Forthcoming in J. Ashbourn (ed.), HAPP 10th Anniversary Commemorative Volume, *Journal of Physics: Conference Series*, Institute of Physics Publishing.

Taking up Space: The Case of the Ether

Ioannis Votsis (Northeastern University London)

Abstract: In this paper, we obtain some methodological lessons in theory construction and modification (generally called 'theory choice'), using the ether as a test case. We focus on this posit both because it has a long history but also because it is associated with some spectacular theories and theorists. In view of the fact that the ether has been expunged from contemporary science, we ask why it survived as long as it did, whether its discussion was a waste of space, and where we do go from here. The last of these questions concerns not just the ether but also similar posits currently in circulation or those that will enter circulation in years to come.

1. Introduction

Are there any methodological lessons that can be drawn from the history of science? The aim of the current paper is to identify a handful of such lessons with the help of a case study. The case study concerns the history of the ether, as it is rich in twists and turns, and, more than anything, it is a history of a posit's utter doggedness. The plan of the paper is as follows. Section 2 offers a brief introduction to the history of the ether. Its purpose is not to exhaustively scrutinise the entire timeline, as even book-length treatments find it difficult to do that [1] [2] [3]. Rather, its purpose is to select some noteworthy accounts and details of the concept that will inform the methodological discussion that follows. Section 3 provides a rationale for the longevity of the ether and, as a result, offers the first lesson in theory construction and modification. Section 4 adds three more lessons, gleaned from pertinent historico-philosophical discussions, into the mix. The paper concludes with Section 5, where a concise summary of the main points is given.

2. A Brief History of the Ether: From Aristotle to Newton

In the ensuing paragraphs, we sketch out some interesting waypoints in the history of the ether. It's worth repeating that our aim is not to do justice to all the conceptions of the ether out there but rather to identify some salient features that will serve as grist to our mill in the discussion of theory construction and modification that follows.

We begin by stepping back into antiquity. The notion of the ether, also aether, makes its first appearance during this period. The term comes from the ancient Greek *αίϑήρ*, which translates roughly as upper, bright and/or purer air, but also *αἴϑω*, which translates roughly as to ignite, kindle, light, burn or shine. It was used to denote the sky, which was thought of as containing clean luminous air. One of the first thinkers to theorise about it was Heraclitus, who took the ether to be a pure form of fire that orders and animates the cosmos. Not much else is known about Heraclitus' views on the matter. Speculations were made by other philosophers, including Empedocles, Anaxagoras and Anaximenes, but none amounted to anything like a fully articulated theory. That honour befell Aristotle, who, insisted on the existence of a distinct fifth element, to be added to the other four: earth, fire, water and air (note: the idea that the universe is made up of four elements is credited to Anaximenes). This was special in Aristotle's theory as, unlike the other elements, it was not only thought of as unchanging but also as disposed towards eternal circular motion.

A small digression is in order to give some context to these ideas. For Aristotle, bodies have a natural tendency to move either in a straight line or in a circle depending on their composition. There was already a tradition in place, going back at least to the Pythagoreans but also found in Plato (Aristotle's teacher), to treat circles and spheres as perfect shapes. Among other reasons offered to prop up this belief was the geometrical realisation of their highly symmetric nature. There are, after all, an infinite number of ways a circle or sphere can be rotated in space without changing its appearance. Unsurprisingly then, circular motions were treated as perfect motions. Looking up at the heavens and observing that stars, both those that appear fixed but also those that wander (what we today call 'planets' plus the Sun and the Moon), follow circular paths could not but reinforce that belief. Indeed, Plato famously posed a puzzle to his students: What combinations of circular uniform motions would be required to precisely account for those movements of the wandering stars? To be precise, Plato asked his students to explain the *apparently* less than perfect movements of the planets, which included the phenomena of stations of retrogressions, with some combination of perfect circular motions. In so doing, he set the cosmological theorising agenda for the next two millennia. The result was several geocentric models of astronomy, of which Aristotle's was one of the best known. Borrowing the idea of concentric crystalline interconnected spheres from another of Plato's students, Eudoxus, Aristotle sought to explain the observed motion of the stars by placing them into and onto those spheres. The fixed stars were placed onto the outermost largest sphere, while the wandering stars were placed onto and into the inner spheres. Rotational motion to the spheres was mechanically conveyed via one or more prime movers, who caused all motion but who were themselves uncaused.

In Aristotelian cosmology, the world is made up of bodies without any vacuums. That's because a vacuum is thought of as nothing and, as such, it cannot exist. Bodies predominantly made of earth or water move downward towards the centre of the Earth and in a straight line. Those predominantly made of fire or air move upward away from the centre of the Earth and in a straight line. These imperfect motions were observed in the sublunar sphere, the sphere containing the imperfect Earth. Since the heavens, i.e. the region beyond the sublunar sphere that encompasses the Moon, the Sun, the wandering stars and the fixed stars, and its motions were perfect, they could not be constituted by any of those elements. Instead, a fifth element was posited: the ether. A perfect, unchanging and incorruptible substance that makes up the heavenly spheres and even the stars themselves. The names Aristotle used for the ether were 'the primary body' or the 'first element'. Despite some scholars claiming that he never used the term 'ether', he clearly identifies this additional element with the ether. For example, in *On the Heavens* [4], he asserts: "The name, too, of that body seems to have been handed down right to our own day from our distant ancestors who conceived of it in the fashion which we have been expressing... implying that the primary body is something else beyond earth, fire, air, and water, they gave the highest place the name of aether" (p. 992). And in *Meteorology* [4] he asserts: "We have already described the first element and its powers... This is an opinion we are not alone in holding: it appears to be an old belief and one which men have held in the past, for the word 'ether' has long been used to denote that element" (p. 1220). Aristotle's ether later came to be called *quinta essentia*, i.e. the fifth essence, quintessence or the firth element, by his followers.

