need to distinguish Space from Time to overcome this difficulty.

inalists who would otherwise accept symmetric Arntzenius Continuum, but who reject the group-theoretic characterisation of symmetric aether, and who take de Sitter Space-time to be more probable than Minkowski Space-timefor the non-quantum 4 dimensional case.

If some or all of the 'worlds' are of infinite extent, either spatially or temporally, then we might require an infinite dimensional Minkowskian Space-time to contain them all.

Conclusions

There is variety of structures that the aether might, for all we *know*, have is quite bewildering. Using a combination of a priori reasoning and contemporary physics I have, however, reached the following conclusions about its likely structure.

1. Either the aether is granulated, and so point-free, or it is continuous and, in that case, probably also point-free.
2. Even if the aether has point parts, we should reject the thesis that all regions (i.e. parts of aether) are grounded in points by summation. We should also reject the Orthodoxy that the regions are in one to one correspondence with the sets of points and even the weaker position that the regions are in one to one correspondence with measurable sets of points.
3. The above orthodoxies could, to be sure, be defended by rejecting the Axiom of Choice but even so the resulting hypotheses are not especially attractive.
4. If the aether is point-free and continuous its most likely structure is that of Arntzenius Continuum
5. If the aether is granulated, its most likely structure is that of Pseudo-set Granulation.
6. Both Arntzenius Continuum and Pseudo-set Granulation have some counter-intuitive features, but there is no intuitively perfect hypothesis about the nature of the aether.
7. The aether may be identified with Space-time as traditionally understood if: realism about past and future is correct, if the aether is continuous; and if every region is the sum of points. In all other cases Space-time is either a fiction or a construct. In the case of discrete aether, it is a fiction. In the continuous case, Space-time is a construct – more precisely spatio-temporal locations are genuine properties analysable in terms of the aether.
8. Probably, the aether is symmetric if and only if it is continuous, in which case it has the structure of affine, de Sitter Space-time or anti de Sitter Space-time. On a many worlds interpretation of

quantum theory we may, however, take the aether to have a higher dimensional Minkowski Space-time structure.

9. If the aether is continuous we should reject the interpretation of General Relativity according to which the geometric structure of Space-time is contingent and determines gravity. Gravity is not, on this interpretation, a force. This rejection is based upon the serious difficulty of characterising a differentiable manifold without hypothesising further non-contingent structure.
10. Contemporary physics undermines an objection to the discrete aether based on scale invariance, but does not offer much positive support for discrete aether. For what it is worth, the current state of physics still suggests that the aether is continuous, and so symmetric. If instead we rely only upon well-established physics and a priori reasoning the choice between discrete and continuous aether depends on how we weigh up simplicity, which favours discretion against symmetry, which favours continuity
11. The hypothesis of a continuous symmetric aether might be problematic if its shape is not flat but instead (part of) de Sitter or anti de Sitter Space-time. For these shapes can be characterised in a straightforward way using a group of symmetries but I do not know how they can be characterised axiomatically. So those (extreme) nominalists who reject quasi-realism about universals have a problem. They might like to invoke a higher dimensional Minkowski structure for the aether in which de Sitter or anti de Sitter Space-times are embedded, as in the many worlds interpretation of quantum theory.
12. Conversely, if the aether is flat but without an infinitely extended past or future there are difficulties with point-free continuous aether hypotheses such as Arntzenius Continuum. In that case we might prefer Borel Continuum.

Glossary

≈. x ≈ y if x and y are approximately equal.

¬. u = ¬v if u is the *complement* (q.v.) of v.

u « v. u is an *interio*r part (q.v.) of v.

Absolute priority. Region u is *absolutely prior* to region v if with respect to every frame of reference every part of u is earlier than every part of v.

Adjacency. Two regions are *adjacent* if they are not *separated* (q.v.)

Œ. The set of coordinate quadruples representing œ.

œ. A region of aether to which attention is restricted.

Affine aether. The aether is said to be *affine* if it is both flat and infinite in all directions.

Affine connection. An affine connection specifies a way of correlating energy-momenta at q with energy-momenta at a point q′, displaced from q by a small amount. It must preserve the affine structure for the space of energy-momenta. It must also preserve the Minkowski structure of this space, and so is, strictly speaking, a Minkowski connection.

Aleph Null hypothesis. The hypothesis that œ (q.v.) is the sum of a countable infinity of points.

*Alexandrov interva*l. A non-empty meet of a past and a future *light cone.*(q.v.)

Almost identity. X is almost identical to Y if they differ by a set of *Lebesgue measure* (q.v.) zero, that is if X – Y and Y – X are both of zero measure.

Arbitrarily Fine Covering (Premise Five of Ch. 2). The principle that if N is a positive integer and if u is a *globule* (q.v.) of finite diameter then u is part of the sum of at most a countable infinity of globules of diameter less than 1/N.

