UTRECHT UNIVERSITY
MSC PROGRAM IN HISTORY AND PHILOSOHY OF SCIENCE

Empirical Equivalence and Underdetermination of Theory Choice

A Philosophical Appraisal and a Case Study

Pablo Acuña
Supervisor: Dennis Dieks
August 2012

The illustration in the cover, Alice and Tweedledum and Tweedledee, is taken from the original drawings by John Tenniel in Lewis Carroll's *Through the Looking Glass.*

Acknowledgments

I am deeply grateful to my supervisor, Dennis Dieks, for all his support, help and fruitful criticisms along the writing of this thesis. It was an honor to have the chance of being guided in this big learning adventure by such a remarkable scholar. Moreover, that in all of our meetings and discussions he always received me with a warm and friendly smile is something that I cannot thank enough.

I would also like to show my gratitude to Jeroen van Dongen. Besides being a great teacher –I discovered how fascinating the case of Lorentz *vs.* Einstein is in one of his courses–, he offered me his academic and personal support and orientation in times when I was somewhat confused about my thesis plans.

I owe much also to Roberto Torretti. His work has been of great inspiration ever since my early days in the world of philosophy, and his generous suggestions and criticisms concerning some subjects in this thesis were certainly crucial to improve it.

I am also very grateful to my good friends and mates in the HPS program. All the nice conversations (about physics, epistemology, football, life in Holland, and even about the weather) with Vincent Schoutsen, Chao Kang Tai, Giulia Paparo, Nick Evans, Steve Shapiro and Tom Stekenburg contributed to the completion of this work.

I would like to express my special gratitude to Alfred van Herk and his wife Marjia, for their beautiful friendship has helped me not to feel always so far from home. The same holds for my "Hungarian family": Janos, Marta and Mate in Budapest.

Finalmente, quiero agradecer a todo mi *clan* en Chile. Sin el amor de mis padres-hermanos-primos-amigos nada de esto sería posible ni valioso.

For Kata, for all the joy

You know that I care what happens to you,
and I know that you care for me.
So I don't feel alone,
or the weight of the stone,
now that I've found somewhere safe
to bury my bone.
And any fool knows a dog needs a home,
a shelter from pigs on the wing.

Roger Waters.

TABLE OF CONTENTS

INTRODUCTION

Predictive equivalence between scientific theories –along with the empirical underdetermination of the choice to be made that it is supposed to entail– is one of the deepest and most debated issues in the philosophy of science. This is hardly surprising, for the rationality of theory choice is at risk if the problem is not solvable: if the problem at issue is intractable, it might be necessary to quit to efforts in order to provide an account of theory acceptance that is grounded on empirical evidence.

In spite of all the attention and pages that have been devoted to this problem, it is awkward that most of that attention and those pages deal with the subject only from an abstract and conceptual point of view. None of the solutions provided has been directly tested in a case of 'real-life' science. This is rather curious, since there are well-known cases of predictive equivalence in the history of science that could be used as a *case study* for the problem at hand. The example of Einstein's special relativity *vs*. Lorentz's ether theory is symptomatic. Even though there is plenty of excellent historical work available in this case, it has not been used in order to evaluate how the different philosophical positions regarding the general problem of empirical equivalence and underdetermination fit in real science. It is true, though, that authors like Schaffner, Zahar and Janssen have proposed analyses of the Einstein *vs*. Lorentz case in which they argue for what are the reasons that can be rationally invoked to make a decision. However, none of them has approached the subject as a specific instance of the problem of empirical equivalence and undedetermination of theory choice.

This is a gap that I will try to fill in this work. I argue for a solution for the problem and I carry out a test of that solution in the actual case of the two empirically equivalent theories mentioned: Einstein's special relativity and Lorentz's ether theory. In the first chapter I undertake a philosophical assessment in which I try to provide a clear and precise exposition of the problem, an examination of the most important possible solutions that have been proposed in the relevant literature, and an evaluation of these solutions in order to determine which of them work and which of them do not. The solution I defend is based on a very influential argument introduced by Larry Laudan and Jarret Leplin in 1991. Even though I think that their line of reasoning is essentially correct, I introduce important provisos and qualifications which I think clarify the real scope and nature of the solution that these authors offer.

In the second chapter I offer a historical and conceptual outline of the Einstein *vs.* Lorentz case –based on the available work that I mentioned above. This outline includes an exposition of the main tenets of both theories, and a comparison between them in order to show that they are different and rivals –but at the same time predictively equivalent. On these bases, I examine the reasons that can be invoked to make a decision between the contending theories, and I offer an evaluation of the situation in which I argue that some of those reasons are objectively grounded and that some others are not fully justified.