Despite the repeated denial, by some scholars, of any continuity between ancient and modern conceptions of the ether [5], five core ideas that subsequent thinkers would associate with the ether are already present in Aristotle. These are likely to have influenced them, given the immense effect Aristotelianism had on the Middle Ages and beyond. The five ideas are: its distinctness, its ubiquity, its unique role in explicating at least certain types of motion, the mechanical nature of that explanation and, potentially, its subtlety. It is a distinct element in that it is unchanging and unlike anything we encounter on Earth. It is ubiquitous, as it is present virtually everywhere, extending from the Moon out to the edge of the cosmos. It is key to understanding the motion of celestial objects as it makes up the celestial spheres and even the stars themselves. That motion is explicated in a mechanical way, through action by contact initiated by the prime movers. Finally, it is presumably subtle, though this idea is rather less grounded than others. As Grant [6] notes "Aristotle nowhere says [that it is subtle], but it seems an implication of his ordering of the four elements – namely, earth, water, air, and fire – which, as we move up from the earth, become… more subtle. Since the celestial ether extends beyond fire, it should exceed the latter in… subtlety" (p. 172, 62f).

A whole raft of thinkers, from antiquity to the renaissance, have ruminated on the nature of the ether, though their discussions are often no more than a commentary on Aristotle's conception. They include Plotinus, Simplicius and Francis Bacon. It's not until Descartes comes along that we get a more developed and, in some respects, provocative conception of the ether as well as its place in the cosmos. As is well known, Descartes is the father of modern mechanical philosophy, which shifted explanations away from occult posits and forces, so popular in medieval times, to the more mundane explanations of material particles that interacted by pressure and impact. In the posthumously published *Le Monde* [7] (and to a lesser extent in other works like the *Principia Philosophiae* [8]), Descartes sets out his conception of the ether in the clearest terms. It's important to note that he never actually employs the term 'ether', but he does describe what is effectively ether. On his account material particles are infinitely divisible and perfectly inelastic (solid and incompressible) but have no other qualities (properties), besides being extended in space. Indeed, it is through such differences in the extension of bodies, i.e. the "motion, size, shape, and the arrangement of their parts" (p. 26), that all other qualities can be explicated [9]. Moreover, there are three 'fundamental' types of particles, distinguished only by their shape, size and motion. These are: earth, air and fire. (Note: In the earlier *Discours de la Methode* [10], particularly in the essays *La Dioptrique* and *Les Météores*, Descartes fails to distinguish between air and fire.). Earth particles are the largest, possess irregular shapes, and move the slowest. Fire particles, which are the smallest of the three, have variable shapes and move the fastest. Finally, air particles – to be clear, not the ordinary air which we breathe – are spherical and have middling size and speed. What comes to be known as 'Cartesian ether' consists of air particles.

Once again, a little context is helpful in trying to understand his conception of the ether. Descartes marries his elemental view of the world with his vortex theory of planetary motion. On this theory, voids are rejected, as material particles of all three elements fill up the whole of space. In place of a vacuum between planets and stars, there is what we now call a 'Cartesian plenum': a continuous, somewhat chaotic, interlocking set of swirling crystalline (and therefore largely imperceptible) fluid vortices made up of air particles – see Figure 1. Sometimes, the Cartesian plenum is described as consisting of air particles punctuated by fire particles since the latter are always present in between the former – see explanation of light below. The Cartesian plenum is utilised to explain all sorts of phenomena, including those relating to light but also planetary motion. Light is explicated in a somewhat convoluted way. Fire particles fall towards the centre of these vortices and contribute towards star formation. The stars then generate an outward pressure on the surrounding air particles. These, in turn, push other air particles, and so on and so forth, until that push reaches and affects earth particles, including those that make up our very eyes. It is precisely this rectilinear propagation of pressure through air particles that Descartes deems to be light. Otherwise put, light is a sort of vibration in the Cartesian ether. The explication of planetary motions is a little less convoluted. Each vortex rotates around a star. In so doing, it carries celestial bodies like planets and comets with it. To be precise, such bodies exhibit a centrifugal tendency to escape the centre of the vortex but are often kept in a stable orbit by the counterbalancing action of the accumulated matter at the ring of the vortex, which itself is kept in place by the pressure exerted from the neighbouring vortices and their rings.

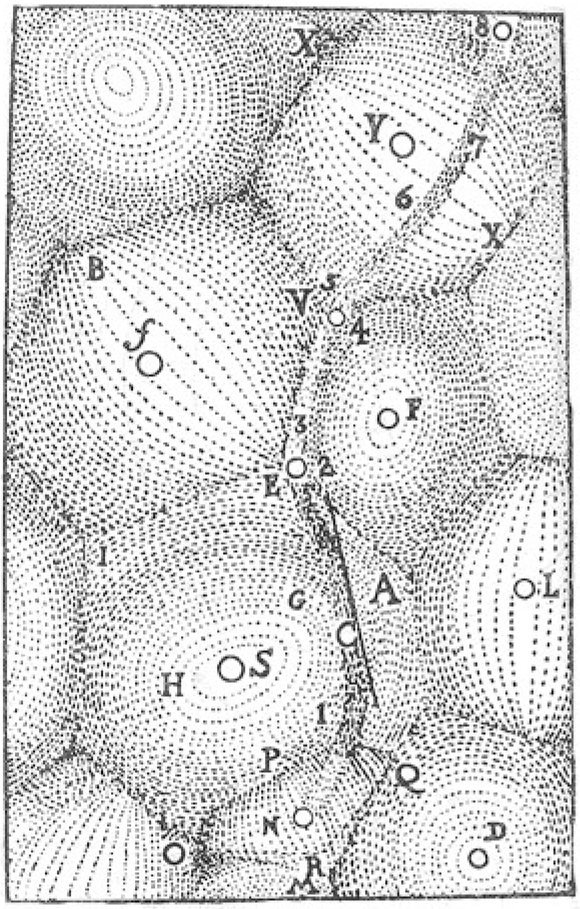

Figure 1. The Cartesian plenum, populated with various celestial objects, including the Sun S and other stars (e.g. F and f). Drawing taken from [8].

A few things are worth noting here. One is the shift towards a heliocentric model of the Solar system, though not the universe. Another is the shift towards a conception of the universe as being populated by multiple Sun-like stars. These two shifts are already of course associated with Copernicanism, which had been around for nearly a century. According to Vermij [11], Descartes and his followers were instrumental in helping make "the heliocentric theory into an acceptable and indeed dominant theory" (p. 140). Yet another thing to note [12] is that Descartes' vortex theory provides a prima facie plausible explanation as to why "planets lie in approximately the same plane" and why they "orbit the sun in the same direction [and in the same direction as the Sun's spin]" (p. 18). As the fluid vortices spin unidirectionally, the explanation went, so do the planets and any other celestial bodies. We now know, of course, that not all celestial objects lie in approximately the same plane or indeed orbit their parent star in the same direction including the direction of its spin. Despite the view's many flaws, it's still extraordinary how it foreshadows our modern conception of certain cosmogonical processes. It posits that the chaotic swirling motion of particles in the universe eventually gives rise to bodies like stars and planets. Moreover, it posits that bodies coalesce or rarefy via impacts and pressures. Similarly, the modern nebular disc model explains how collapsing interstellar clouds condense into protoplanetary discs with stars at their centre, ultimately giving rise to planets.