Arbitrarily Thin Boundaries (Premise Six of Ch. 2). The principle that if u is a *globule* (q.v.) of finite diameter, and N is a positive integer, then there is a region v of hypervolume less than 1/N such that any *connected* (q.v.) region w that overlaps u, but is not part of u, overlaps v.

Arntzenius Continuum. Arntzenius Continuum is obtained from the *Borel Continuum* (q.v.) by first ignoring parts of zero hypervolume and then identifying those that differ by zero hypervolume.

Atom. An atom of aether is a region u that is not the sum of the proper parts of u. If we assume classical mereology every atom is a *simple* (q.v.)

Automorphism. A 1 to 1 onto *homomorphism* (q.v.) from a system to itself whose inverse is also a homomorphism.

Ball. A *closed hyperball* b of radius r centre <a, b, c, d> is the set of co-ordinate quadruples <t, x, y, z> such that:

$$(t - a)^2 + (x - b)^2 + (y - c)^2 + (z - d)^2 \leq r^2.$$

An *open hyperball* b of radius r centre <a, b, c, d> is the set of co-ordinate quadruples <t, x, y, z> such that:

$$(t - a)^2 + (x - b)^2 + (y - c)^2 + (z - d)^2 < r^2.$$

The analogs for triples are the closed and open *balls*.

Banach Tarski theorem. The theorem that a *ball* (q.v.) of radius 1 is the union of five sets congruent to five other sets whose union is two balls of radius 1.

Boolean algebra of sets. A set X of subsets of B form a Boolean algebra if the complement, B - Z of any member, Z, of X is also in X and if the union of any two members of X is also in X.

Boolean lattice. A *distributive lattice* (q.v.) is said to be *Boolean* if for every member x of the lattice, ¬¬x = x. (See *complement.)*

Boolean mereology. A mereology that becomes a *Boolean lattice* if the fictitious *empty* region ø is adjoined.

Borel Continuum. The hypothesis that the regions are faithfully represented by the non-empty *Borel sets* (q.v.) of coordinate quadruples.

Borel sets. The Borel sets of coordinate quadruples form the smallest σ-*algebra* (q.v.) containing all the open sets of coordinate quadruples.

Characterization problem. The problem of describing structure without reversing the order of explanation.

cl(U). The topological *closure* (q.v.) of a set U.

Classical mereology. (a.k.a. *general extensional mereology*). The system obtained by adjoining to the transitivity and anti-reflexivity of parthood the axiom that any regions have a unique *fusion* (q.v.)

Closed. A *closed* set is one whose complement (q.v.) is open.

Closure. The *closure* of a set is the intersection of all the closed sets that include it.

Comm. A (often *the*) maximal commutative group of symmetries.

Compact. In point-set topology, a set X is said to be *compac*t if every *open covering* (q.v.) Y of X has a finite subset Z that is also a covering of X. In the point-free case, a region x is said to be compact if every *exterior covering* (q.v.) Y of x has a finite subset that is also an exterior covering.

Complement. The complement ¬x of a region x is the join of all the regions disjoint from x.

Complete Heyting lattice. A *distributive lattice* (q.v.) in which meets distribute over arbitrary joins.

Complete Heyting mereology. A mereology that becomes a *complete Heyting lattice*(q.v.) if the fictional empty region ø is adjoined.

Connected Parts (Premise One of Ch. 2). The principle that every region has a part that is connected.

Connected region. A region that is not the *sum* (q.v.) of two parts separated from each other.

Consistent Renormalization research program. The research programmed based on the core assumption that the standard approach to quantum field theory can be developed without divergent (i.e. infinite) integrals.

Convex region. A region u is *convex* if given any two parts x and y of u every region between x and y is also part of u.

Countable Subadditivity. The following principle. Consider a sequence of regions, u_1, u_2, etc such that, for all j, u_j has hypervolume. Then the hypervolume of the *sum* (q.v.) is no greater than the sum of the hypervolumes of the u_j.

Cross Section principle. The principle that given any equivalence relation on a set Y there is a subset X of Y that contains precisely one member of each equivalence class.

Curmudgeon variant. A variant on a hypothesis about aether structure that treats disconnected regions as fictions.

Definite Range. The principle that the value of an 'observable' is definitely in the range from the great lower bound to the least upper bound of the eigenvalues.

Dependent Quantity. If something depends ontologically on some disjoint parts then its quantity is independent of the relations between those parts.

diam(u). The diameter of region u.

Diameter Hypervolume Nexus (Premise Three of Ch. 2). The principle that given any positive integer M there is some positive integer N such that any region of diameter less than 1/N is part of a region of hypervolume less than 1/M.

Difference. The *difference* of regions x and y, x – y is the join of all the parts of x disjoint from y. In a *lattice* (q.v.) $x - y = x \wedge \neg y$.