The general conclusions of this work are given by an assessment of the conceptual solution proposed in the first chapter by means of the case study presented in the second. The main result I obtained is that the solution I defend does work, and that the provisos and qualifications that I introduce to restrict and clarify the scope of Laudan and Leplin's reasoning are correct: these authors argue that the problem gets *definitively* refuted, but I claim that their argument shows that some features which are a common and essential part of scientific practice are able to provide rational and evidential foundations for theory choice in cases of predictive equivalence. However, the solution at issue leaves open the possibility that empirical equivalence *could* be a source of underdetermination. But even though the problem at hand remains a possible scenario in actual science, the solution I defend shows that some typical methods and procedures of scientific practice are able to provide an evidentially-grounded choice; and this is a clear indication that the problem -and its solution– are more a matter of *science* than a matter of *epistemology*.

CHAPTER 1

A Philosophical Appraisal

This chapter is divided in five sections. In the first one I provide a precise formulation of the problem in terms of an argument constituted by two premises and a conclusion. I also undertake an analysis of what is precisely at stake in the problem, and of what are the presuppositions underlying both the premises. In the second section I argue that three possible solutions that have been provided do not work: the view that predictively equivalent theories are simply two different formulations of the same theory, the recourse to simplicity and explanatory power as non-empirical virtues which could be used to pick one of the theories, and Jarret Leplin's attack on the logical soundness of the argument which constitutes the problem. In the third section I offer a detailed exposition of the solution that Larry Laudan and Jarret Leplin proposed in 1991, and I add an argument by Richard Boyd as a useful complement for that solution. In the fourth section I deal with three *remaining challenges* of empirical equivalence and underdetermination which, at face value, are of a different nature than the main argument that settles the problem: the *van Fraassean* alternative formulations of Newton's theory, the 'Poincare-Reichenbach parable', and the case of *total theories* or *systems of the world*. I argue that none of these remaining challenges implies difficulties beyond the scope of the solution provided in section three. The fifth and final section of this chapter consists on a re-evaluation of the solution provided by Laudan & Leplin (and Boyd). I argue that these authors assign a scope to their position which goes too far. They claim that the problem at hand has been *refuted*, but I show that even though their argument is essentially correct –in the sense that science has the tools to dissolve it–, empirical equivalence as a source of underdetermination of theory choice remains as a possible scenario.

I. Sketching the problem

The problem of underdetermination, as a consequence of empirical equivalence between two competing theories, is a very simple one, and, at face value, also very deep and difficult. If two different theories entail exactly the same observational consequences they are also equivalent from a confirmational point of view –provided that the standard hypothetic-deductive model of evidential support is adopted–; that is, everything which confirms or disconfirms one of the theories also confirms or disconfirms the other. Therefore, there are no possible evidential resources to accept one of them and to reject the other. This conclusion threats the rationality of the decision to be made between the competing theories. Moreover, the argument is commonly presented as establishing that *any* theory has an actual or virtual empirically equivalent competitor, so in this case *any* theory whatsoever is underdetermined;

and consequently, theory choice as a general feature of scientific practice could be seen as ungrounded.

In a more structured way, the argument of empirical equivalence (EE) as leading to underdetermination (UD) can be put in terms of the following two premises which entail the problematic conclusion: 1) *for any theory T and any body of observational evidence E, there is another theory T′ such that T and T′ are empirically equivalent with respect to E*; and 2) *the entailment of the evidence is the only epistemic constraint on the confirmation of a theory*.

The rationale for the first premise comes from two main sources. Some authors claim that given any theory that entails certain observational consequences, there are algorithmic procedures which produce another theory with exactly the same observational consequences. On the other hand, what is commonly known as *the Quine-Duhem thesis* also seems to support this premise. Duhem has shown that the logic of evidence is holistic: a theoretical hypothesis can entail empirical consequences only with the help of auxiliary hypotheses, so that, *logically speaking*, any evidence could be accommodated by any theory given the necessary arrangements on the auxiliary assumptions[1]. This last conclusion leads to the effectiveness of the first premise: theory *T*, compounded by hypothesis *H* and auxiliary assumptions *A*, entails *E*; and another theory *T′* compounded of hypothesis *H′* and auxiliary assumptions *A′* also entails *E*, so that *T* and *T′* are empirically equivalent.

The second premise is supported by the traditional hypothetic-deductive model of evidential confirmation. Roughly and briefly speaking, this model, introduced and supported mainly by logical positivists, asserts that an observational report can count as evidence for a certain hypothesis if and only if a sentence expressing that report is entailed by the hypothesis at issue (along with auxiliary assumptions, of course). In spite of the many problems that the logical positivist program had to face, this model of confirmation –maybe because of its simplicity and *prima facie* obviousness– has remained a milestone in the philosophy of science.

