Reconsidering the Fresnel-Maxwell Theory Shift: How the Realist can have her cake and EAT it too*

Juha Saatsi†
School of Philosophy, University of Leeds, LS2 9JT, UK

June 2, 2008

Abstract

This paper takes another look at a case study which has featured prominently in a variety of arguments for rival realist positions. After critically reviewing the previous commentaries of the theory shift that took place in the transition from Fresnel's ether to Maxwell's electromagnetic theory of optics, it will defend a slightly different reading of this historical case study. Central to this task is the notion of explanatory approximate truth, a concept which must be carefully analysed to begin with. With this notion properly understood, it will be finally argued, the popular Fresnel-Maxwell case study points towards a novel formulation of scientific realism.

If the realist is going to make his case for convergent epistemological realism, it seems that it will have to hinge on approximate truth, rather than reference.

Laudan (1981)

1 Introduction

Responding to the notorious argument from *pessimistic induction* (PMI) has arguably been the most central task in the recent realist agenda. Most realists have attacked the argument as presented by Laudan (1981), for example, by attempting to directly undermine its rather pessimistic premise that there are a significant number of past theories which were successful

*Forthcoming in *Studies in History and Philosophy of Science*.
†I want to thank Steven French and an anonymous referee for helpful comments and Joseph Melia for stressing the importance of multiple realisability. The research for this paper was supported in part by the Finnish Cultural Foundation.

but cannot be considered to be approximately true in any reasonable sense. Although there is no apparent verdict to be found on what this significant number would be—i.e. how large the "inductive basis" of the pessimistic "induction" needs to be—it seems that many realists are bothered by the argument to the extent that they want to deal with *each and every* item on Laudan's pessimistic list in order to leave nothing for the anti-realist to work with.

Allegedly a particularly troublesome item on Laudan's list consists of the 19^{th}-century ether theories of optics, featuring both bona fide novel successful predictions and a prime example of a completely rejected key theoretical concept. The current consensus is that the realist wishing to deal with this "counter example" to the No-Miracles Argument (NMA) head-on needs to clarify her conception of approximate truth and/or of reference in order to deal with the intuition that the ether theories are false simpliciter. This intuition, shared by Laudan, is simply due to the fact that in our world there is no pervasive elastic mechanical medium the existence of which the successful predictions seem to have relied on—i.e. the theoretical term 'ether' seems prima facie non-referring. The challenge, in other words, is to delineate *some* theoretical content to accompany the undeniable cumulative continuity at the empirical level, enough to *explain* the success of science in a non-miraculous fashion. The challenge thus posed immediately presents itself in form of the following two questions: (1) What does it take to explain a particular success of science? (2) If appeal to *some* theoretical content is indeed required as the explanans (as the realist argues), then on exactly what *principled grounds* should this content be delineated? These questions are at the heart of this present paper.

A notable array of realists have taken pains at elaborating their position in the face of this challenge, producing a variety of responses to both Laudan as well as to each other. (Hardin and Rosenberg 1982; Kitcher 1993; Worrall 1989, 1994; Chakravartty 1998; Psillos 1999) This paper takes yet another look at a case study which has been particularly popular in this context: the theory shift from Fresnel's prediction of reflection/refraction amplitude relations for polarised light to their modern derivation from Maxwell's electrodynamics. This is a case well featured in recent realist dialectic between Worrall and Psillos about how we should properly respond to Laudan.

These two authors, with their respective commentaries on the theory shift, arrive at formulations of realism radically at odds with each other. Worrall's *epistemic structural realist* knows next to nothing about the nature of unobservable entities in themselves and is not bothered by abandonment of central theoretical terms like 'ether' in scientific revolutions, as long as the *formal structure* of the theory exhibits appropriate continuity in theory change. Psillos's more orthodox formulation of realism, on the other hand, allows evolving conception of the nature of the ether only to the extent that it allows him to argue to the prima facie surprising conclusion that 'ether' (*pace* Laudan) is not a non-referring term after all! The present paper takes the stand that both of these positions are ultimately too extreme to be tenable: Worrall's structuralism tends to collapse into tracking a trivial kind of continuity with no real explanatory power, whilst Psillos is inclined to find unwarranted levels of continuity over and above Worrall's structure. But there is no need for pessimism either, for a more natural realist position—at least as far as the key case study of the Fresnel-Maxwell theory-shift is concerned—is to be found somewhere between these two opposites.

Despite arriving at two positions so widely at variance, Psillos and Worrall both actually follow the same overall strategy to undermine the contribution of this particular item to Laudan's pessimistic premise. Namely, their shared objective is, broadly speaking, to arrive at a notion of approximate truth that would provide the realist with the necessary resources required to explain the success of the earlier theory in light of the present theory. Psillos dubs this strategy '*divide et impera'*, and explicates that

> ... it is enough to show that the theoretical laws and mechanisms which generated the successes of past theories have been retained in our current scientific image. I shall call this the 'divide et impera' move. (Psillos, 1999: 108)

The relevant notion involved is called *explanatory approximate truth* (EAT) in this paper, and a careful analysis of this is a prerequisite for fruitfully considering the proper implications for realism of the historical case study.

The paper hence begins by analysing the notion of explanatory approximate truth. After critically reviewing both Worrall's and Psillos's elaboration of this notion (§2.2 and §2.3), it goes on to reconsider the details of the case study itself (§3). The lesson to be learned from the Worrall-Psillos juxtaposition is that in order to avoid the triviality of Worrall's structuralist construal of explanatory approximate truth, the realist does not have to defend a (problematic) reference invariance formulation of EAT as Psillos would have it. Rather, the analysis of EAT exposes the possibility of elaborating a variant of this notion which is sensitive to the hierarchy of theoretical properties appealed to in scientific theorising. The kind of hierarchy referred to here is best explicated by pointing to the practice of reductive explanation in science (§4). It will be concluded that the details of the case study, properly understood, point towards a novel formulation of scientific realism (§5).

2 Explanatory Approximate Truth

Most of the recent commentators agree that the realist should counter PMI by implementing some form of *divide et impera* –strategy. (Worrall 1989;

Kitcher 1993; Chakravartty 1998; Psillos 1999) They disagree, however, on the exact form this strategy should take and—regarding Fresnel's theory in particular—whether or not 'ether' refers and whether this matters or not. This section first gives an account in the abstract of the requirements placed on this strategy (§2.1) and then critically reviews the previous proposals made in this regard (§2.2-4).

2.1 Explaining theoretical success of rejected theories

The basic realist intuition behind explanatory approximate truth is readily understood on the basis of the popular realist argument of No-Miracles. According to this abductive argument the best (or only) explanation for a theory's success is its approximate truth. As far as PMI threatens to pose a problem to this intuition, it is natural for the realist to try and respond by refining her notion of approximate truth. Worrall, for example, suggests that

> ...the realist needs to show that, from the point of view of the later theory, the fundamental claims of the earlier theory (in so far as they played integral roles in that theory's empirical success) were—though false—nonetheless in some clear sense "approximately correct". He needs to show that, from the point of view of the later theory, we can still explain the success enjoyed by the earlier one. (1994: 339)

The basic idea is hence to show that the success of past science, by and large, did not depend on what we now take to be fundamentally flawed theoretical claims, and coupled with the no-miracles argument such an understanding of past successes licences a realist belief in the theoretical elements involved in the explanation of these successes. This rough idea of responding to PMI by elaborating a workable account of EAT is subscribed to by many contemporary realists, albeit with diverging details.

Before attending to the form these elaborations take in Psillos and Worrall, the crucial notion of explanation should be clarified further in the abstract. Two questions arise immediately. First, since the spirit of EAT is to restrict the realist commitment to those and only those things which are used to explain the success of a theory, we need to ask how to characterize 'the theoretical laws and mechanisms responsible for success' to begin with. And this is clearly closely tied up with the question of what explaining the success of a past theory exactly amounts to.

To begin with the obvious, I take it to be indisputable that the realist really needs to provide a *realist* explanation of the success of past theories on pain of inconsistency: success is a success, whether of a past or present theory, and the best explanation simply ought to be the same for both.[1] So arguably the best explanation is that a past, rejected but successful theory is false yet approximately true. The chronological order of explanation here is from the present to the past: the success of a *past* theory is to be explained *in the light of* our current best theories, the latter assumed to be approximately true.[2] Given that the past theories, taken at face value, can be really wide of the mark in presenting the world as we now take it to be, the critics really are right in demanding an explication of the sense of approximate truth applicable here. Luckily for the realist, the anti-realist intuitions about approximate truth are gratuitously pessimistic—EAT suffices for the realist purposes.

The realist explanation of a theory's success is a logical one: valid arguments with true premises always lead to true conclusions.[3] Explaining the success of a past theory in this spirit now involves identifying truth-content in the theory which enables this form of logical explanation to be applied, without being compromised by the immanent falsehoods.