Cartesian cosmology is deeply ethereal since two of the three elements involved in its conception, fire and air, make up the plenum and help mechanically explain all the effects on the third element, earth. Despite his general hostility towards scholastic Aristotelian ideas like the distinction between form and matter, Descartes follows Aristotle in rejecting vacuums and in attributing to the ether versions of the five aforementioned ideas. The Cartesian ether is distinct (as it is unlike the elements of fire and earth), subtle (as it is imperceptible), ubiquitous (as it fills up interstitial space), crucial in explicating at least certain types of motion (e.g. the centrifugal motion of celestial bodies as they ride the vortices) and mechanical in nature (as it conveys action by impact and pressure). Having said this, he also goes further in postulating that the ether is a fluid in constant motion and that it is the carrier of light. As we saw earlier, the association with light is not entirely original as Heraclitus and other pre-Socratic thinkers linked the ether with pure fire or light. This link, as we shall soon see, will be pivotal in the theoretical construction of specifically light-bearing forms of the ether in the centuries that follow.

Although a number of other post-Cartesian thinkers wrote about the ether, e.g. Huygens, LeSage and Riemann, in what follows we shall only focus our attention on Newton; we return in the next section with a brief discussion of Fresnel and Maxwell. Just as Descartes rejects some Aristotelian ideas to break new ground, so does Newton reject or at least sideline several Cartesian ideas. Gone are the plenum, vortices, and the insistence that all motion must be explained by contact. In their stead, we get the vacuum of space, the gravitational attraction between masses and action at a distance. As is well known, Newton was not a fan of occult or spooky explanations of phenomena and was every bit as eager to explain away action at a distance with some mechanism. On that issue, at least, he was an aspiring mechanical philosopher, much like Descartes. Even so, Newton refrained from unequivocally committing himself to a mechanical hypothesis since it could not be supported by either argument or evidence, an attitude nicely encapsulated in his pronouncement *hypotheses non fingo* (I do not feign hypotheses). But that didn't stop him from discussing some such hypotheses, including the hypothesis that gravitational interactions were due to some disturbance in an underlying medium. As Jourdain [13] notes: "… he did not pretend to know what the cause of gravity might be, but it seemed to him incomprehensible that matter should act on other matter without the intervention of a medium" (p. 418).

It's worth dwelling a little on some of the details of these hypotheses, which, needless to say, involve the ether. Rosenfeld [14] notes that "the mechanisms [Newton] proposed for explaining gravitation [over the years] exhibit considerable and significant differences" (p. 29). Indeed, Newton appears to seriously entertain the notion of an ether at various stages of his life, both prior and after the publication of the *Principia* [15], though it hardly gets a mention in any of this book's editions. Discussion of the ether makes an appearance as early as 1675 in an essay titled 'An Hypothesis explaining the Properties of Light' [16]. It also appears elsewhere, as, for example, in his correspondence with Bentley [13]. The Opticks appears to be the last time it is brought up. We here focus on the latter period and, in particular, its culmination in the second edition of *Opticks* [17]. In this book, Newton suggests that gravitation could be explained as the result of a density gradient in an ethereal medium. He conceived of that medium as all-pervading and elastic, proposing that it was constituted by particles that are exceedingly small – smaller even than particles of air or light – and that can interact with material particles. In more detail, he suggested that such a medium would be rarer where bodies with mass reside but denser away from them. His own words in Query 21 are rather illuminating:

> Is not this medium much rarer within the dense bodies of the sun, stars, planets, and comets than in the empty celestial spaces between them? And in passing from them to great distances, does it not grow denser and denser perpetually and thereby cause the gravity of those great bodies toward one another and of their parts toward the bodies, every body endeavoring to go from the denser parts of the medium toward the rarer?... if the elastic force of this medium be exceeding[ly] great it may suffice to impel bodies from the denser parts of the medium towards the rarer with all that power which we call gravity.

The idea here is that there is a mutually repulsive force between the ether particles, whose cumulative effect in the denser regions of the medium is to push bodies with mass, presumably via another repulsive force, towards each other and towards the less densely packed parts of the medium.

Having said this, and as many scholars have noted, Newton supposed that even the denser parts of the ethereal medium could not have been really dense, as that would have created significant retarding forces on the motion of celestial bodies. As such, this created a problem, described aptly by Rosenfeld [14], who notes that Newton's postulation of the universal law of gravitation in the *Principia* meant that he "had thereby [been] forced… to the unwelcome conclusion that since the aether did not oppose any appreciable resistance to the passage of the celestial bodies, it must be a medium of such extremely low density that its role as the agent of gravitation was in jeopardy" (p. 32). One can't help but wonder whether this is why he never unconditionally committed to the ether.

Like Descartes, Newton does not restrict the scope of his explanation to gravitational phenomena but extends it to cover other phenomena such as light and heat. Unlike Descartes, he posits that light is made up of corpuscles (particles) that shoot out of sources such as stars at extremely high velocities. Since they pass through, and interact with, the ethereal medium, their default rectilinear paths may be redirected in ways that amount to reflections, refractions and diffractions. For example, in Query 29, Newton explained the rings of colour that appear when light is projected on thin glass plates by arguing that rays of light are put into "Fits of easy Reflexion and easy Transmission" through their interaction with the ether, creating vibrations that resemble waves in water. As for his take on heat, he argues, via a thought experiment, that its transmission could be explained by the presence of an ethereal medium. Newton imagines two inverted cylindrical vessels of glass, each of which contains a thermometer that is not touching its encapsulating vessel. Only one of the vessels has the air sucked out of it. He argues that if both vessels were moved from a cold to a warm environment and back again, heat would be transferred towards and then away from the thermometers almost as quickly in the one that is suspended in the vacuum as in the one that is not. From this Newton infers that there must be an ethereal medium that facilitates the transference of heat. In his own words: "Is not the Heat of the warm Room convey'd through the Vacuum by the Vibrations of a much subtiler Medium than Air, which after the Air was drawn out remained in the Vacuum? And is not this Medium the same with that Medium by which Light is refracted and reflected, and by whose Vibrations Light communicates Heat to Bodies" (Query 18).