Distributive lattice. A *lattice* (q.v.) of regions in which the distributive laws hold: $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ and $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$.

Distributive mereology. A mereology that becomes a *distributive lattice* if the fictitious empty region ø is adjoined.

Einstein Aether. A physical theory with a *foliation* (q.v.) of Space-time into Space-like hypersurfaces specified as *normal* (i.e. ‘perpendicular’ with respect to the ‘metric’ on Minkowski Space-time) to a distinguished vector field.

Einsteinian manifold. A differentiable manifold equipped with a general relativistic ‘metric’.

Endurance. Something that persists from one time to another is said to *endure* if it lacks proper temporal parts (except in so far as it acquires or loses proper spatial parts.) Contrast *perdurance* (q.v.)

Exemplars. The four *exemplars* are the hypotheses that are judged best at an early stage in the inquiry.

Extended Simples. The hypothesis that the aether is the *sum* (q.v.) of disjoint extended granules and that these granules are not merely atomic but *simple* (q.v.)

Exterior covering. An exterior covering Y of x, is a set of regions for which there is a set W of regions such that: for all $w \in W$ there is some $y \in Y$ for which w « y; and x is part of any region z of which every $w \in W$ is part.

Faithful representation. A 1 to 1 representation.

Fibres. Space-time is *fibrated* if it is the *sum* (q.v.) of a set of disjoint parts F, the fibres, where F itself has a Space-like structure, and each member of F is as we ordinarily take Space-time to be, namely 4 dimensional (or more in Supergravity theories.)

Filter. A filter, w.r.t. a transitive anti-reflexive relation <<, is a set W such that:

(1) if $x \in W$ and $x << y$ then $y \in W$; and

(2) if $x \in W$ and $y \in W$ then there is some $z \in W$ such that $z << x$ and $z << y$.

Fine Structure. Structure at the infinitesimal scale.

Finite Globules (Premise Four of Ch. 2*)*. The principle that there is a globule of finite positive diameter and finite positive hypervolume, namely œ (q.v.)

Finite Subadditivity (Premise Nine of Ch 2). The principle that if x and y have hypervolumes, then $hvol(x + y) \leq hvol(x) + hvol(y)$.

Flat aether. Flatness is an intuitive notion but may be explicated as having a unique *transitive* (q.v.) commutative group of symmetries.

Fusion (in mereology). The fusion of the Xs is some region that overlaps all and only the regions that overlap some X.

Geometric Correspondence. The principle that any extended object has parts corresponding to the parts of the region it occupies.

Geometrodynamics. The project of using gravity to provide a 'theory of everything'.

Globules. Regions that are topologically equivalent to *hyperballs* (q.v.).

Granulated Aether. Extended Simples (q.v.) and its variants. On these hypotheses there are said to be significant small (presumably Planck scale) regions that are the *granules*.

Granules. Extended *atoms* (q.v.) of aether.

Granuloid. An extended region that is not the *sum* (q.v.) of two or more regions of approximately the same quantity.

Grounding (ontological). See *ontological dependence.*

Growing Block. The hypothesis that the past and present *exist* in the pre-tensed sense of 'exists' but not the future.

Gunk. Something is said to be gunk if it has no parts that are atoms. Point-free aether may loosely be said to be gunky, but that presupposes that we ignore *fine-structure* (q.v.)

Gδ Continuum. The hypothesis that the regions of the aether are represented faithfully (q.v.) by the non-empty Gδ sets of coordinate quadruples, that is the countable intersections of open sets.

Homomorphism. A mapping from a mathematical system X to a mathematical system Y that preserves whatever structure is being considered in the context.

Hume's Razor. Necessities are not to be multiplied more than necessary.

hvol(u). The hypervolume of u.

Hyperball. See *ball.*

Hybrid Granulation. A hypothesis on which there are not just simple points but also simples of higher dimension such as the edges.

Hypervolume Subadditivity. The principle that hvol(x + y) ≤ hvol(x) + hvol(y).

Hypervolume Supplementation (Premise Eight of Ch. 2).The principle that if x is part of y and hvol(y) > hvol(x) > 0, then y has a part z disjoint from x such that hvol(z) > 0.

Interior Part Supplementation. The principle that if x is an interior part of y then there is some part z of y disjoint from x.

Interior part. Region x is an interior part of y (x « y) if x is *separated* (q.v.) from every region z disjoint from y

Interior. The interior of a set is the union of all the open sets included in it.

Join. The join of the Xs*,* ∨X, is least upper bound of the Xs, that is, an *upper bound* (q.v.) of the Xs, that is part of every other upper bound.

Lattice mereology. A mereology in which any two regions with a *lower bound* (q.v.) has a *meet* (q.v.), and any two regions have a *join* (q.v.)

Lattice of regions. The partially ordered system obtained from the mereology by adjoining the fictitious empty region ø is said to be a lattice if for any two regions x and y there is a *join* x∨y (q.v.) and a *meet* x∧y (q.v.)