Explanation thus understood, let us initially characterise EAT broadly as "truth of a theory in those qualitative aspects directly reflecting the unobservable reality which can explain the theory's success". The generality of this depiction is a virtue since it allows that the character of EAT can be to a large extent specific to each particular novel success to be explained, rather than being given *tout court*. On the other hand, however, the characterisation above seems so weak as to deflate realism into a virtually empty promissory note! Are our realist commitments with respect to some presently entertained successful theory to be summarised merely as "truth in those aspects which will *turn out* to explain (in realist fashion) the theory's success from any later vantage point"? Let us call this position *minimal explanatory realism*. Although rather deflated, it is still a realist position by virtue of both preferring a realist explanation of success over miracles, and sticking its neck out vis-à-vis historical as well as future science. For all the PMI instances the realist claims to be able to identify elements of reality in the past theories which explain their success. But the minimalist requires initially nothing of these elements other than that they can fulfil the essential explanatory function, and hence they may in principle form a gerrymandered lot. The only principled constraint is for theories which

[1] Some have tried to argue their way around this widely shared assumption (cf. Lange 2002, Lewis 2001, and Saatsi forthcoming for a rejoinder).

[2] Although this may prima facie seem circular, this worry is unfounded. The reason is that the realist assumes that the present theory is closer to the truth, or at least more comprehensive than the past theory. Also, the task of extracting the exact truth content of theories falls on the scientist, not the philosopher. (cf. Psillos 1999: 113)

[3] Here we are *not* concerned with defending the important and contentious issue of explaining the success of a *present* theory in this way. Rather, we are only concerned with explaining the successes of past theories on the basis of the present ones, assuming that NMA regarding the latter makes sense.

themselves are used to explain the success of their predecessor(s): the success of the current must be explicable (in light of any future theory) in a way that is *compatible* with the explanation(s) offered by the present theory.[4]

The realist positions examined in the rest of this paper, my own proposed variant included, can be viewed as building on such minimal explanatory realism by fixing a principled framework which provides initial guidance for the explanatory endeavour. Instead of being potentially gerrymandered in form, the various instances of EAT are unified in order to pronounce more definitive realist commitments with respect to the current science. Fixing this framework is a matter of answering the initial question of how the potential success-fuelling theoretical constituents should be conceived and characterised to begin with. This paper has grown out of the conviction that this question is a non-trivial one and that a refined attitude towards it will be for the benefit of the realism debate. The remainder of this subsection serves to motivate this conviction which, it will turn out, is fulfilled by the case study considered later.

The realist faces the task of extracting truth content from past theories to explain their successes. Clearly there is nothing in this kind of logical explanation itself that dictates the nature of the truth content eligible: we only need to be able to derive a conclusion which constitutes the theory's (novel) success from the truth content, by way of a valid argument. This suggests that we really need to look at particular theoretical derivations in detail to discern the kind of content they depend on.

It is commonplace that theories typically speak of various unobservable entities, kinds of physical particulars endowed with qualitative properties and relations, interacting in various ways. This is usually our preferred way of conceptualising theoretical content. What is less well appreciated is the fact that our access to theoretical properties and relations is a complex one. Typical fundamental theoretical properties (e.g. charge, spin, magnetic field amplitude) are defined not only through some dynamical, causal force laws via which the entities instantiating them interact and are ultimately indirectly observed, but also through a mixture of theoretical principles of conservation and symmetry, for example. Acknowledging the complexity of a typical theoretical description gives gravity to the question of how best to characterise the success-fuelling theoretical constituents. In particular, it is not by any means obvious that successful derivations operate on theoretical propositions best conceived of in terms of interacting entities with causal properties. Are the explanatory ingredients approximately true descriptions of the kinds of unobservable objects the theory speaks of, or approximately true descriptions of the properties and the theoretical principles involved in

[4]Perhaps this constraint is actually stronger that it initially seems. Most current theories are needed to explain various theoretical success-stories of the past, and the mere requirement of explanatory consistency may fix the realist commitments of minimal explanatory realism surprisingly precisely.

the derivation of novel successes?

Different answers to this question can lead to alternative realist positions. The explanatory account of derivation congruent with "standard" realism has always been construed as approximate truth about unobservable objects and processes denoted by the central kind terms. From Psillos (1999), for example, something akin to the following can be distilled.[5] The referents of kind terms are seen (informally, not necessarily metaphysically) as bundles of properties, and EAT gets spelled out in terms of sets of properties: if a subset s of all properties S attributed to a posit E of a past theory T_P is found to be fuelling a novel derivation in T_P, in light of the current best theory T_C, and s is also a subset of essential properties S' attributed to a posit E' of T_C, then T_P is approximately true assuming that "E" can be seen as referring to E'. Explaining a derivation in these terms boils down to finding a continuous set of properties s that (from the current perspective) were necessary for the derivation. There are then two options in the *standard realist account* of EAT. Either some entity allegedly referred to in the past theory does not actually have *any* of the properties that play a part in producing a particular success taken as the explanandum—in which case the term in question is non-referring but unproblematically so—*or* this entity has these significant properties but was also incorrectly attributed other "misleading" properties in the past theory, and the term in question refers to whatever satisfies the correct description. To ensure that continuity in the success-fuelling properties entails also referential continuity an appropriate theory of reference of theoretical terms is required as a fundamental part of this framework. Implicit in this account of EAT is the idea that theories cannot be approximately true (in a sense worthy of realist attention) unless the central terms denote something out there in the actual world—where "centrality" is understood as having a success-fuelling referent.

But it is not prima facie clear that this standard account with its referring-kind-term construal of EAT is at all optimal in trying to understand why the logico-mathematical reasoning leading to certain novel success took place within a particular theoretical world view. In practice one attempts to understand why the reasoning followed in the derivation leads from one step to the next in a way that ultimately explains the end result in a non-miraculous manner. Applying the *divide et impera* at each step throughout the deriva-

[5]Psillos's standard formulation of realism is adequately summarised for the time being as follows: scientific theories describe unobservable entities, their properties and causal interactions / processes. Theoretical terms have putative factual reference (the semantic component of realism), successful theories are approximately true and 'entities posited by them, or, at any rate, entities very similar to those posited, inhabit the world' (the epistemic component of realism). This formulation seeks to keep EAT as close to the intuitive correspondence notion of approximate truth as possible. The above discussion on EAT should make it clear that the referential, semantic component of realism does not come automatically with the epistemic component, but depends on the way the latter is spelled out.

tion, we may find that the derivation of a mathematical relation suggested by the theoretical picture is best explained by *approximately true description of the properties and theoretical principles involved*, rather than of kinds of objects as conglomerates of these properties. That is, approximately true theories in this sense give a false characterisation of the right property, and only (some of) the true aspects of this false characterisation are essential for the derivation. The referential question of whether we have two different descriptions of the same entity instantiating these right properties, or descriptions of different entities, is a red herring here. Rather, the explanatory work would be done at the level of properties.

These initial general remarks are best explicated further via the actual case study undertaken in section §3. Before proceeding to this, however, we should review in more detail the elucidating views of Worrall and Psillos on EAT.

2.2 Worrall on Explanatory Approximate Truth

Worrall (1989, 1994) appeals to the Fresnel-Maxwell theory shift in his argument for a structuralist account of explanatory approximate truth. After defending the *divide et impera* intuition (as captured by the quote above) the argument proceeds by ruling out the standard account of EAT as appropriately capturing the case study in question.

> A natural assumption is that such an explanation requires a demonstration either that the parts of the earlier theory rejected by the later one were redundant or that no real "rejection" was involved (but only a "re-description"). However, in this particular historical case at least, the most straightforward and least revisionary account of the explanation ... fits neither of those patterns. (1994, 339)

In its stead, Worrall suggests that the most natural explanation of the success of Fresnel's theory is given in terms of a structuralist construal of EAT: that Fresnel 'misidentified the *nature* of light, but his theory nonetheless accurately described not just light's observable effects but also its *structure*'. (1994, 340)

Worrall is right in claiming that Fresnel should not be interpreted as having spoken about the electromagnetic field all along, or so it will be argued later on (against Psillos). But it is not clear that Worrall has given a fair run to the alternative possibility of demonstrating that 'the parts rejected were redundant'. The problem is that evaluating this option requires a careful consideration of how these "parts" should be construed in the first place, and Worrall provides no discussion of this matter.

For example, if we take the ether as a theoretically posited entity to be a rejected part of the theory, then indeed it is difficult to argue that this part was ("as a whole") redundant for the derivation of the success of Fresnel's theory. Such a demonstration, I must agree with Worrall, would have a serious air of "whiggishness" about it. But why would the identification of rejected parts and the ensuing demonstration that those parts were redundant operate at the level of objects and kinds—as the standard realist account would have it—as opposed to the level of properties these objects were taken to instantiate? Theoretical reference, of course, is to the unobservable entities and the existence of non-referring terms has been traditionally taken to constitute a strong indication of the failure of approximate truth. But Worrall, with his structuralist theory of approximate truth, already marks a departure from this tradition; the structural realist is not bothered about non-referring central terms, anyway, since the structuralist's explanatory endeavour takes place not at the deep level of entities, but at the level of "structure". Similarly, if we managed to explain the success of Fresnel's theory by demonstrating that those *properties* of the ether which fuelled the derivation of the novel empirical prediction were retained, whilst those rejected were redundant, then that presumably would amount to a legitimate realist elucidation of EAT despite 'ether' possibly turning out to be a non-referring term. But this would still be in the spirit of the first strategy that Worrall rejects! If this approach were to go through there would be no need to be radical and adopt a structuralist theory of EAT.[6]

Psillos (1995, 1999, 2001) has argued against the nature-structure distinction that Worrall commits to. Although I am far from convinced that there is no useful such distinction to be had, against Worrall this criticism bites since there is no proper explication of the central notion of structure to be found, but rather some serious ambiguity. Mostly Worrall speaks of 'formal' or 'mathematical' similarities in Fresnel's and Maxwell's theories. He says, for example, that

> ... disturbances in Maxwell's field do obey formally similar (mathematically identical) laws (1994: 340)

and that

> ... there is structural, mathematical continuity between the two theories. (loc.cit.)