To summarise Newton's stance on the ether, he is definitely enamoured by the concept but never seems to exhibit unequivocal commitment. Like Aristotle and Descartes, he attributes to it the five core ideas. The Newtonian ether is distinct (as it is made up of particles unlike those of ordinary matter), subtle (as its particles are miniscule and imperceptible), ubiquitous (as it penetrates all matter and can even be found in vacuums), crucial in explicating at least certain types of motion (e.g. gravitational, light and heat related motions) and mechanical in nature (as it meant to provide a mechanical explanation for action at distance phenomena). Unlike Aristotle and Descartes, he does not reject vacuums. Two things are worth noting here. First, vacuums are conceived of as being devoid of all ordinary matter but not ethereal matter. As such, it may be said that Newton's take is not so different to Aristotle and Descartes in that he does not permit a pure vacuum; at least not when the distances concerned exceed the smallest distance holding between two ethereal particles. Second, for the same reason, i.e. since ethereal particles never touch each other, his mechanical explanation of action at a distance phenomena merely postpones the problem to a smaller scale. Unlike Descartes, he postulates neither a tripartite distinction between elements, nor that the ether is a dense fluid swirling in vortices. Moreover, he does not try to explain the presumed unidirectionality of planetary orbits. Unlike Descartes, he argues that planetary motions can be explicated via a repulsive force that holds between ether particles and even between ether and ordinary particles but also argues that light is particulate, not a propagated pressure. Ultimately, Newton takes mechanical explanations to be ideal but is forced to endorse action at distance, which seems hard to shake off, even in the presence of an ethereal medium, which is meant to obviate it.

This concludes our brief history of the ether. Although patchy in some respects, it highlights some of the important similarities and differences that existed between the various conceptions of the ether throughout the years. It is now time to turn our attention to a critical examination of what these similarities and differences can teach us about theory construction and modification.

3. The Ether's Longevity: A Lesson in the Value of Hidden Posits

This section picks up from where we left off in the historical timeline, as we follow conceptions about the ether in the works of Augustin-Jean Fresnel and James Clerk Maxwell. The objective here is not to get into as much detail about these conceptions as we did earlier but rather to note some developments that ultimately led to the downfall of the ether. We then ask the crucial question of why the ether survived for as long as it did. Our answer contains both mundane and provocative explanations that are not necessarily in competition with each other. We conclude this section with the first lesson in theory construction and modification, namely that there is value, and potentially great value, in putting forth hidden posits.

Newton's success in mathematically describing and unifying celestial and terrestrial phenomena with his laws of motions and gravitation led to an aura surrounding his methodology. It's not surprising then that in the hundred or so years after Newton's publication of the *Principia*, the popularity of the ether begins to wane. As Torretti [18] notes:

> By 1771 the enlightened founders of Encyclopedia Britannica thought it appropriate to explain 'ether' as 'the name of an imaginary fluid, supposed by several authors […] to be the cause […] of every phenomenon in nature'. Not without irony, Joseph Priestley extolled the 'fine scene' that ether afforded 'for ingenious speculation':
>
> 'Here the imagination may have full play, in conceiving of the manner in which an invisible agent produces an almost infinite variety of visible effects. As the agent is invisible, every philosopher is at liberty to make it whatever he pleases, and ascribe to it such properties and powers as are most convenient for his purpose.'

This led to the construction of theories about various phenomena, including electricity and magnetism, that eschewed the ether and favoured explanations based on action at a distance. Ampère's electrodymanics as well as various other theories are examples of this change in attitude.

Despite that slump, the ether regained momentum in the 19th century. This was partly due to the successful re-introduction of wave theories of light by Thomas Young and, particularly, Augustin-Jean Fresnel. On Fresnel's view, the ether was a purely luminiferous, i.e. light-bearing, medium. Light consists of vibrations conveyed through this all-pervading material medium that undulates to produce transverse waves. These were put to use to explain the hitherto puzzling phenomena of polarised light. The throwback to earlier concepts and ideas was so strong by the middle of the 19th century that James Clerk Maxwell had not only embraced the ether but even attempted to produce a vortex model [19] of his newly composed and highly successful theory of electromagnetism. The ether was by now assumed to be an absolute frame of reference.

Fresnel's, Maxwell's and the efforts of various others scientists to keep the ether afloat, however, came to naught, as the conceptual and empirical problems kept accumulating. Fresnel was already forced to postulate that the ether was partially dragged by the very bodies it penetrated in order to account for aberration phenomena (e.g. when starlight is displaced towards the direction of a moving observer). In more detail, influenced by Young, Frensel posited that any body with a refractive index n greater than a vacuum (where n = 1) would have greater aether density than that found in a vacuum. It was then hypothesised that this excess density would be pulled along by a moving body (e.g. a glass prism) in such a way so as to account for the said aberration. Although at first confirmed by the Fitzeau and similar experiments, partial ether dragging (but also total ether dragging) was eventually disconfirmed by numerous other experiments, most famously (but not yet definitively) starting with Michelson and Morley – see, for example, [20]. The final nail in the ether's coffin was driven by Albert Einstein, who articulated the special theory relativity and explicated stellar aberration phenomena without the need for an absolute frame of reference. (Note: I am, of course, glossing over various contributions by, among others, Fitzgerald, Larmor, Lorenz and Poincaré.).

At this point, it is worth asking why the ether survived as long as it did. There are two broad explanations that can be offered, and they are not necessarily in competition. That is, both may very well be in operation. First, this may be down to a general form of conservativism in grand theorising [21] [22]. Conservativism is an umbrella term here, as it covers a number of different factors, e.g. cognitive limitations, conceptual prejudices and biases, whether these be intentional or unintentional. Second, it may be down to the fact that, in some sense, the idea of postulating the ether was justified. Let us consider each of these cases in turn.