Lebesgue Continuum. The hypothesis that all and only the non-empty *measurable* (q.v.) sets represent regions.

Lebesgue measure of a set of quadruples. The hypervolume of a region in a fictitious 4 dimensional Euclidean space whose points are represented using Cartesian coordinates t, x, y and z.

Light cones. The past and future light cones are stipulated to be solid, that is their location includes all points separated from the vertex by either a light ray or a time-like geodesic.

Limiting Hypervolumes (Premise Ten of Ch. 2). The following principle: Consider some totally ordered regions (i.e. given any two of them one is a part of the other) each of which has hypervolume less than or equal to k. Then, if the regions have a *sum* (q.v.) this sum has hypervolume less than or equal to k.

Linear Ordering. A relation $x < y$ is a linear ordering if (a) it is transitive; and (b) the derived relation , *neither* $x < y$ *nor* $y < x$ is an equivalence relation.

Locale Continuum. The hypothesis that the regions of aether are represented faithfully by the non-empty open sets of coordinate quadruples.

Lower bound. A lower bound of a set of regions X is a region u that is part of every member of X.

Maximal open set. An open set U such that if V is an open set such that $U \subset V$ then $V - U$ is of positive *Lebesgue measure* (q.v.)

Maximal point. A point that is not part of any other point.

Measurable set. A set with a *Lebesgue measure* (q.v.).

Meet. The meet $\wedge X$ of a set X of regions is the greatest lower bound, that is, a lower bound of which every *lower bound* (q.v.) is part.

Mereology Hypervolume Nexus (Premise Seven of Ch. 2). The conjunction of the following.

(1) If x is part of y, which is part of z, if x and z have hypervolumes, and if hvol(x) = hvol(z) then y has a hypervolume and, hence hvol(x) = hvol(y).

(2) If y is part of z and z is of zero hypervolume then y has a hypervolume, so y has zero hypervolume.

Metathety. The relation holding of x, y and z if y is *between* x and y.

Norm. $\Re^4$ is equipped with a *norm*, that is, the distance from <0,0,0,0>, defined by $||<t,x,y,z>|| = \sqrt{(t^2 + x^2 + y^2 + z^2)}$,

Normal subgroup of G. A subgroup N of G is said to be *normal* if for all $g \in G$ and $x \in N$, $x \circ g \circ x^{-1} \in N$.

Ontological dependence. To say that the Ys *depend ontologically* on the Xs is to say that the Ys exist in virtue of the Xs, or that the Xs are the ontological grounds for the Ys. A necessary but insufficient condition for the Ys to depend ontologically on the Xs is that the Ys supervene on the Xs.

Open covering. Y is an open covering of X if every member of Y is open and $X \subseteq \cup Y$.

Orthodoxy. The hypothesis that parts of the aether are in one to one correspondence with the non-empty sets of all quadruples of real numbers

Pentatope (a.k.a. 5-simplex). 4 dimensional analog of a tetrahedron, i.e. a 4 dimensional *polytopic* (q.v.) region with only 5 vertices.

Perdurance. A thing x that persists over an interval of Time is said to *perdure* if it is the *sum* (q.v.) of temporal parts none of which persist over the interval in question.

Pinwheel tiling. A certain tiling made up cells of the same intrinsic shape and size but with some the mirror images of others. (See Diagram One.)

Point. A region of zero diameter,

Point Discretion. The hypothesis that any part of the aether of finite diameter is the *sum* (q.v.) of finitely many points.

Polytope. Analog of a convex polyhedron in any number of dimensions: a convex region with only a finite number of vertices.

Presentism. The thesis that neither past nor future but only the present *exists* in the pre-tensed sense of 'exists'.

Prime matter. The stuff of which everything is made. If prime matter pervades the whole universe I call it the aether.

Pseudo-set Granules. The hypothesis in which closed *polytopic* sets represent aether atoms but these atoms are not *simple* (q.v.) because the hyperfaces, faces edges and vertices of the polytopic sets also represent regions.(See *polytope*.)

Quasi-realism about universals. The thesis that we may talk as if there are universals even if this can be paraphrased by nominalists.

Quotient space. If H is a subgroup of a group G then the *quotient space* G/H has for its 'points' the equivalence classes of members of G, where g and k are equivalent if $k \circ g^{-1} \in H$.

$\Re^4$. The topological space of all quadruples of real numbers.

Realism about universals. The thesis that universals exist, not that they are fundamental, which I call *fundamentalism* about universals.

Regular open set. In a topological space an open set is said to be *regular open* (or *perfectly open*) if it is the interior (q.v.) of its closure (q.v.).

Regular regions. if x = ¬¬x then x is said to be *regular*. (See *complement*.)

Resilience. A belief is *resilient* if it is not easily defeated.

SAD. The hypothesis that fundamental reality can only be described mathematically.