[6] If we choose to ignore this option and opt for a more radical line of thought, as Worrall does, then there are problems to be faced with the suggestion that, in general, truth about formal, mathematical structure of a theory is truth enough for the realist, and that this construal of EAT adequately *explains* the success of at least Fresnel's theory, in particular. Worrall does not explicate in any detail this structuralist notion of approximate truth or the sense of explaining this notion fulfils—rather, he just relies on our intuition on this matter. I am afraid these intuitions do not bear a closer scrutiny. See Psillos (1999: 151ff) for criticism.

But there are also remarks of quite a different spirit to be found, such as the idea that although

> . . . Fresnel was as wrong as he could have been about what oscillates, he was right, not just about the optical phenomena, but right also that those phenomena depend on the oscillations of something or other at right angles to the light. (loc.cit.)

In the usual philosophical terminology a truth about spatiotemporal relations such as those expressed by "oscillation at right angles to the direction of propagation" is *not* a formal, mathematical truth. It is a truth about some *properties* instantiated where there is phenomenon of light. Oscillating, like "having a velocity", is a property that many different kinds of entities can instantiate; both water and the electromagnetic field can oscillate, and so can the distribution of colour intensity on a computer screen. Oscillating is a *higher-order spatiotemporal property* of a system, and it can be *multiply realised* by various lower-order properties (e.g. location of water molecules, electric and magnetic field amplitude, colour intensity).[7] If there is a decent sense in which such higher-order properties could be termed structural, it is not at all apparent but rather in serious need of explication. Furthermore, it is not clear which one of these divergent readings of structure-as-opposed-to-nature is meant to undertake the explanatory work of EAT. In the next section (§3) an explanation of the success of Fresnel's theory is put forward according to which this can be understood in terms of such multiply realisable spatiotemporal properties which are correctly circumscribed by Fresnel's theorising. Whether or not there is a decent sense in which this explanation could be called a structuralist one, the fact remains that Worrall's move from our knowledge of such properties to the knowledge of formal, mathematical structure is a non sequitur.

2.3 Psillos on Explanatory Approximate Truth

After criticising Worrall's attempt to put a structuralist spin on explanatory approximate truth, Psillos proceeds to salvage the standard account of EAT from Worrall's (and others') criticism, thereby doing away with the principal motivation to 'go structural' in the first place.

Psillos advocates the standard view through a set of detailed case studies. The case study of *dynamical optical ether theories,* in particular, 'suggests that the most general theory—in terms of Lagrangian dynamics and the satisfaction of the principle of the conservation of energy—which was the backbone of the research programme . . . has been retained'. (1999:

[7]In a broader sense the term is sometimes used without the non-spatiotemporal connotation, referring not to periodic change in spatial properties but in any quantifiable property, as in 'oscillating exchange rates', for example.

113) This, together with a suitably tailored theory of reference, provides the grounds for Psillos to argue that 'ether' can be taken to refer to the electromagnetic field.

Prior to presenting the case studies, Psillos does not much elaborate on the notion of EAT in general, apart from the obvious logical characterisation of indispensability of truth-like success fuelling constituent.

> Suppose that H together with another set of hypotheses H' (and some auxiliaries A) entail a prediction P. H indispensably contributes to the generation of P if H' and A alone cannot yield P and no other available hypothesis H^* which is consistent with H' and A can replace H without loss in the relevant derivation P. (1999, 110)

This admittedly makes good sense but does not give much of a handle on how the actual case studies should be conducted. Psillos, of course, aims to extract from the case studies a level of continuity required to defend his orthodox realist position. That is, Psillos aims to retain referential invariance for those posited *kinds*—like the luminiferous ether—which cannot be ruled outside of success-fuelling constituents without violating the standard account of EAT. In the next section we demonstrate that these case studies, at least as far as the central Fresnel-Maxwell example goes, require a more open-minded construal of the explanatory success-fuelling constituents. When looked at in closer detail this case actually does not conform to the mould offered by the standard realist. More specifically, issue will be taken with the reference invariance claim of the standard account of EAT—the claim that armed with a suitable causal-descriptive theory of reference we can plausibly take 'ether' to refer to the electromagnetic field.[8] Before entering the case study itself the source of the problem will be expressed in more general terms, below.

2.4 Explanatory approximate truth and success-fuelling properties

It is the contention of this paper that a more natural explanation of the success of Fresnel's theory can be had by adopting a more refined notion of success-fuelling theoretical constituent. In particular, it seems to be a mistake to view these as *propositions about unobservable objects and kinds* featuring in the theory, in such a way that an indispensable hypothesis must be linked to successful reference of the kind names featured in the hypothesis

[8]There is also a general worry about this kind of reference-insisting construal of EAT, raised by some interesting recent arguments to the conclusion that the realist is playing an illegitimate semantic game in trying to salvage her realism by tailoring theories of reference to ensure referential invariance. Cf. Bishop & Stich (1997), Bishop (2003) and Cruse (forthcoming).

in question. Rather, the realist should adopt a more open-minded characterisation of these constituents, and make the working/idle –distinction at the level of *properties* instantiated by these kinds, for those properties involved in the derivations.

Something along these lines has been proposed by Chakravartty (1998), also commenting on Worrall's structuralism and the Fresnel-Maxwell case study. Chakravartty's demarcation between *detection* and *auxiliary* properties provides a useful reference point regarding the present suggestion, and he, too, declares that 'distinguishing kinds of properties may in fact distinguish forms of realist commitment'. (394) The crucial distinction for Chakravartty is between those properties 'upon which the causal regularities of our detections depend, or in virtue of which these regularities are manifested' on one hand, and those 'associated with the object under consideration, but not essential (in the sense that we do not appeal to them) in establishing existence claims', on the other. (394-5) Properties of the latter kind—the auxiliary ones—then only 'supplement our descriptions, helping to fill out our conceptual pictures of objects under investigation', whereas 'only the former are tied to perceptual experience'. (395) The detection properties hence form the subject matter of realist commitment. The insightful observation that the explanatory identification of success fuelling constituents should operate directly at the level of properties involved in a theoretical derivation is certainly a move to the right direction. But the distinction between the auxiliary properties and the rest, as it stands, is only rough and ready and it is not quite clear how this differentiation at the property level is meant to interact with the object level talk. The rest of this paper is an attempt to flesh out the details of this basic idea, by looking in closer detail at the various distinctions that can be made regarding properties as the explanatory constituents of EAT.

For one thing, it seems that Chakravartty's analysis of detection properties does not probe deep enough into the explanatory structure. The characterisation of these properties as directly contributing into causal regularities, as opposed to purely metaphysical images of the entities thus contributing, provides a useful starting point, but does not exhaust the dimensions of EAT. A crucial facet of properties often ignored in this context is their hierarchical nature (for lack of a better word). Theoretically posited objects and kinds, described as bundles of first-order properties, do not *as such* fuel a successful derivation. These bundles come and go in radical theory change, and the realist is naturally led to consider which particular kind-defining properties are invariant over an otherwise radical theory shift. But it is not necessarily these properties *as such* that fuel the derivation either, for an instantiation of a property can in turn necessitate the instantiation of higher-level properties and *these* can be the explanatory elements in the final analysis, whilst the lower-level properties exhibit variance in theory change. For example, spatiotemporally distributed systems typically instantiate important spatiotemporal higher-level properties: a spatiotemporal distribution of lower-level properties fluctuating in a particular manner, for instance.

Once again, these (somewhat cryptic) preliminary remarks are best clarified in terms of revisiting the actual case study under debate. After that a general framework to accommodate the emerging picture is offered in terms of reductive explanations in science (§4).

3 Explaining the success of Fresnel's theory

The realist is challenged to explain how Fresnel was able to predict the intensity of reflected and refracted polarised light from seemingly false presuppositions. This section offers an explanation which draws solely on those premises of Fresnel's derivation which are not refuted by the modern understanding of the phenomenon. It is imperative to examine the theory-shift and the electromagnetic understanding of optics in some detail (§3.1), to gain a grip on the kind of continuity this explanation capitalises on (§3.2-3). The notion of EAT considered in the abstract above is finally applied to the case study in §3.4.