We begin, briefly, with the explanation of conservativism. No matter what form this conservatism takes, lack of imagination, biases, etc., it should be clear that none of these thinkers carried out their grand theorising in a vacuum. Aristotle works, at least partly, in the shadow of his teacher Plato as well as their predecessors, and that affects both the questions he is trying to answer, i.e. how to explain planetary motion with circular motions, and the ideas that are available to him, i.e. the ether as a pure substance that can help animate the cosmos. Similarly, Descartes, like many other thinkers at the time, is desperate to distinguish himself from the Aristotelian scholasticism that had dominated western thought for the better part of the last thousand years. He dismisses many occult forces and properties along the way but cannot completely escape their teachings. His mechanisation of the world is, after all, a project that started with the pre-Socratics, especially the Atomists, and left an indelible mark on Aristotle's conception of the heavens. His rejection of vacuums is similarly Aristotelian in origin. Finally, his endorsement of the ether is clearly continuous with the past. Newton is also preoccupied with the past as he seeks to discredit Cartesian ideas about physics, particularly the plenum. Even so, he also falls for the ether and nearly 'everything under the sun' related to it. He takes the ether to be light-bearing, and even proffers some rudimentary explanations of its involvement in the generation of heat.

The more interesting, and certainly more contentious point, is the idea that the ether's postulation was in some sense justified and that's why it survived as long as it did. Allow me to elaborate. The ether, much like similar concepts throughout history including the caloric and phlogiston, is a hidden posit. That is, it is posit that is not directly visible to our unaided sensory observations or directly detectable by instruments. At best, it is detectable only indirectly, through its effects on things we can observe and measure. Such posits have a genuine reason to appealed to in science. That's because it was clear, since ancient times, that not every part of the world is immediately accessible to our senses or instruments. Moreover, the usefulness of conjecturing such posits is that they can stand for the unknown cause of sets of phenomena that, for better or for worse, appear to cluster into a cohesive whole. Finally, having such a concept is the first step in guiding and refining the design of experiments, which are the vehicles that can ultimately get us to the truth, or at least help us inch forward.

It's worth dwelling a little on the claim that science needs hidden posits to stand for the unknown cause of phenomena. We can call such posits 'placeholders' because their exact details, though initially unspecified or underspecified, are ultimately replaceable and indeed often replaced. In its simplest form, a placeholder posit is one about which we cannot assert anything, other than that it is responsible for an array of phenomena. Since that assertion is hopelessly generic, an obligation immediately builds up to conjecture some details about it. One way to do so is by constructing a mechanically coherent, and therefore prima facie intuitively plausible, account. The more details we add to this account, the closer we get to something testable, each detail imposing much-needed constraints on the posit's behaviour. Another, often complementary, way to do so is by mathematically rendering that account into a model. That enables a precisification of its content and commitments, thereby raising its overall testability.

The history of the ether, as it is outlined above, exemplifies just such a process. Several celestial phenomena were known all throughout this period, including the diurnal motion of the stars as well as the stations and retrogressions of the planets. It was thus natural for theorists to suppose a placeholder posit, namely that *some thing or other* was responsible for that array of phenomena. They then had to come up with a concrete hypothesis about this posit and its workings. Given the otherworldliness and remoteness of the heavens, it was hypothesised that a posit unlike no other must fill that role. That, of course, was the ether, which was already thought to be associated with phenomena of light, burning and the heavens. The next step was to postulate a mechanism that helped explain how that posit leads to the aforesaid phenomena. Different theorists articulated different mechanisms. Aristotle assumed crystalline spheres rotating around the earth, Descartes had his plenum of swirling vortices and Newton his elastic solid medium with variable density. The final piece of the puzzle was mathematisation. Though Aristotle did not provide much by way of a detailed mathematical model, other theorists working in the same geocentric tradition did, including most famously Claudius Ptolemy and his followers. In like manner, Descartes and Newton, who were both accomplished mathematicians, sketched out their own mathematical models. At the end of the day, all of these models were inadequate but also needed continual adjustments, a tell-tale sign of a degenerating research programme [23]. But that is beside the point. The point we are trying to make here is that the ether played a useful role as a hidden placeholder posit, whose presence allowed the development of theory.

It's also worth dwelling a little on the claim that science needs hidden posits to guide and refine the design of experiments. To start off, posits, whether hidden or visible, are a prerequisite for experiments. There is, in fact, something paradoxical in trying to carry out experiments without posits, as experiments are meant to decide between such posits or properties thereof. Even when a posit is in place, that may still leave much to be desired. That's because experiments are not likely to possess any probative force where the posits (and hypotheses) concerned are ill-defined. Positively stated, experiments are at their most probing when they are designed to test well-defined posits (and hypotheses). Since hidden placeholder posits are often not well defined, this makes them unsuitable for experimental tests. Indeed, until such time as they become better defined, any experiments carried out are likely to be met with suspicion and incredulity. This has happened several times in the history of science. For example, in relation to phlogiston, the absence of details regarding its weight created space for its advocates to employ ad hoc manoeuvres. To be precise, despite the expected loss of phlogiston from metals during oxidation, the observed weight gain in those metals was met with the reply that phlogiston had negative weight by its proponents. Needless to say, their competitors were far from impressed.

In the case at issue, the ether played a positive role in guiding and refining the design of experiments that eventually led to its downfall. Having such a hidden posit, allowed thinkers like Aristotle, Descartes, Newton, and eventually Fresnel and Maxwell, to focus their efforts in search of a mechanism that would explain all the aforementioned phenomena. The proposals ebbed and flowed over the centuries (e.g. from crystalline spheres to a plenum to an ethereal medium of variable density), and sometimes even regressed to earlier ideas (e.g. the revisiting of vortices in Maxwell's work), but one constant was the inexorable push towards a well-honed account of the posit, making it ripe for experimental testing. As we indicated above, and among other things, this involved the idea of ether drag and its effects on the speed of light. The probative force of those experiments that started in the 19th century and continue to this very day [24] has thereby been all but guaranteed.

The lesson that can be adduced from the above discussion is the following:

Lesson 1: Hidden posits (like the ether) are valuable in the development of both theory and experiments, and, as such, they cannot, and should not, be jettisoned from science, at least not until they reach a certain level of maturity.

This is a lesson that may seem trivial to some, but it is certainly not one that has been firmly rooted in the scientific psyche, as demonstrated by the various calls to dismiss more recent hidden posits, such as strings [25], that have not been given sufficient time to mature.

4. Lessons from Historico-Philosophical Debates

In this section, we identify three other lessons in theory change, using the case of the ether as the point of departure. We focus on lessons that emerge in the context of historico- philosophical debates concerning theory construction and modification, particularly the scientific realism debate – for an overview of this debate see [22]. The three lessons are not meant to be exhaustive. They are perhaps the easiest lessons to build a strong case around, but they are certainly not the only ones.