Scattered Object Argument. The argument that there are scattered objects and so not all regions are connected.

Separation. Two regions y and z are *separated* if there are real numbers ξ and η that are possible values of the diameter function, such that $\xi < \eta$ and any region overlapping both y and z has diameter greater than η.

Sigma algebra (σ-algebra) of sets. A set X of sets such that given any set U in X the complement of U, X – U is also in X, and given any sequence of members of X, U_1, U_2 etc, indexed by the positive integers, the union of the U_j is also in X.

Simple. A region x is *simple* if it has no part other than x itself.

Skeletal Granulated Aether. The hypothesis that all aether *atoms* (q.v.) are of 1 dimension.

Smoo. The ring (better real algebra) of smooth (i.e. infinitely differentiable) real functions on a differentiable manifold.

Smooth boundary condition. A necessary condition for function f to be differentiable is that: (1) for every real number k the sum of all points x such that $f(x) < k$ has a smooth boundary, and (2) so does the sum of all the points such that $f(x) > k$.

Solovay's Axiom. An alternative to the Axiom of Choice in which all sets are measurable.

Space. Because mathematicians use the term 'space' freely to refer to systems with properties that are reminiscent of geometry, I am adopting the convention that 'Space' is written with an upper case 'S' when used literally.

Sparse Continuum. The hypothesis that the regions are represented faithfully by the non-empty members of the smallest *Boolean algebra* (q.v.) of sets of coordinate quadruples that contains all open convex sets.

Stab. $Stab_u$, the *stabilizer* of u is the subgroup of all the symmetries that leave u fixed.

Sum. The term 'sum' is reserved for the non-technical idea of the combination of two regions and not be used as synonymous with *fusion* (q.v.).

Summation mereology. A *lattice mereology* (q.v.) in which any regions have a *sum* (q.v.). See also *Universal Summation.*

Superadditivity. The principle that if regions u and v are disjoint and if u and v are both parts of w, then hvol(w) ≥ hvol(u) + hvol(v).

Supersponge. A *supersponge* for a region u of positive hypervolume is some part v of u that is of positive but strictly less hypervolume than u, such that there is no connected part of u that is of positive hypervolume but disjoint from v.

Symm. The group of symmetries of the aether with respect to some designated structure.

Tarski Continuum. The hypothesis that the regions are represented faithfully by precisely those sets that are non-empty *regular open* (q.v.) sets of coordinate quadruples.

Temporary Intrinsics, *problem of*. If regions *endure* (q.v.) then the same region would have to have different intrinsic properties at different times, contrary to the Indiscernibility of Identicals.

Tesseract. 4 dimensional analog of the cube.

Torsion. Ambiguous between the torsion component of an *affine connection* (q.v.) and the existence of a positive integer n such that for some non-zero member x of a designated commutative group nx = o.

Transitive action. A group of symmetries G *acts transitively* on a set X if for any y ∈ X and z ∈ X, there is some g ∈ G such that g(y) = z.

Triangle Inequality. The principle which holds for genuine metrics but not the 'metric' of Relativity, that for any points x, y and z the distance between x and z is less than the sum of the distances between x and y and between y and z.

Ultrafilter. A *filter* (q.v.) W such that if W ⊆ V and V is also a filter then W = V.

Universal Summation (Premise Two of Ch. 2). The principle that any regions have a *sum* (q.v.)

Upper bound of the Xs. A region of which every X is a part.

Weak Hypervolume Supplementation (Premise Eight* of Ch. 2). The principle that if x is part of y and hvol(y) > hvol(x) > 0, then y has a part z such that hvol(z) > 0, and either x and z are disjoint or hvol(x∧z) = 0.

Weak Supplementation. The principle that if x is a proper part of y then there is some part z of y disjoint from x.

Weyl Tile problem. The problem for discrete aether that a suitably regular arrangement of granules seems to imply a metric that is not even approximately that of a Euclidean space.

Zooming-in. The principle that that we can reliably zoom-in on ever smaller regions treating them as if they were macroscopic.

References

Alekseevskij, D. V., Vinogradov, A. M. and Lychagin, V. V., *Geometry I: Basic Ideas and Concepts of Differential Geometry*, (trans. E. Primrose), Berlin: Springer-Verlag, 1988.

Ambjørn, J., Carfora, M., A., and Marzuoli, *Quantum Geometry A Statistical Field Approach,* Cambridge: Cambridge University Press 1997.

Armstrong, David, *A Theory of Universals: Universals and Scientific Realism Volume II*, Cambridge: Cambridge University Press, 1978.

Arntzenius, Frank, 'Gunk, Topology and Measure', *Oxford Studies in Metaphysics* (ed. Dean Zimmerman) 2008.

Ashtekar, A., Baez, J., Corichi, A. and Krasnov, K., 'Quantum Geometry and Black Hole Entropy', *Physical Review Letters*, 80, 904-907 (1998): 904-907.