3.1 Optics of reflection and refraction from Fresnel to Maxwell

Augustin Fresnel's successful derivation of the reflection and refraction amplitudes for polarised light in the early 1820s stands out as a significant chapter in the development of optical ether theories. Fresnel's successful conception of light as a *transverse* oscillatory mechanism was considered as one of the most fundamental discoveries about the nature of the elastic ether, and it was a crucial test for any proposed ether model that it could reproduce *Fresnel's equations* (1823) for the amplitudes of reflected and refracted light. For the two components of the reflection amplitudes $A^{reflected}$, for instance, Fresnel derived the equations

$$A_{\perp}^{reflected} = \frac{-\sin(i-i')}{\sin(i+i')} A_{\perp}^{incident} \tag{1}$$

$$A_{\parallel}^{reflected} = \frac{-\tan(i-i')}{\tan(i+i')} A_{\parallel}^{incident} \tag{2}$$

> The terms here refer to: A's are amplitudes, with $A_{\parallel}$ and $A_{\perp}$ the components parallel and perpendicular to the plane of incidence (spanned by incident and reflected rays); i is the angle of incidence (as well as reflection) and i' is the angle of refracted light (cf. Figure 1 in §3.2). The amplitudes (the vector sum of components) are linked to observations so that $A^2 \propto$ intensity.

An ether theory that eventually managed to reproduce the laws of reflection and refraction was that of McCullagh's *rotational ether*, but this in fact turned out to be just part of Maxwell's theory in disguise, since the dynamical assumptions employed by McCullagh turned out to be unrealisable by a system of material, elastic ether, but satisfied exactly by the field of Maxwell's theory as shown by Fitzgerald in 1878 (cf. Stein, 1982). Maxwell's theory indeed produces equations formally equivalent to those of Fresnel's theory, as first shown by Lorentz in his Doctorate thesis (1875). This formal correspondence has been appealed to by Poincaré (1952) and Worrall (1989, 1994) in defence of their structural realist position.

The equations (1) and (2) of Fresnel's theory were derived via theoretical principles and heuristic analogies from what was known about elastic mechanics (e.g. conservation laws of energy and momentum in oscillations) and material wave motion, insisting all the time on purely transverse oscillatory mechanism (despite thus creating a severe disanalogy with the nature of mechanical wave motion known). The result is a quasi-mechanical theory of the ether which is not based on a set of exact, consistent principles and force functions, but rather on some fitting dynamical boundary conditions applied to a close geometrical analysis of the problem. It is the interpretation of these boundary conditions that renders Fresnel's construction specifically an ether theory. We will come back to this below in more detail.

In Maxwell's theory, on the other hand, the corresponding equations follow by imposing a natural continuity condition on a dielectric interface of reflection and refraction. There exist solutions to Maxwell's equations which are taken to correspond to waves which propagate at constant velocity $c = 1/\sqrt{\varepsilon_0\mu_0}$ in free space (regardless of reference frame) and with velocity $v = 1/\sqrt{\varepsilon\mu} < c$ in a dielectric (non-conducting) medium. The simplest such solutions represent plane waves, i.e. monochromatic light. Any electromagnetic field must satisfy all the four equations everywhere and reflection and refraction of light on an interface of two dielectric media (air and glass, say) is handled by fitting the solutions of Maxwell's equations in both domains together at the boundary. These boundary conditions follow from the requirement that the fields obey all Maxwell's equations everywhere. For example, *Faraday's law* for the electromagnetic induction can be written in integral form

$$\oint \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{S}$$

where S is a complete surface spanning the loop over which the line integral on the left-hand side is taken. For this integral identity to hold smoothly in the limit at which a loop intersecting the boundary approaches to a line lying at the boundary (so that the surface integral on the right-hand side vanishes) it must be the case that the components of the electric field tangential to the interface are continuous across it. Similarly we can derive continuity conditions for the tangential component of the magnetic field intensity $\mathbf{H} = {}^1\!/_{\mu}\mathbf{B}$, and for the normal components of both $\mathbf{B}$ and the electric displacement field $\mathbf{D} = \varepsilon \cdot \mathbf{E}$. Imposing these continuity conditions for plane wave solutions hitting the surface of a dielectric material eventually gives us the modern Fresnel's equations, where the amplitudes are, of course, electric field amplitudes. (cf. Jackson, 1999)

In order to properly assess the correspondence between the two theories it will prove useful to go through the equations involved in broad outline.[9] For example, for a plane wave

$$\mathbf{E} = \mathbf{E}_0 \mathrm{e}^{\mathrm{u}\mathbf{k}\cdot\mathbf{x}-\mathrm{i}\omega \mathrm{t}}$$

$$\mathbf{B} = \sqrt{\mu\varepsilon}\frac{\mathbf{k}\times\mathbf{E}}{k}$$

we get the boundary conditions

$$(\mathbf{E}_0^{incident} + \mathbf{E}_0^{reflected} - \mathbf{E}_0^{refracted}) \times \mathbf{n} = 0$$

$$\left[\frac{1}{\mu}(\mathbf{k}^{incident} \times \mathbf{E}_0^{incident} + \mathbf{k}^{reflected} \times \mathbf{E}_0^{reflected}) - \frac{1}{\mu'}(\mathbf{k}^{refracted} \times \mathbf{E}_0^{refracted})\right] \times \mathbf{n} = 0$$

for the tangential components of $\mathbf{E}$ and $\mathbf{H}$, where $\mathbf{n}$ is a unit vector normal to the interface. If the light is linearly polarised with $\mathbf{E}$ perpendicular to the plane of incidence, these conditions yield

$$E_0^{incident} + E_0^{reflected} - E_0^{refracted} = 0$$

$$\sqrt{\frac{\varepsilon}{\mu}}(E_0^{incident} - E_0^{reflected})\cos\ i - \sqrt{\frac{\varepsilon'}{\mu'}}E_0^{refracted}\cos i' = 0$$

These give as the relative amplitudes

$$\frac{E_0^{reflected}}{E_0^{incident}} = \frac{n\cos i - \frac{\mu}{\mu'}n'\cos i'}{n\cos i + \frac{\mu}{\mu'}n'\cos i'} \cong \frac{\sin(i-i')}{\sin(i+i')}$$

where we have taken $\frac{\mu}{\mu'} \cong 1$ (which holds for dielectrics and optical frequencies) and used Snell's law $n' = n\sin i/\sin i'$. Apart from the minus sign—resulting from a particular choice of orientation of $\mathbf{E}$ and $\mathbf{B}$—this is equivalent to Fresnel's result (1), assuming that the electric amplitude E_0 squared is proportional to intensity, which is the case (for each medium of propagation).

In Maxwell's theory there is also a continuity equation for a quantity that is interpreted as the energy of electromagnetic field. The flow of this quantity is represented by the so-called *Poynting vector* which depends on the electric and magnetic fields:

$$\mathbf{N} \ = \mathbf{E} \times \mathbf{H}$$

[9]Cf. Jackson (1999), for example, for more detail.

The interpretation of this vector as representing energy flow through unit area per unit time agrees with the idea that the energy contained in static electric and magnetic fields (in volume V) is

$$U = \frac{1}{2}\int_V (\mathbf{E}\cdot\mathbf{D} + \mathbf{B}\cdot\mathbf{H})d\tau$$

as well as with the continuity equation

$$\frac{\partial U}{\partial t} = -\int_S \mathrm{N} \times dS$$

For plane waves we get simply

$$u = \frac{\varepsilon}{2}|E_0|^2$$

as the time-averaged energy density. It should be finally noted—and this is relevant for our later assessment of Psillos's analysis of the Fresnel-Maxwell theory-shift—that this continuity equation for energy is not *as such* employed in the derivation of Fresnel's laws from Maxwell's equations (although it undeniably follows from these equations just as well).[10] This is in stark contrast to Fresnel's original theorising in which a continuity equation for a quantity he interpreted as kinetic energy plays a central role, as will be seen below.

3.2 Deriving Fresnel's equations

The modern derivation of Fresnel's equations was sketched in some detail above, but very little was said about the original theorising in question. The present objective is to explain, on the basis of our current best understanding of light phenomena, how Fresnel was able to derive his equations for the amplitudes of reflected polarised light. If such an endeavour manages to succeed by unveiling theoretical constituents of Fresnel's derivation which can be taken to directly reflect the unobservable reality, then there is a decent sense in which Fresnel's theoretical construction is approximately true. Many of the assumptions Fresnel made about the nature of light may have to be left outside of such explanation, which only appeals to those properties and principles that are not only of heuristic use but truly indispensable for the derivation. When it comes to distinguishing these elements in practise, I whole-heartedly agree with Chakravartty (1998) who suggests that

> ... we must turn to the equations with which we attempt to capture phenomenal regularities, and ask: what do these mathematical relations minimally demand. We must consider not

[10]There are also some interpretation difficulties associated with the interpretation of **N** as energy flow, cf. Lange (2002, 136ff).

> what possible metaphysical pictures are consistent with these equations, but rather what kinds of property attributions are essential to their satisfaction—i.e. to consider not what is possible, but what is required. (396)

To implement this idea of *minimal interpretation*, the following strategy is adopted: an abstracted reconstruction of Fresnel's derivation of the equation (1) is initially considered from premises which say as little as possible about the nature of light (below). It is then argued (§3.3) that the theoretical properties appealed to in this reconstructed derivation are also realised in Maxwell's theorising, as well as in a multitude of other possible theoretical constructions which all agree on the explanatory basis required for the derivation in a way that is compatible with the Maxwellian understanding sketched above. Finally, the notion of EAT is revisited in the light of this case study (§3.4).