The second lesson concerns the truth-content of our best, i.e. most empirically successful, theories. This includes theories like quantum mechanics and relativity (both in its special and general forms). It can be summarised as follows:

Lesson 2: The most empirically successful theories in our possession may be very far away from the whole truth.

The ‘whole truth’ here refers to the truth about both visible and hidden posits that populate the domain targeted by the theory. To the best of our knowledge, nobody in the various historico-philosophical debates claims that our best theories encode the whole truth about their respective domains. What there is less agreement on is whether those theories are close to the whole truth. Many, but not all [26] [27] [28] [29], scholars claim that our best theories are indeed close to the whole truth and thus deny Lesson 2. To evaluate this claim we first need to qualify what we mean by being ‘very far away from the whole truth’. There are, roughly, two ways in which theories may fall well short of this ideal: under-description and misdescription. First, a theory may severely under-describe the target domain. That is, it may be oblivious to the existence of objects, properties or relations in that domain. A good example are theories of physics prior to the discovery of the sub-atomic domain. These theories treated atoms as the most fundamental constituents of reality. As such, they missed out on a whole range of objects that are now discussed under the banner of the Standard Model of particle physics. This includes not only the various types of sub-atomic particles, i.e. quarks, leptons and bosons, but also two of the four fundamental interactions, namely the strong and weak nuclear forces. Second, a theory may make many or grievous mistakes about the target domain. That is, it may posit the existence of objects, properties or relations that do not exist. A good example, as we have already seen, seems to be the positing of the ether.

Cases of misdescription and under-description, though logically distinct, tend to interact in the real world. For, to misdescribe the target domain, means to postulate the wrong posits, properties or relations. These will obviously need to be replaced with the right ones. As such, a theory that misdescribes also under-describes its target domain in that it does not postulate the right posits, properties or relations. Similarly, if to under-describe the target domain means to fail in postulating certain posits, properties or relations, a misdescription is bound to ensue in that those posits, properties or relations cannot be employed to correctly describe the target domain. In sum, many scholars deny Lesson 2 by claiming that our theories are approximately true and hence that they cannot be radically misdescribing or severely under-describing their target domains.

There are various ways to motivate Lesson 2. We explore two such ways here. The first draws inspiration from the history of the ether. If we return to those early conceptions of the cosmos in antiquity, what we are struck by how little of the world was accessible to those thinkers. For example, Aristotle could see the diurnal motion of the fixed stars, but he could not detect the stellar parallax that would allow the realisation that those stars may be situated at different distances from the Earth. As a consequence, it was hard for him and his followers to think in completely novel ways that would advance their knowledge of the cosmos. The same story permeates the history of the ether, and, more generally, the history of science. Descartes could conceive of planets and comets going around numerous Sun-like stars, but he could not imagine that these orbits were the result of a mutual attraction between objects with mass. Newton could theorise that the same laws applied to terrestrial and celestial phenomena, but he could not suppose that the speed of light (in vacuo) was invariant for all observers, regardless of any motion. At each and every point, these thinkers and their followers may have thought they were close to the ultimate conception of the world (Aristotle's geocentrism, Descartes' centrifugal explanation of planetary orbits and Newton's absolute reference frame), only for such judgements to be eventually upturned. Going by what has happened before, there is good reason to think that the boundaries of knowledge will continue to be recast, at least in the foreseeable future.

The second way to motivate the Lesson 2 is by means of a hypothetical argument. Suppose we create a vast virtual environment bit-by-bit, and then let people explore and build maps of it. The result is a collection of maps that do in fact contain some features that correspond to the features of the virtual environment. Now, since we, qua creators, have full knowledge of this environment, we would also be able to judge the overall extent to which such maps are faithful. That is, we would be able to say whether a given map is fully correct, approximately correct or neither. But the cartographers themselves are not in the same position. That's because their map-making endeavours do not commence with a perfect and complete knowledge of the virtual environment. Having such knowledge would, of course, defeat the purpose of making a map. Indeed, the vaster the environment in relation to the time the cartographers are given to explore it, the less likely that any of them is going to be in a position to judge the overall extent to which their map is faithful.

A similar problem of faithfulness exists in the case of scientific knowledge. Like the cartographers in our toy example, scientists have not created the universe. They thus do not have the luxury of being able to compare scientific theories to a pre-existing theory that contains a perfect and complete description of the universe. And since the universe is, by all accounts, incredibly vast in relation to the time we have spent investigating it, neither scientists nor philosophers of science are likely to be in a position to judge the overall extent to which our theories misdescribe or under-describe it. As such, and given the abovementioned interactions between misdescription and under-description, we cannot claim that these theories are true or approximately true. In fact, we cannot claim the opposite, i.e. that these theories are false or approximately false, either.

The third lesson concerns the inner structure of theories. It can be encapsulated as follows:

Lesson 3: At least some parts of the most empirically successful theories in our possession must be replaced.

Unlike in the case of the previous lesson, there is perhaps perfect unanimity here. On the one hand, highly pessimistic scholars like Feyerabend [30] and Kuhn [31] are notorious for accepting, or at least making remarks that suggest, the wholesale replacement of our best scientific theories (and even paradigms) in the wake of scientific revolutions. Likewise, Stanford [32] and van Fraassen [33] insist that the parts of theories that make claims about unobservables are, as a matter of fact, entirely epistemically dispensable (though they are of course pragmatically indispensable). On the other hand, there are those who are more optimistic, like Psillos [34] and Worrall [35], and emphasise the piecemeal fashion in which false and idle posits of past theories get abandoned and ultimately replaced in successor theories – see Figure 2. The ether is, once again, a case in point. It is pointed out, for example, that both Fresnel's theory of light and Maxwell's theory of electromagnetism have been rendered ether-less, without loss of predictive success. As such, the ether is unnecessary for the production of that success, and, therefore, undeserving of any credit associate with it.