Baez, John C, 'The quantum of area?', *Nature* 421 (2003): 702-703.

Bishop, John, *Believing by Faith*, Oxford: Clarendon Press, 2007.

Braddon-Mitchell, David and Miller, Kristie, 'The Physics of Extended Simples', *Analysis* 66 (2006): 222-226.

BonJour, Laurence, *In Defence of Pure Reason: A Rationalist Account of a Priori Justification,* Cambridge: Cambridge University Press, 1998

Callender, Craig, 2002, "Philosophy of Space-Time Physics," in *The Blackwell Guide to the Philosophy of Science*, P. Machamer (ed.), Cambridge: Blackwell. 173-198.

Cartan, Élie, 'Sur une généralisation de la notion de courbure de Riemann et les espaces à torsion.' *C. R. Acad. Sci.,* 174 (1922): 593-595.

Casati, Roberto and Varzi, Achille, *Parts and Places: The Structures of Spatial Representation*. MIT Press, Cambridge: 1999.

Coppel, W. A., *Foundations of Convex Geometry,* Cambridge: Cambridge University Press, 1998.

Craig, William Lane, *The Kalam Cosmological Argument*. Eugene: Wipf & Stock Publishers, 2000.

———, *Time and the Metaphysics of Relativity,* Dordrecht: Kluwer 2001.

Donaldson, Simon, 'Self-Dual Connections and the Topology of Smooth 4-Manifold.' *Bulletin of the American Mathematical Society*, 8 (1983): 81-83.

Dorr, Cian, 'Physical Geometry and Fundamental Metaphysics', *Proceedings of the Aristotelian Society*, 111 (2011): 135-159.

Dougherty, Randall and Foreman, Mathew, 'Banach-Tarski paradox using pieces with the property of Baire', *Proc. Nati. Acad. Sci. USA* (1992): 10726-28.

Douglas, Michael R. and Kachru, Shamit, 'Flux Compactification', *Rev. Mod. Phys.* 79, (2007): 733-796.

Field, Hartry, *Science without numbers: a defence of nominalism*, Princeton: Princeton University Press, 1980.

Fine, Kit, 'Ontological Dependence', *Proceedings of the Aristotelian Society, New Series*, 95, (1995): 269-290.

Forrest, Peter and Armstrong, David, 'The nature of number', *Philosophical Papers* 16 (1987): 165-186.

Forrest, 'Is Space-Time Discrete or Continuous? — An Empirical Question', *Synthese* (1995): 327-354.

———, 'Nonclassical Mereology and Its Application to Sets', *Notre Dame Journal of Formal Logic*, 43 (2002): 79-94.

———, 'Grit or gunk: Implications of the Banach-Tarski paradox', *The Monist* 87 (2004): 351-370.

———, *Developmental Theism: From Pure Will to Unbounded Love*, Oxford: Oxford University Press, 2007.

———, 'Relativity, the Passage of Time and the Cosmic Clock', in Dennis Dieks (ed.), *The Ontology of Space-time II*, Amsterdam: Elsevier, 2008: 245-254.

———, 'Razor Arguments', in Robin Le Poidevin (ed.), *The Routledge Companion to Metaphysic,* Routledge, 2009.

———, 'Mereotopology without mereology', *Journal of Philosophical Logic* (2010)

———, 'Conflicting Intuitions about Space', in Shieva Kleinschmidt (ed.) *Mereology and Location*, Oxford (2012)

Gallois, André, "Identity Over Time", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/identity-time/>.

Goldman, Alvin, 'Discrimination and Perceptual Knowledge', *Journal of Philosophy,* 73 (1976): 771-91.

Grätzer, G. *Lattice Theory: First Concepts and Distributive Lattices*, San Francisco, W. H. Freeman, 1971.

Grünbaum, Adolf, 'Is a Universal Nocturnal Expansion Falsifiable or Physically Vacuous?" *Philosophical Studies* 15 (1964) 71-79.

———, Adolf 'The Denial of Absolute Space and the Hypothesis of a Universal Nocturnal Expansion: A Rejoinder to George Schlesinger,' *Australasian Journal of Philosophy 4* (1967): 61-91

Hales, Steven D, 'The Problem of Intuition', *American Philosophical Quarterly*, 37 (2000): 135-46.

Hazen, Allen, 'Hypergunk', *The Monist* 87 (2004): 22-38.

Helgason, Sigurdur, *Differential Geometry and Symmetric Spaces*, New York: Academic Press, 1962.

Hudson, Hud, *The Metaphysics of Hyperspace*, Oxford: Oxford University Press, 2005.

———,‘Simple Statues’, *Philo* 9 (2006): 40-46.

Jacobson, Ted, ‘Einstein-æther gravity: a status report’ 2008, URL = <arXiv:0801.1547v2 [gr-qc]>.