Fresnel derived his equations from some rather elegant and simple premises.[11] The molecules of elastic ether were taken to have mass, so that in oscillation they obtain a mixture of kinetic and potential energy presumably very much like a harmonic oscillator. They also obtain some momentum. The key assumption is that the maximum velocity of the oscillating ether molecules is directly proportional to the amplitude of light, which in turn is proportional to the square root of intensity. Armed with this '*minimal mechanical assumption*' (Psillos 1999:158) Fresnel goes on to exploit a mechanical analogy of elastic collision (in which both momentum and energy are conserved) to derive the equations (1) and (2) above.

In his earlier attempt Fresnel derived (1) from the assumption of momentum conservation for *longitudinal* oscillations, but this result is mitigated by checking the results against the assumption of energy conservation. On his later attempt he took the ether oscillations to be *transverse* in character, and he derived the equation (1) (without the minus sign) from the principle of conservation of energy (or *force vive*). A further continuity equation (or boundary condition) requiring continuity in the components of oscillation parallel to the interface yields, together with the energy equation, eventually the equations (1) and (2).

Here the original derivation is not rehearsed verbatim; rather, we proceed by deriving the equation (1) from some truly minimal metaphysical premises. This derivation is nevertheless in the spirit of Fresnel: the mathematical relationships employed are the very same, and they are motivated by physical continuity principles abstracted from Fresnel's theorising. But it is also in the spirit of Lorentz's derivation of these equations from Maxwell's theory, in that it represents a kind of common core upon which different metaphysical interpretations are then stacked.

[11] Cf. Fresnel (1923). A useful secondary source is Buchwald (1989).

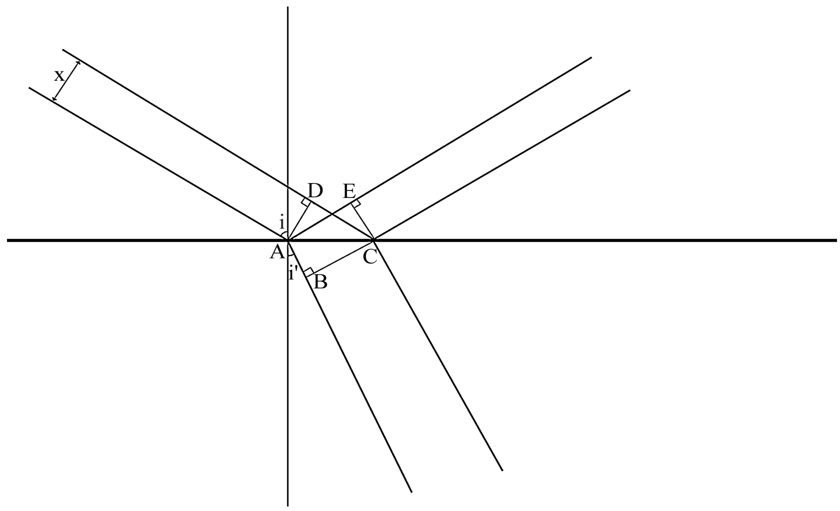

Figure 1. The geometrical reasoning behind Fresnel's derivation. As light front propagates the distance y=DC above the interface, an area of $a_0 = \Delta(\text{ADC})$ of the front reflects and refracts, producing areas $a_1 = \Delta(\text{AEC}) = a_0$ and $a_2 = \Delta(\text{ABC})$, respectively.

Fresnel's Energy Equation. Let's start with the following basic assumptions. The speed of light depends on the medium of propagation. Snell's law of refraction, $n_1 \sin i = n_2 \sin i'$, quantifies refraction phenomena in terms of experimentally determined refractive indices n_i. Let's attribute quantity Q to an area a (suppressing the third dimension for simplicity) swept by a beam (of width x) of light in some time interval Δt. Let Q depend not only of the area a, but also of some quantifiable "density property" q of light which is proportional to some attribute characterising the medium of propagation and to some attribute characterising the quality of light. For example, we might take refractive index n to characterise the medium and light intensity I to characterise the quality of light. One particularly simple way to assign Q is then simply

$$Q \propto a \cdot I \cdot n^2 = a \cdot q \tag{3}$$

Looking at the geometry of the situation (*Figure 1*)—given that the speed of light depends on the medium—it is natural to assume that a_0 is the area of the beam above the interface which through the process of partial reflection and partial refraction gives rise to the reflected area of $a_1 = a_0$ and the refracted area of a_2, respectively. Now, let's proceed by proposing a simple *continuity equation* for Q with respect to this process. Since the incident light "splits into two" it could be assigned quantity Q_0, which is then divided into Q_1 and Q_2, proportional to areas a_1 and a_2, respectively. Hence the continuity in Q, over time interval Δt, area-wise, is captured as

$$Q_0 = Q_1 + Q_2 \tag{4}$$

From the above assumptions alone one can derive, starting with this continuity equation, *Fresnel's energy equation*[12]

$$\sin i' \cdot \cos i \cdot (I_0 - I_1) = \sin i \cdot \cos i' \cdot I_2 \tag{5}$$

Fresnel, of course, derived this from the aforementioned mechanical principles. For him the conserved quantity Q was proportional to the density of the ether (molecules) and the vibration amplitude A of the ether (molecules) squared. Given the *minimal mechanical assumption* the continuity equation (4) clearly amounts to conservation of energy.

It is undeniable that Fresnel' mechanical analogies provided extremely helpful heuristics for the assumptions needed to arrive at (5), but even then the derivation is not a straightforward deduction. The properties of the elastic ether, as is well-known, present a somewhat curious mixture of analogies and disanalogies to the known properties of elastic solids. But even if heuristically crucial, it is obvious that these assumptions about the energy stored in some density of vibrating ether are not indispensable, for the above assumptions spelled out in much less specific terms led to the same equation! Here reference is made only to features of the continuous property of light (quantified in Q) such that Q is proportional to intensity I and the square of refractive index n^2. These features can be satisfied by energy, perhaps, or some other character of light, for all we know. All that is observed (or directly inferred) of light are its geometrical paths and intensity and state of polarisation. As far as light is said to carry energy or momentum, for example, this should be viewed only in a hypothetical or metaphorical sense, since these "theoretical properties" were in no way linked to observations at Fresnel's time, in the sense of not having the status of Chakravartty's *detection properties*. Nowadays we can meaningfully talk about the efficiency of solar powered machines, for example, but in the early 19^{th} century there was obviously no way to transform the alleged mechanical energy of light into some better understood form. For sure there were characteristics of light *suggesting* it carried some form of energy—the typical warmth of the sun, say—but the point is that these characteristics were not systematically and scientifically linked to the observed regularities understood as quantifiable forms of energy.

Fresnel's Amplitude Equations. So far there has been no mention of anything having to do with the transverse amplitude of light or its polarisation. The continuity equation thus far presented does not at all refer to the

[12] The derivation (which is very short) proceeds by first writing the ratio of the areas $a_0 (=a_1)$ and a_2 in terms of the angles of incidence i and refraction i', and then eliminating the refractive indices by using Snell's equation.

underlying mechanical character of light or what happens at the interface. To derive Fresnel's equation (1) one needs to appeal to the transverse vectorial property $\boldsymbol{A}$ ("amplitude") of light related to intensity so that intensity is proportional to $\boldsymbol{A}$ squared. But again we need to assume *very little* about this property of light!

We want to set up another continuity equation for a particular component of a vector quantifying this property, and in effect *all* we need in order to do this is to assume that polarised light is somehow "spread out" (i.e. not fully represented by 1-dimensional rays) asymmetrically so that this spatial asymmetry is quantified by a transverse vector. *Nothing* needs to be said about what the direction in question relates to in the underlying allegedly mechanical mode of propagation. We further assume that this vectorial quantity satisfies the principle of superposition; i.e. that it can be broken down to components each of which describes a possible state of light and that these can be added together to get the original vector.

We now impose another continuity condition for the components of amplitude $\boldsymbol{A}$: the components of $\boldsymbol{A}$ parallel to the interface $A_\perp$, that is, perpendicular to the plane of incidence, must satisfy ('the no-slip condition')

$$A_\perp^{incident} + A_\perp^{reflected} = A_\perp^{refracted} \tag{6}$$

That is, Fresnel demands that 'the horizontal velocity of the incident wave added to the horizontal velocity of the reflected wave must be equal to the horizontal velocity of the transmitted wave' (1923: 773, my translation). For light polarised in the plane of incidence we get, combining this with the energy equation, by elementary algebra and trigonometry[13]

$$A_\perp^{reflected} = \frac{-\sin(i - i')}{\sin(i + i')} A_\perp^{incident}$$

The equation for light polarised parallel to the plane of incidence is obtained similarly, although with some extra trigonometric manipulation to get the form (2).

Again, the heuristics used by Fresnel to come up with the 'no-slip condition' for the vector components $A_\perp$ may well have been indispensable heuristically at the time, although in this case our intuitions about mechanics surely do not suggest this boundary condition as anything obvious, and the striking disparity in the treatment of the "vertical" versus "horizontal" velocities is a notorious weakness in Fresnel's argument. The point is, however, that we can naturally describe this continuity principle in more abstract and less specific terms: to impose this fairly natural boundary conditionwe only need to subscribe to the asymmetric polarisation of light that

[13]The trick is to choose unit incident amplitude, for which we have $(1-v)^2 = (1-v)(1+v)$ on the left-hand side and $(1+v)^2$ on the right hand side of (5). To get (1) the trigonometric identities $\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$ are used.

can be represented vector-wise; what this asymmetry consists of, that we need not say.