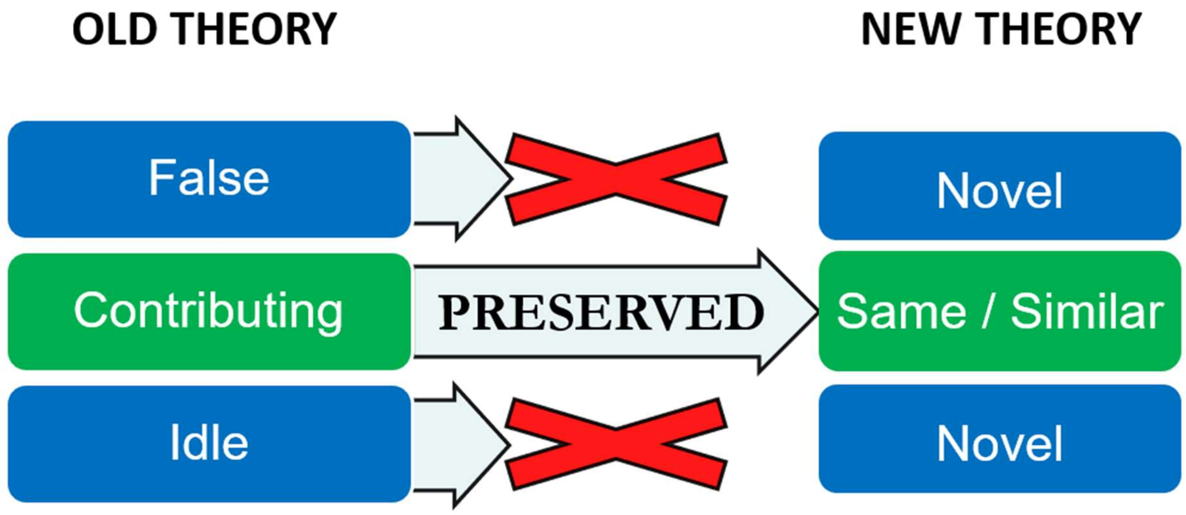

Figure 2. This represents the optimistic view of theory change. Those parts of a predecessor theory that contribute towards its empirical success are expected to be preserved across a scientific revolution. By contrast, idle or false parts (eventually) get discarded.

Going back to the key message, even if we accept that our theories are approximately true, that would still mean that at least some replacements must be made to make those theories true simpliciter. This lesson is hard to resist because no amount of interpretational gloss can overcome the fact that even our best scientific theories are incomplete and imperfect. The best of the best, quantum field theory and general relativity, are still at odds with each other, so we know that non-trivial modifications must be made for a unified conception of the universe. Modifications entail that some posits will need to be replaced. That's just another way of saying that our best theories are incomplete (since some correct posits are missing) and imperfect (since some current posits will be thrown out).

The same point can also be supported via historical considerations. All past successful theories contained at least some wrong posits. These were eventually removed either because they were ultimately epistemically dispensable (pessimist rationale) or because they had to make way for the right posits (optimist rationale). Take the shift from classical to relativistic physics as an example. Some of the key ideas of classical physics had to be given up. One such idea is the conservation of mass, i.e. that the total mass of an object or a collection of objects remains invariant even if we rearrange their parts. In relativistic physics, this idea is replaced with mass-energy conservation, as mass and energy are now thought to be interconvertible. This is one of many examples of how our best theories of the past have had at least some parts replaced – a lesson that the majority of optimist and pessimist scholars find compelling.

The fourth lesson also concerns the relation between successor and predecessor theories but addresses it in a more direct way. It can be formulated thus:

Lesson 4: Empirically successful successor theories must be such that they either straightforwardly reduce to or degenerate into the well-confirmed parts of their empirically successful predecessors.

This lesson also goes by the name of the 'generalised correspondence principle' and can be traced back to the 17th century. Fadner [36] argues that scientists, from Newton, Young and Clausius to Bohr, Heisenberg and Dirac, to name but a few, employed this principle in one form or another to ensure that theory development preserved the successes of theories past. In the philosophy of science, this principle gained prominence in the work of Post [37] and has since been developed and analysed in various directions by [38] [39] [40] [41]. For obvious reasons, it is widely endorsed by optimists, e.g. Ladyman [42], Redhead [41], Schurz [29] and Worrall [35]. Pessimists like Bueno [43] and van Fraassen [33] are also happy to endorse it so long as the well-confirmed parts of predecessor theories are circumscribed in a way that they deem aligned with their epistemic commitments, i.e. excluding unobservables. Both parties do so in recognition of the fact that the relation between new and old theories may be straightforward in some cases, e.g. deriving an equation from the new theory that also holds under the old one, or more complex in others, e.g. deriving an equation (again from the new theory) whose solutions approximate those of a corresponding equation in the old theory.

The rationale for endorsing Lesson 4 is simple. Science is an area of human activity where the expectation to produce tangible results is ever-present. This means that there is great pressure to incorporate any genuine empirical success a theory enjoys into subsequent theories. Were that not the case, successor theories would not be able to increase their usefulness. That is to say, successor theories would not be able to get closer to either the whole truth (which is what optimists want) or some severely restricted version of it (which is what some pessimists want). One celebrated example along these lines concerns Niels Bohr's use of the correspondence principle in the development of quantum physics. Classical physics could account for (and hence was genuinely empirically successful in relation to) the frequencies of some, but not all, atomic spectra. When Bohr constructed a mathematical model of these frequencies on the basis of quantum principles he made sure to not only account for the atomic spectra that classical physics had trouble with but also those that could be successfully predicted by it. In short, for those parts of atomic spectra where the classical model was approximately correct, the quantum model would itself be approximately identical to its predecessor. Thus, even though the newly-formed quantum principles contradicted some of the assumptions made by the classical principles, e.g. the former asserts (against the latter) that electrons can only occupy a finite number of discrete energy states, the two sides were still in approximate agreement over a range of other spectra.

4. Conclusion

In this paper, we explored some influential conceptions of the ether, spanning a timeline from antiquity to the 19th century. A case was made for the continuity of several core ideas (distinctness, subtlety, ubiquity, explicating motions, mechanical nature) throughout the ether's history, and despite several rather fundamental changes in the ways in which it was put to use. We then asked the question why the ether endured for as long as it did. The answer offered, involved, among other things, the claim that the ether was a hidden posit that played a key role in the evolution of the respective scientific domains. That is, the theoretical preoccupation with the ether was not a waste of space. Out of this answer, we distilled a general lesson, namely that because hidden posits provide invaluable assistance in the development of theory and experiment, they cannot, but also should not, be discarded as eagerly as it is sometimes suggested. This lesson was followed by three others, themselves distilled from historico-philosophical debates concerning theory construction and modification. The second lesson advanced the claim that even our current, most empirically successful, theories may not be close to the whole truth. The third lesson put forth the claim that at least some parts of such theories ultimately end up getting replaced. Finally, the fourth lesson proposed the claim that empirically successful successor theories are designed to be continuous with any well-confirmed parts possessed by their predecessors. We sincerely hope that when taken together, these lessons can help illuminate the difficult task of constructing and modifying theories that awaits the theorists of the future.