Johnstone, Peter, *Stone Spaces*, Cambridge Studies in Advanced Mathematics 3, Cambridge University Press, Cambridge, 1982.

Kant, Immanuel, *Gesammelte Schriften*, ed. Akademie der Wissenschaften, Berlin: Reimer, 1910.

Lakatos, Imre and Musgrave, Alan, ‘Falsification and the Methodology of Scientific Research Programmes’ in *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 1970: 91-196.

Le Poidevin, Robin, ‘Space, supervenience and substantivalism’, *Analysis* 64 (2004): 191-98.

Lewis, David, *On the Plurality of Worlds*, Oxford: Blackwell, 1986.

———, *Parts of Classes,* Cambridge, Massachusetts: Basil Blackwell, 1991.

Luce, R. Duncan, and Suppes, Patrick, ‘Representational measurement theory’, in H. Pashler and J. Wixted (eds.), *Stevens’ handbook of experimental psychology: Vol. 4. Methodology in experimental psychology* (3rd ed.), New York: Wiley, 2002: 1-41.

Lurie, Jacob, ‘Whitehead Triangulations (Lecture 3)’, 2009. URL = <www.math.harvard.edu/~lurie/937notes/937Lecture3.pßdf>

Malament, David, 'The class of continuous timelike curves determines the topology of Space-time', *Journal of Mathematical Physics* 18 (1977): 1399

Markosian, Ned, 'Simples', *Australasian Journal of Philosophy* 76 (1998): 213-228.

Maudlin, Tim, *The Metaphysics within Physics*, Oxford: Oxford University Press, 2007.

McCall, Storrs, *A model of the universe: Space-time, probability, and decision* Oxford: Clarendon Press 1994.

McDaniel, Kris, 'Extended simples', *Philosophical Studies* 133 (2007a): 131-41.

———,'Distance and Discrete Space', *Synthese* 155 (2007b): 157-162.

McLaughlin, Brian and Bennett, Karen, "Supervenience", *The Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/ entries/supervenience/>.

Moschella, Ugo 'The de Sitter and anti-de Sitter Sightseeing Tour', *Séminaire Poincaré*, 1 (2005): 1-12.

Nash, John, 'The imbedding problem for Riemannian manifolds", *Annals of Mathematics*, 63 (1956): 20-63.

Nerlich, Graham, *The Shape of Space,* Cambridge: Cambridge University Press, 1994.

———, 'Space-Time Substantivalism,' in *The Oxford Handbook of Metaphysics*, M. J. Loux (ed.), Oxford: Oxford University Press. 281-314, 2003

Newton, Isaac, *Universal Arithmetic: Or, a Treatise of Arithmetical Composition and Resolution*, London (1728).

Nolan, Daniel, ‘Classes, Worlds and Hypergunk’, *The Monist*, 87 (2004), 3-21

North, Jill, ‘The "Structure" of Physics: A Case Study’, *Journal of Philosophy* 106 (2009): 57-88.

Norton, John D., ‘The Hole Argument’, *The Stanford Encyclopedia of Philoso*phy Winter 2008 Edition, (Ed. Edward N. Zalta), URL = <http://plato.stanford.edu/archives/win2008/entries/Space-time-holearg/>.

Nozick, Robert, *Philosophical Explanations*, Cambridge (Mass.): Harvard University Press, 1981.

Panti, Giovanni, ‘Denominator-preserving Maps’, 2011, URL = <arXiv:1103.5473v2>.

Petkov, Vesselin, ‘On the Reality of Minkowski Space’, *Foundations of Physics*. 37 (2007): 1499-1502.

Plantinga, Alvin, ‘Reason and Belief in God,’ in Faith *and Rationality*, eds. Alvin Plantinga and Nicholas. Wolterstorff, Notre Dame: University of Notre Dame Press 1984: 16-93.

Pollock John L., ‘Criteria and our knowledge of the material world’, *Philosophical Review*, 76 (1967): 28-62.

Pontrjagin, Lev, *Topological groups*, (trans Emma Lehmer), Princeton: Princeton University Press, 1939.

Quine, W.V.O., ‘On What There Is', *The Review of Metaphysics*, 2 (1948): 21-38.

Radin, Charles, ‘Symmetry and Tilings’, *Notices of the American Mathematical Society*, 42 (1995) 26-31.

Ratzsch, Del, "Teleological Arguments for God's Existence", *The Stanford Encyclopedia of Philosophy (Spring 2012 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2012/entries/teleological-arguments/>.

Reicenbach, Hans, *The Philosophy of Space and Time*, New York: Dover, 1957.

Rideout, David; *Dynamics of Causal Sets*; PhD Thesis (Syracuse University 2001); URL = <arXiv:gr-qc/0212064>

Robb, Alfred, A *theory of time and space*, Cambridge: Cambridge University Press, 1914

Roeper, Peter, 'Region-Based Topology', *Journal of Philosophical Logic* 26 (1997): 251-309.