3.3 Comparing Fresnel to Maxwell

Fresnel's original derivation relied heavily on crucial assumptions about the geometrical configuration of light rays: that the incident, reflected and refracted light lie on a plane, and that the angle of incidence equals the angle of reflection. It also incorporated Snell's law—relating the refractive indices and the angles of incidence and refraction—which together with the geometrical reasoning encapsulated in the Figure 1 ties together assumptions about the speed of light, refractive indices and the two angles. By way of contrast, in the modern derivation these "observed" facts about the behaviour of light all flow out from the field equations. But apart from the fact that the Maxwellian derivation is "deeper" in that way, what can be said of the correspondence between the two derivations? And in particular, can the latter be used to explain the former in the way earlier alluded to, in terms of multiply realisable success-fuelling properties?

First of all, it is clearly not the case that the modern derivation just follows the path of Fresnel's deduction with some substituted set of theoretical properties. For example, no use is made of the conservation of energy and there are not one but two boundary conditions imposed for the vectorial quantities $\mathbf{E}$ and $\mathbf{B}$ related through Maxwell's equations. These equations tie the two vector fields together in a way that describes a self-inducing orthogonally oscillating system of fields propagating at the speed $v = 1\big/\sqrt{\varepsilon\mu}$. We have seen that although Fresnel also spoke of *waves* of the ether, the real work in his derivation is done by the tangential boundary condition for $\boldsymbol{A}$ plugged into the continuity condition for Q,regardless of what these two quantities *exactly quantify*—as long as they relate to observations and the geometrical nature of the phenomena by the constraints $\boldsymbol{A}^2 \propto I$ and $Q \propto a \cdot I \cdot n^2$. So even though both Fresnel's and Maxwell's theory speak of the oscillatory nature of light, in our understanding of EAT this constitutes a kind of accidental correlation that is not to be appealed to as an explanatory correspondence.[14] On the other hand, for the realist appealing to explanatory approximate truth—for her to have her cake and eat it too—it is required that the quantities $\boldsymbol{A}$ and Q are there to be found in both theories.

The claim now is that these minimal explanatory properties could be realised by different kinds of systems. *Wave theory* of light represents one

[14]Explanatory approximate truth thus *diverges* here from the intuitive notion of approximate truth: whilst the latter is a matter of sufficient matching of the theoretical story and the world, *simpliciter*, the former takes into account the *explanatory value* of the matching. Fresnel's hypothesis of the oscillatory nature of light is not explanatory of his successful derivation in the required sense of being essential to it.

such possible realisation.[15] Understood in a wave theoretical framework (regardless of whether these waves are *further* described by Maxwell's equations and regardless of what these waves are understood to be waves *of*), the success of Fresnel's theory can be explained in a very straightforward fashion: we can point out that the boundary condition for $\boldsymbol{A}$ can be effectively arrived at by purely wave theoretical reasoning based on the principle of superposition. Furthermore, we do not need to employ the geometrical reasoning (Figure 1. above) involving the variable speed of light, or Snell's law for that matter, to use the conservation of energy condition in conjunction with the superposition principle. Hence the mere fact that in Maxwell's theory light is understood in terms of electromagnetic waves obeying the superposition principle—manifested as linearity of Maxwell's equations, of course—can be already used to explain why the theoretical assumptions of the minimal derivation lead to the right prediction! Fresnel himself, however, did not employ such explicitly wave theoretical superposition reasoning to arrive at the boundary condition, and therefore should not be considered as propounding a properly wave theoretical understanding along these lines.

The electromagnetic theory of light represents a further step in pinning down the properties that can realise the minimal description of Fresnel's derivation. We saw above that although the electromagnetic waves have, well, a wave nature too, and although Maxwell's equations do obey the linear superposition principle as well as energy conservation law, the derivation of Fresnel's results here is not just a matter of mimicking Fresnel by deriving one boundary condition and coupling it to energy conservation. Actually, we saw that both the wave nature of light and the energy conservation only come into play only in so far as these facts also undeniably follow from the Maxwell's equations, and therefore the respective derivations have in fact widely different theoretical bases. But we can nevertheless easily find a vectorial property $\boldsymbol{A}$ in the electromagnetic theory of light which (approximately) realises the boundary condition $A_{\perp}^{incident} + A_{\perp}^{reflected} = A_{\perp}^{refracted}$ and the conservation law $Q \propto a \cdot I \cdot n^2 = a \cdot q$ connected by $\boldsymbol{A}^2 \propto I$, which moreover relates $\boldsymbol{A}$ to the observable intensity of light I and its state of polarisation. Looking at the sketch of the modern derivation in §3.1 we see immediately that a natural candidate for such property is the electric field $\mathbf{E}$. This satisfies $(\mathbf{E}_0^{incident} + \mathbf{E}_0^{reflected} - \mathbf{E}_0^{refracted}) \times \mathbf{n} = 0$ and is related to the time averaged energy density u by $u = {}^{\varepsilon}\!/_{2}\,|E_0|^2$. The former is clearly equivalent to the boundary condition for $\boldsymbol{A}$, and the latter is directly proportional to q with good approximation—that is, when we put $n = \sqrt{\varepsilon\mu/\varepsilon_0\mu_0} \cong \sqrt{\varepsilon} \cdot \sqrt{1/\varepsilon_0}$, which holds for optical frequencies in typical dielectric matter for which $\mu/\mu_0 \cong 1$. By directly tapping into these properties of the electric vector field and the energy of electromagnetic field,

[15]Cf. Feyman 1964, Vol. 1:33–6 for a wave theoretical derivation of Fresnel's equations along the lines described here.

Fresnel's derivation managed to predict the correct observable intensity relations in a way that has thus proved to be much short of a miracle.

3.4 Fresnel and explanatory approximate truth

How does the minimal derivation of Fresnel's equations now accord with the various accounts of EAT reviewed in the previous section? Beginning with Worrall, it should be obvious that the minimal derivation appeals to crucial unobservable properties and theoretical principles besides formal, logico-mathematical structure, and that we appeal to these crucial theoretical constituents in our explanation of Fresnel's derivation from the modern perspective. In terms of EAT, to say it again, it is certainly not the case that Fresnel's theory is only true about the structure, as opposed to nature, of light.

Perhaps some find it difficult to appreciate the distinction between the minimal derivation and Worrall; perhaps the minimal derivation appears *too minimal*—so much so that it threatens to collapse into triviality. The worry might be that in the derivation of the energy equation (5), for example, the theoretical assumptions employed are so minimal that all the substantial theoretical explanatory content just evaporates! If all there is to the "density property" q is expressed in terms of proportionality to broadly speaking observable attributes of light—its intensity I and refractive index n—then where is the theoretical content proper to be found? Where is the explanatory causal mechanism, for example? It may appear, that is, that the explanatory derivation proposed is *so* minimal that the anti-realist could just as well buy into it!

But actually it is not the case that *nothing* theoretical is said about Q, for example. It is not a trivial theoretical assumption to make that there is a quantifiable attribute of light which is thus distributed across space. The continuity equation (4) for Q expresses a property of Q that is minimal—yes—but certainly not trivial. Similarly the component-wise required continuity in $\boldsymbol{A}$ expresses a higher-order property of light: whatever the asymmetry of polarised light amounts to, at the rock bottom, it satisfies the boundary condition $A_{\perp}^{incident} + A_{\perp}^{reflected} = A_{\perp}^{refracted}$. And the logically prior requirement that this asymmetry can be described in linear *vectorial* terms in the first place is already a theoretical assumption about the *nature* of light: satisfying the principle of superposition with respect to spatial components is not a matter of triviality or just a formal logico-mathematical fact about our description of a system, but rather best understood as a higher-order property.

Coming now to Psillos's account of EAT, differentiating the present proposal from his is a more subtle affair. In criticising Worrall's analysis, Psillos looks in some detail into Fresnel's derivation with the conclusion that the theoretical assumptions employed in conjunction with the minimal mechanical assumption were (1) *the principle of conservation of energy,* and (2) a *geometrical analysis* of the configuration of the light-rays in the two media. Hence, according to Psillos, there

> . . . is no sense in which Fresnel was 'just' right about the structure of light-propagation and wrong about the nature of light, unless of course one understands 'structure' so broadly as to include the principle of the conservation of energy and the theoretical mechanism of light-propagation. . . . At any rate, all of these properties of light-propagation were carried over in Maxwell's theory, even though Maxwell's theory dispenses for good with ethereal molecules. (1999: 159)

Small differences apart what we have said above seems to agree by and large with these sentiments. Indeed, one cannot but fully agree with Psillos when he says that 'we can clearly say that [Fresnel] was right about *some* of the fundamental properties of the light-waves, and wrong about some others'. (159)

Where the real disparity comes in is the set of conclusions drawn by Psillos regarding the most defensible form of realism *vis-à-vis* the case study in question. Psillos argues for *standard realism*, and part and parcel of that is the semantic, referential component: scientific theories are to be taken at face value and the central theoretical terms featuring in them have putative factual reference. Hence it is crucial for him to ensure that theoretical terms such as 'ether'—which is arguably a basic term in the theoretical part responsible for the successfulness of optical ether theories—turn out in one way or another to invariably refer! Thus he goes to great pains to argue that *ether referred to the electromagnetic field.* It is part of Psillos's realist project to show that once we adopt a particular causal-descriptive theory of reference in which referential continuity is guaranteed by 'substantial continuity in those properties which ground the causal role attributed the posited entities' (294)—so that a 'term which is employed to denote the posited entity is associated with a *core causal description* of the properties by virtue of which it plays its causal role vis-à-vis the set of phenomena'—then it turns out that "luminiferous ether" and "electromagnetic field" 'refer to the same entity precisely because their referents share the same core explanatory structure'. In other words: reference of a term is fixed by 'kind constitutive properties by virtue of which the entity denoted by the term is intended to play its causal role'.