References:

---

[1] Grant, E. (1981). Much ado about nothing: Theories of space and vacuum from the middle ages to the scientific revolution. Cambridge University Press.
[2] Schaffner, K. F. (1972). Nineteenth-Century Aether Theories: The Commonwealth and International Library: Selected Readings in Physics. Pergamon.
[3] Whittaker, E.T. (1910). A History of the Theories of Aether and Electricity. From the Age of Descartes to the Close of the Nineteenth Century. Longmans, Green and Co.
[4] Barnes, J. (1984). The complete works of Aristotle (Vols. 1 and 2). Princeton: Princeton University Press.
[5] Watling, J. (1983). Review of Conceptions of Ether: Studies in the History of Ether Theories 1740-1900, Mind 92(367): 467-470.
[6] Grant, E. (1987). Celestial Orbs in the Latin Middle Ages. Isis, 78(2), 153-173.
[7] Descartes, R. ([1664] 1909). Le Monde. Oeuvres de Descartes. Charles Adam and Paul Tannery (eds.). Paris: Vrin.
[8] Descartes, R. ([1644] 1983). Principia Philosophiae, V. R. Miller and R. P. Miller (trans.), Dordrecht: Kluwer Academic Publishers.
[9] Descartes, R. ([1664] 1909). Le Monde. Oeuvres de Descartes. Charles Adam and Paul Tannery (eds.). Paris: Vrin.
[10] Descartes, R. ([1637] 2006). Discours de la Methode. OUP Oxford.
[11] Vermij, R. (2002). The Reception of the New Astronomy in the Dutch Republic, 1575-1750 (Amsterdam, Royal Netherlands Academy of Arts and Sciences).
[12] McAllister (2007). Introduction. Dilemmas in Science: What, Why, and How, in Knowledge in Ferment, Leiden: Leiden University Press.
[13] Jourdain, P. E. (1915). Newton's Hypotheses of Ether and of Gravitation from 1693 to 1726. The Monist, 418-440.
[14] Rosenfeld, L. (1969). Newton's views on aether and gravitation. Archive for history of exact sciences, 6(1), 29-37.
[15] Newton, I. ([1687] 1999). The Principia: mathematical principles of natural philosophy. Univ of California Press.
[16] Newton, I. (1757). An hypothesis explaining the properties of light. The History of the Royal Society, 3(1675), 247-305. This was written in 1675 but published posthumously.
[17] Newton, I. (1717). Opticks: or, a treatise of the reflections, refractions, inflections and colours of light. The third edition. London: William and John Innys.
[18] Torretti, R. (2007). Getting rid of the Ether. Could Physics have achieved it sooner, with better assistance from Philosophy?. THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia, 22(3), 353-374, p. 357.
[19] Maxwell, J. C. (1861/62). ‘On physical lines of force’. Philosophical Magazine, 21: 161-175, 281-291, 338-345; 23: 12-24, 85-95.
[20] Holton, G. (1969). Einstein, Michelson, and the “crucial” experiment. Isis, 60(2), 132-197.
[21] Chang, H. (2003). Preservative Realism and Its Discontents: Revisiting Caloric. Philosophy of Science 70 (5): 902–12.
[22] Votsis, I. (2011). The Prospective Stance in Realism. Philosophy of Science 78(5): 1223-1234.
[23] Lakatos, I. (1968). Criticism and the methodology of scientific research programmes. In Proceedings of the Aristotelian society (Vol. 69, pp. 149-186). Aristotelian Society, Wiley.
[24] Swenson, L. S. (1972). The ethereal aether: a history of the Michelson-Morley-Miller aether-drift experiments, 1880-1930. University of Texas Press.
[25] Krauss, L. M. (2005). Hiding in the mirror. Viking Press.
[26] Kuipers, T. (2000). From Instrumentalism to Constructive Realism, Dordrecht: D. Reidel.

[27] Niiniluoto, I. (1999). Critical Scientific Realism. Oxford: Oxford University Press.
[28] Leplin, J. (1997) A Novel Defense of Scientific Realism, New York: Oxford University Press.
[29] Schurz, G. (2009). When empirical success implies theoretical reference: a structural correspondence theorem. British Journal for the Philosophy of Science. doi:10.1093/bjps/axn049.
[30] Feyerabend, P. (1974). Against Method, Verso.
[31] Kuhn, T. (1962). The Structure of Scientific Revolutions, Chicago: University of Chicago Press.
[32] Stanford, K. (2007). Exceeding our grasp: Science, history, and the problem of unconceived alternatives. Oxford: Oxford University Press.
[33] van Fraassen, B. (2006). Structure: its shadow and substance. British Journal for the Philosophy of Science, 57(2), 275–307.
[34] Psillos, S. (1999) How Science Tracks Truth, London: Routledge.
[35] Worrall, J. (1989). Structural realism: The best of both worlds? In D. Papineau (Ed.), The philosophy of science. Oxford: Oxford University Press. 1996.
[36] Fadner, W.L. (1985). 'Theoretical Support for the Generalized Correspondence Principle', American Journal of Physics, 53, pp. 829-838.
[37] Post, H.R. (1971) 'Correspondence, Invariance and Heuristics', Studies in History and Philosophy of Science, 2, pp. 213-255.
[38] Batterman, R.W. (2002) The Devil in the Details, Oxford: Oxford University Press.
[39] Bokulich, A. (2008) Reexamining the Quantum-Classical Relation: Beyond Reductionism and Pluralism, Cambridge: Cambridge University Press.
[40] Radder, H. (1991) ‘Heuristics and The Generalized Correspondence Principle’, British Journal for the Philosophy of Science, 42: 195–226.
[41] Redhead, M. (2001) ‘The intelligibility of the universe’, In A. O’Hear (Ed.), Philosophy at the new millennium (pp. 73–90). Cambridge: Cambridge University Press.
[42] Ladyman, J. (1998). What is structural realism? Studies in History and Philosophy of Science, 29, 409–424.
[43] Bueno, O. (1997). Empirical adequacy: A partial structures approach. Studies in History and Philosophy of Science, 28A(4), 585–610.