Rowland, Todd. 'Fiber Bundle', *MathWorld*, a Wolfram Web Resource, created by Eric W. Weisstein, 2009, URL = <http://mathworld.wolfram.com.FiberBundle.html>

Schlesinger, George, 'It Is False That Overnight Everything Has Doubled in Size', *Philosophical Studies,* 15 (1964): 65-71.

Schmidt Hans-Jürgen*,* 'On the de Sitter Space-Time - the Geometric Foundation of Inflationary Cosmology' *Fortschritte Der Physik*, 41 (1993), 179-199.

Schwarz, F, 'On Discrete Subgroups of the Lorentz Group', *Lettere Al Nuevo Cimento,* 1976.

Sider, Theodore, 2007,'Parthood', *Philosophical Review* 116 (2007): 51-91.

Simons, Peter, *Parts: a study in ontology*, Oxford: Clarendon Press, 1987.

———, 'Extended Simples: A Third Way between Atoms and Gunk', *The Monist* 87 (2004): 371-384.

Solère, Jean-Luc, 'Scotus Geometres, The longevity of Duns Scotus's geometric arguments against indivisibilism', 2010, URL = <https://www2.bc.edu/~solere/docs/PAPERS/Solere_Scotus%20Geometres%202.pdf>.

Solovay, Robert, 'A model of set-theory in which every set of reals is Lebesgue measurable', *Annals of Mathematics.* Second Series 92 (1970): 1-56.

Sorensen, Roy, Seeing *Dark Things: The Philosophy of Shadows*, New York: Oxford University Press, 2008.

Sorkin, Rafael, 'Causal Sets: Discrete Gravity' in *Proceedings of the Valdivia Summer School*, (eds. A. Gomberoff and D. Marolf) 2003.

Stone, Marshall, 'The Theory of Representations of Boolean Algebras,' *Transactions of the American Mathematical Society*, 40 (1936): 37-111.

Suppes, Patrick and Zinnes, Joseph, 'Basic Measurement Theory', in R. D. Luce, R. R. Bush, and E. H. Galanter (eds.), *Handbook of Mathematical Psychology, Vol. 1*, New York: Wiley, 1963: 3-76.

Suppes, Patrick, 'Review of *Foundations of special relativity: Kinematic axioms for Minkowski Space-time* by J. W. Schutz', *Bulletin of the American Mathematical Society*, 82(1976): 459-462.

Surya, Sumati, 'Directions in Causal Set Quantum Gravity', 2011, URL = <arXiv:1103.6272v1>.

Szekeres, Peter, 'Discrete Space-time', in *Confronting the Infinite*, (eds. A.L. Carey, W.J. Ellis, P.A. Pearce and A.W. Thomas), Singapore: World Scientific, 1995: 293-303.

Tarski, Alfred, 'Foundations of the Geometry of Solids', in J. H. Woodger (translator) A. *Tarski, Logics, Semantics, Metamathematics, Papers from 1923 to 1938*, Oxford: Clarendon, 1956.

Thiemann, Thomas, 'Lectures on Loop Quantum Gravity', *Lecture Notes in Physics,* 631 (2003): 41-135.

Van Bendegem, Jean-Paul, 'Zeno's Paradoxes and the Weyl Tile Argument', *Philosophy of Science*, 54 (1987): 295-302.

———, 'Ross' Paradox Is an Impossible Super-Task', *The British Journal for the Philosophy of Science*, 45 (1994): 743-748.

———, 'Finitism in Geometry', *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2009/ entries/geometry-finitism/>.

Varzi, Achille, 'Mereological Commitments', *Dialectica* 54 (2000): 283-305

———, 'Mereology', *The Stanford Encyclopedia of Philosophy* (Summer 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2009/entries/mereology/>.

Wagon, Stan, *The Banach-Tarski Paradox*, Cambridge: Cambridge University Press, 1985.

Weisstein, Eric W., 'Menger Sponge', from *MathWorld – A Wolfram Web Resource*, 2009, URL = <http://mathworld.wolfram.com/ MengerSponge.html>.

Weyl, Hermann, *Philosophy of Mathematics and Natural Sciences*, Princeton: Princeton University Press, 1949.

Whitehead, Alfred North, *Process and Reality*, New York: Social Science Book Store, 1929.

Wolfram, Stephen, *A New Kind of Science*, Champaign: Wolfram Media, 2002.

Yamabe, Hidehiko, 'On an arcwise connected subgroup of a Lie group', *Osaka Mathematical Journal* 2, (1950): 13-14.

Zimmerman, Dean, 'Temporary Intrinsics and Presentism' in (eds. Peter van Inwagen and Dean Zimmerman), *Metaphysics, the Big Questions*, Oxford: Blackwell, 1998: 206-220.

Index