Prima facie it seems that the idea of 'ether' and 'electromagnetic field' denoting the same class of entities is ludicrous—given what we know of the role of the ether in the transition from Galilean to Special Relativity—but it is good to spell out in some detail just how far-fetched this idea becomes once we realise how *minimal* the explanation of Fresnel's derivation actually is.[16] Given the foregoing derivation of Fresnel's equation, can we really take seriously the idea that 'the core causal description' of the elastic ether in Fresnel's theory, for example, just consists of spatial transverse vectorial asymmetry and a boundary condition with respect to that asymmetry in one component? Is *everything else* in the connotation of the elastic ether to be taken as merely heuristic and dumped into the heuristic models? Psillos wants to say something like that with respect to the mechanical ether models in general, but for him the core causal description of the ether has a significant dynamical component: 'the luminiferous ether was the repository of potential and kinetic energy during light-propagation'. But we have seen that, apart from heuristics, this component plays no role in explaining the derivation of Fresnel's law! Furthermore, that notion of mechanical energy had nothing to do with the notion of energy that is nowadays attributed theoretically to the electromagnetic fields.

Perhaps it is worth delving into this last point in a bit more detail. Whilst it is true that Fresnel appealed to a well-known energy conservation law of mechanics in his formulation of (4) and (5), it is not really the case that we can make sense of the 'energy of the ether' in Fresnel's framework as anything but an *auxiliary* property—to now borrow a useful piece of terminology from Chakravartty. When speaking of the energy of the electromagnetic field, on the other hand, we have at our disposal various ways to link it to observations and other forms of energy, by virtue of which we can elevate this property of the electromagnetic field to the *detection* -category. It is quite possible of course that an auxiliary property matures into a respectable detection one as science advances. Chakravartty is fully sympathetic with this, and indeed regards the auxiliary properties of theories, as far as his *semirealism* goes, 'not as substantive knowledge, but as methodological catalyst'. (1998: 404) A closely associated point was also made above in explicating the distinction between EAT and the intuitive notion of approximate truth: not all the conserved theoretical elements necessarily play an explanatory role and hence automatically gain the epistemic warrant that goes with it. (cf. footnote 14)

Moving now on to Chakravartty's account, he insists on a minimal interpretation of Fresnel's derivation much like I do.[17]

[16]da Costa and French (2003: 169-170) acutely press the point that in taking 'ether' to refer to the electromagnetic field one draws an illegitimate line between the kinematical and dynamical properties as opposed to mechanical properties of the ether, a manoeuvre which obscures the significance of the transition from classical to relativistic physics. They take this as a point in favour of a version of structural realism, as opposed to standard realism with its focus on entities—'if the mechanical properties are shunted off to models, as it were, in what sense can we still say that the scientist is still referring to the ether *as an entity?'*—but in my view this is going too far.

[17]Actually Chakravartty speaks of interpreting Fresnel's *equations* (1) and (2), but the real point of focus is, of course, the derivation of these.

> What, then, do Fresnel's equations require? They demand some kind of influence, propagated rectilinearly and resolvable into two components, oscillating at right angles to one another and to the direction of propagation. The property or properties of light in virtue of which such influences are realized are detection properties. (1998: 396)

Here some care should be exercised, however. First of all, the minimal derivation offered above does not even go as far as speculating about the oscillatory nature of light. Secondly, the properties which fuel the derivation are exactly the density property Q and the vectorial property $\boldsymbol{A}$, defined by (3), (4) and (6), respectively. The properties of light in virtue of which these are realized, on the other hand, are undetermined by the derivation. A third and related point has to do with the general characterisation of these *explanatory properties.* Chakravartty emphasises the role of causal relations in his demarcation of detection properties from the auxiliary ones, and he goes as far as claiming that 'all structures of interest may be accounted for in terms of causal relations which identify specific entities'. (p. 401) But this is surely too narrow a construal of the explanatory constituents, given the case study above. The two continuity equations supplying the crucial ingredients for successful prediction are not directly causal in any straightforward sense, for example. The question of how best to characterise these properties in general terms then, is finally taken up next.

4 Analysing reductive explanation in science

The preceding discussion has revolved around the two opening questions: (1) What does it take to explain a particular success of science? (2) If appeal to *some* theoretical content is indeed required as the explanans (as the realist argues), then on exactly what *principled grounds* should this content be delineated? These questions are interlinked and different responses result in a variety of realist positions. Answering the second question, in particular, determines a notion of explanatory approximate truth. In the foregoing it was argued that a more open-minded characterisation of the realist commitment is desirable in this context. In this penultimate section the allusive suggestions above that such characterisation should be given in terms of multiply realisable theoretical properties is elaborated by pointing to the practice of *reductive explanation.*

Scientific theories—regardless of whether or not they have produced novel successes—are required to fulfil an explanatory function: explaining facts at one level by appealing to facts at another is what science does (or one of the things, anyway). Whatever is minimally required to explain the successfulness of some theoretical assumptions and the ensuing logico-mathematical derivation of a novel prediction corresponds to whatever is minimally required to explain (in a sense clarified below) the phenomenon in question. And whatever goes beyond that minimal interpretation with respect to the former explanandum is equally beyond the minimal interpretation required with respect to the latter explanatory endeavour. Hence, although the focus of the present paper is primarily on the task of delineating a notion of EAT in the former sense, this can be approached by asking what is minimally required to explain some phenomenon scientifically and how all this conforms to the actual scientific practice which, the anti-realist reminds us, often proposes explanations wide of the mark.

The idea of *non-explanatory surplus content* is familiar enough from everyday explanations of all kinds. Consider explaining why a philosophy student Owen suddenly begins to express optimism about the prospects of conceptual analysis, by suggesting that he has recently read Frank Jackson's book *From metaphysics to ethics.* This suggestion may be false—and hence disqualify as an explanation as such—yet contain a significant seed of truth: perhaps Owen has recently read another book of Jackson's in which similar ideas were entertained. Or consider explaining why a crammed elevator refuses to move by suggesting that there is 50 kilograms of excess weight on board. This is likely to be strictly speaking false, and hence disqualified (strictly speaking) as an explanation, yet it may contain a significant seed of truth: perhaps there is indeed too much weight on board, but only some 44 kilograms, say. Such everyday examples strengthen the intuition that a false story does not have to be explanatorily empty. In both cases above there is a clear sense in which a *less specific* story would have captured what is true in the explanatory proposals, and thus would qualify as an explanation (although an equally specific and fully true explanation also exists, of course). One could have explained Owen's behaviour by the fact that he has recently read *one of the works* by Jackson in which conceptual analysis is defended, for example. It is also clear that such an explanation can be compatible with a multitude of more detailed stories which are incompatible with each other: perhaps Owen has only read his personal copy of the book, or perhaps he consulted a library copy, etc.

Outlining what is minimally required to explain a theory's success corresponds to extracting the minimal theoretical explanation of the phenomenon successfully predicted (or accommodated). This minimal scientific explanation is compatible with a multitude of stories about the lower-level facts. This can be understood through *multiple realisability of properties*: the explanatory ingredients are properties identified by their causal-nomological roles, and most (if not all) such properties are *higher-order* multiple realisable in the sense that these properties are instantiated by virtue of having some other lower-order property (or properties) meeting certain specifications, and the higher-order property does not uniquely fix the lower-order one(s). This thesis of multiple realisability is taken to be fully uncontroversial here. It is common place that many macro-realm properties are multiply realisable in this sense—e.g. 'being a pen' designates a property instantiated by many types of pen—and multiple realisation arguments form the locus of a central debate between the reductionists and non-reductionists in the philosophy of mind. But multiple realisation is clearly not confined to states of cognition, or the very high-level artefact functional kinds.[18] Some of the properties are multiply realised in the actual world, whereas others have alternative realisations only across possible worlds. In particular, there is a natural sense in which the properties involved in the minimalist explanation of Fresnel's successful prediction are unmistakably multiple realisable, even if only in a restricted modal sense.

To explicate the idea of multiple realisation further we might focus on the realisation relation itself. Adopting the following rough characterisation is adequate for the present purposes:

> A property P of an object o realises a property F of o if and only if (i) it is necessary that, if o instantiates P, then o instantiates F, and (ii) o's instantiating P in some metaphysical sense explains o's instantiating F—being P is one way in which a thing can be F. (Clapp, 2001: 112)

Thus, to iterate the standard example, the property of a fixed amount of gas being in a certain temperature T can be realised by some very discriminating microphysical property of the gas molecules having certain velocities and masses. Furthermore, having a temperature T necessarily follows from having this microphysical property, and having this microphysical property is only one of the ways in which the gas can have the temperature T as there are many other molecule configurations which equally correspond to the same mean kinetic energy. The sense in which a realising property 'metaphysically explains' the instantiation of the higher-order property amounts to the familiar notion of *supervenience*: the asymmetrical relation between an instantiation of a lower-order property necessitating the instantiation of a higher-order one, whilst the latter being compatible with a multitude of lower-order properties, is captured here by saying that temperature (of a gas) supervenes on its molecular configuration.[19] Corresponding to this

[18]Consider the higher-order property designated by 'being a primary colour', for example. This is realized by the properties of being red, being blue and being yellow. More generally, the relationship between some determinable property and the corresponding set of determinates can be viewed from this perspective; there are many ways of being (spatially) asymmetric, for example, or having the mass of nine grams. (Clapp, 2001; Yablo, 1992) More pertinently, the property designated by Q in our reconstruction of Fresnel's derivation, defined by its continuity over the process of reflection/refraction and direct proportionality to intensity, is multiply realisable in the modal sense of having epistemically possible realisations. Cf. footnote 22.

[19]The concept of supervenience allows many degrees, varying with the interpretation of the modality of necessitation between lower and higher-order properties. Cf. Chalmers (1996) for discussion.

metaphysical explanation there is an *epistemic reductive explanation* very much familiar from scientific (as well as everyday) practice.

In addition to various other modes of explanation operative in science much of the effort to understand the workings of nature can be analysed as reductive explanation of an inherent dispositional characterisation of some system of interest. To start with a humdrum illustration: a car engine is characterised by some typical causal powers of the kind ‘press the gas pedal and the car accelerates’, and so on. To explain the workings of such a system it can be broken down to subsystems the contribution of each one of which is then analysed with respect to the functioning (i.e. causal powers) of the engine as a whole. The workings of a subsystem can perhaps be further explained reductively in terms of yet smaller subsystems. In a similar manner, albeit with an increased level of abstraction, the phenomenon of light reflection/refraction, for instance, can be first analysed in dispositional terms and then reductively explained by reference to various lower-order properties upon which an analysandum disposition supervenes. And all this is quite compatible with those explanatory properties being multiply realisable by yet lower-level properties not appealed to in the explanation.

Without doubt there are numerous metaphysical conundrums to be cleared up with the notion of multiple realisability and its cognates, but what should be appreciated are the epistemological lessons of this view as regards the tenability of reductive explanations in the face of prevailing multiple realisability of properties. To continue the humdrum illustration by the engine analogy, it is well-known that there are functioning subsystems (or parts) in an internal combustion engine which were implemented mechanically earlier but electronically later—e.g. the distributor. Such multiple realisation of the distributor function does not at all affect the reductive explanation of the engine at the level on which the causal-role characterisation of ‘being a distributor’ –property is satisfied by both mechanical and electronic implementations. When it comes to theoretical explanations, we have seen that there are property descriptions which can be multiply realisable in a more abstract sense going beyond such everyday causal-functional roles. (For example, consider Q in the reconstructed Fresnel’s derivation.)

Moving on to the actual practice of making reductive explanations of science, focusing on varying theoretical accounts of light in particular, the following perspective is naturally affiliated to the foregoing. A theory (such as Fresnel’s) is put forward as an account of how *possible* low-level facts (the oscillating mechanical ether molecules) would entail the explanandum phenomenon functionally analysed (light passed through polarizer, then partially reflected and refracted). [20] There are typically two parts to such ac-

[20] In actual practice the phenomenon to be explained may not be *fully* known at the time of theorising, of course, but become thus acknowledged only after its *prediction* based on some *less* detailed story of the phenomenon. The initial explanandum is then formed by whatever less detailed high-level facts must be accommodated by the theory.

count: there are (1) higher-order multiply realisable explanatory properties, and (2) a set of lower-order properties representing a possible realisation of the latter. Often, of course, these two parts are subtly interwoven, given the theoretical understanding of the time. It is no trivial matter to disassociate the two; it took almost a century for us to begin to see how non-mechanical properties could underlie the explanatory theoretical properties associated with the ether. If the theory is logically consistent[21] and fully compatible with a given body of evidence, then one can think of the theoretical description as a description of an *epistemically possible scenario* about which the theory is true—a way the world might be for all we know.[22] When the description of the explanandum phenomenon is sharpened as new evidence is incorporated it often happens that a particular epistemically possible scenario is no longer such: the low-level properties of the original proposed theory cannot realise the higher-level properties involved in the revised explanandum. A new theory is proposed and accompanied by a new, more fine-grained division of potential epistemically possible scenarios. By successive iteration science proceeds and gets closer to the ideal division of epistemically possible scenarios compatible with all the available evidence, ever.

But how should the realist apprehend such an iterative process of theoretical development in the face of the PMI challenge?

5 Towards a novel formulation of realism

The preceding discussion of non-explanatory surplus content and level-dependency of reductive explanation, together with the minimalist interpretation of Fresnel's derivation, manifestly points towards a novel realist position. To an anti-realist who picks out the Fresnel example as a case par excellence of the roots of her pessimism, we must respond by questioning the link she draws between approximate truth and reference. The notion of approximate truth

[21]It was noted above that Fresnel's theoretical assumptions were not altogether consistent, given the knowledge of elastic forces at the time. What is inconsistent, to be exact, is his appeal to purely transverse fluctuations in the *elastic* ether, where referring to the property of 'elasticity of ether' carries certain conceptions drawn from study of elastic materials. The minimalist core of Fresnel's derivation above is obviously not logically inconsistent as such.

[22]'The way the world might be for all we know' is called an *epistemically possible world*, as opposed to a counterfactually possible world. The theory of ether does not describe a counterfactually possible world if natural kind terms are considered to be rigid designators: there is no counterfactually possible world in which our term 'light' does not denote certain kinds of electromagnetic oscillations, assuming that this is what 'light' denotes in the actual world. This familiar Kripkean point about rigid designation does nothing to undermine the usefulness of the very clear intuitive notion of epistemically possible world in the context of the present argument. More formal presentation would require the use of some form of two-dimensional semantics. (cf. Chalmers, 2004)

appropriate for the realist—the notion of *explanatory approximate truth*—can and often does diverge from the naive notion the anti-realist intuition relies on. Analysing the nature of the theoretical constituents involved in a scientific explanation, and acknowledging the hierarchy of properties appealed to, makes room for a more fine-grained explanation of a theoretical success. Failing to do that has led realists astray, to make huge efforts in the semantics of theoretical terms in order to prove the anti-realist wrong but still remain true to the intuitive sense of approximate truth which demands that 'ether' refers to something.

But once we see this we can change tack just as well: we need not worry about 'ether' being non-referring *exactly because* it is actually *not* a central term in the right explanatory sense! If 'central' *just means* 'denoting an entity the existence of which is required for the minimal realist explanation of success' the same conclusion follows, because the existence of the ether is *not* required in that explanation. What is required is that there is a common core of theoretical properties appealed to in both Fresnel's and the corresponding modern day theorising. The phenomenon of light under theorising, analysed in dispositional terms, is reductively explained by these properties: they are described by their causal-nomological roles in the respective theories via the boundary and symmetry conditions equally present in both derivations, and they entail the phenomenon of reflection/refraction of polarised light. But there being unobservable properties thus described and related to the observations in question is no conceptual truth about light: these properties really are eligible for carrying the explanatory realist commitment. This is all that needs to be said to the anti-realist.

More can be said, however, about the way in which theorising proceeds in iterative fashion through successive falsities towards the truth. Tracking the nature of entities and processes underlying the appearances has turned out to be a challenge stretching human imagination to its limits in trying to explain phenomena by concepts increasingly far removed from the familiar. It is no wonder that in an attempt to conceptualise strange properties underlying some phenomenon more than enough is often said; it is clear that it frequently helps to focus on a possible lower-order realisation of the higher-order properties employed in a derivation. But the crucial properties are typically multiply realisable, and a proposed explanation is likely to contain some non-explanatory surplus. The art of finding the explanatory content and retaining it whilst discarding or replacing the rest is to be found in the science itself, not philosophy of science. For the latter, and for realism in particular, it is enough to describe a principled discriminatory *framework* to accommodate and explain the successfulness of this scientific practice. This paper has put forward the beginnings of a novel description of such a framework.

In celebration of the aforementioned art of discerning the explanatory, success-fuelling constituents, such a framework could perhaps be called *eclec-*

tic realism. The true measure of it, of course, only comes through a great variety of case studies and only prudent optimism can be expressed in the meanwhile. But when it comes to the highly frequented Fresnel-Maxwell case—the adopted battle ground for Psillos's standard realism and Worrall's structural realism, and widely commented on by others—it can be safely concluded that some *philosophical eclecticism* is rightfully needed to reach a defensible middle ground. Explicating the notion of explanatory approximate truth, and working explicitly at the level of properties as suggested by Chakravartty, and using this to take a closer-look at the details of Fresnel's derivation, shows how there is indeed more than structural continuity in the theory change from Fresnel to Maxwell, yet not enough to warrant the standard formulation of realism with its insistence on the successful reference of 'ether'.