Given this diagnosis of the problem, it is quite clear that any attempt to solve it will have to be put in terms of criticism of one of the premises, or of both. Obviously, such criticism will be directed to a critical assessment of their supporting theses –algorithms, holism and confirmation model. Before undertaking this task, I will demarcate the perspective from which I will tackle the issue with respect to another possibility. The EE and UD problem has usually been used as an argument against the cogency of a realist conception of scientific theories. Realists hold the view that for a theory to be accepted one might have reasons to believe that the theory is true or at least approximately true. Evidential confirmation, of course, looks to be the most robust candidate to offer a reason like that. But if EE is a general pattern for theories, then the support of evidence as leading to belief in their truth gets denied. This is not the worry I will address. I will tackle the problem of EE and UD from a more basic point of view. If the problem is indeed intractable, the

[1] Duhem himself, however, stated that this *logical* possibility is not all that it takes to have an empirically equivalent rival. He argued for a scientific *good sense* that indicates good scientists when is no longer possible, scientifically speaking, to continue saving a hypothesis in the face of disconfirming evidence –see Duhem 1908, 37-9. As we will see, this observation is essential in the assessment of the problem.

rationality of evidential support and of theory choice is threatened, regardless of whether one is a realist or not[2]. A solution or dissolution of the problem restores such rationality, and of course, it would be welcome by a realist. However, the fact that I will offer reasons to deny the problem and to embrace evidential support and theory choice as rationally grounded features of scientific practice does not mean that I share the view that these features are enough for believing theories as true. In simple words, my position about EE and UD is neutral with respect to the debate about realism.

II. Discarding some possible ways out

In this section I will briefly review two commonly attempted ways to avoid the problem of EE and UD, and a general argument intended to show that the EE thesis itself is inconsistent with UD. I think that these three approaches are unsuccessful or at least incomplete, and that the reasons of their failure help to clarify what is the right path to fruitfully block the problem.

a) EE as different formulations of the same theory

If the empirical equivalence between two theories can be shown not to be a case of a competition, but a case in which the very same theory is presented in two different formulations, it is obvious that the UD does not even come up. The "choice" to be done can be grounded on pragmatic considerations such as simplicity or the like, since no theory is really being rejected. Choosing one or the other formulation is not an epistemic issue.

This was the position that many logical positivists held. Its rationale comes from the verificationist criterion of meaning. The meaning of a term, of a sentence, or of a theory in this case, is nothing but the method to verify it. Such method, in the case of scientific theories, is given by their observational consequences. Therefore, if two theories have exactly the same empirical consequences, they have exactly the same meaning. From this semantic point of view, two empirically equivalent theories are just synonyms, so that the choice between them has nothing to do with evidential or epistemic conditions.

Insofar as this view depends on the verificationist criterion of meaning, if such criterion is shown to be untenable, the related view with respect to EE also falls. As it is widely known, this is exactly what happened during the last century. However, if the view that empirically equivalent theories are the same theory can be supported by arguments which do not depend in the logical positivist semantic criterion, this way out can be reintroduced. Some attempts along this line of thinking have been done, especially within

[2] Actually, and as J. Busch shows, the argument of EE and UD does not favor an antirealist constructive-empiricist position. The non-empirical virtues of a theory that the constructive empiricist holds as justifying his choice of one theory over the other –which are mainly related to explanatory virtues, and which are not available for the realist since they do not support the truth of the chosen theory–, cannot be shown to be indicators of empirical adequacy, which is the main goal of science according to constructive empiricism. Consequently, Busch concludes that EE and UD constitute a problem for this antirealist view as well. See Busch 2009.

the semantic conception of scientific theories[3]. For example, John Norton argues that whenever we face a case in which empirical equivalence can be asserted between two theories we might suspect that they are nothing but variant formulations. His argument states that in order to determine the equivalence a set of conditions must be met: 1) we must have a tractable description of their observational consequences; 2) each of the theories must make essential use of its own theoretical language in the entailment of their observational consequences;

> 3. If we are able to demonstrate observational equivalence of the two theories, the theoretical structures of the two theories are most likely very similar. While it is possible that they are radically different, if that were the case, we would most likely be unable to demonstrate the observational equivalence of the two theories. For the theoretical structures are what systematizes the two sets of observational consequences, and a tractable demonstration of observational equivalence must proceed by showing some sort of equivalence in these systematizing structures.
> 4. The two sets of theoretical structures may be inconvertible without loss; or they may not be. In the latter case, there would be additional structures present in one theory but not in the other. However, any such additional structure will be unnecessary for the recovery of the observational consequences. That follows since the additional structure has no correlate in the other theory, yet the other theory has identical observational consequences. Thus any additional structures will be strong candidates for being superfluous, unphysical structures.[4]

This list of requirements, Norton argues, implies that if two theories are so demonstrated to be empirically equivalent, they are likely to be a case of variant formulations of the same structure (theory). Cases in which no structure is lost, like matrix mechanics and wave mechanics in quantum theory are clear and straightforward[5]; but a further step is needed when there is loss of structure. The remnant structure might be considered as representing nothing physical since it is not needed for the observational entailment in the more economic theory. However, Norton himself points out that the debate about the superfluous character of the remaining structure is a very complex one. In any case, his point consists in showing that, at least in principle, EE strongly *suggests* theoretical identity. This does not mean that EE *means* theoretical identity right away. But if the EE and UD argument is to be asserted, the issue of establishing that EE is indeed a case of two rival theories must be independently evaluated.[6]

Norton's view seems to be much more promising and cogent then the logical positivists' one. However, it depends on whether the semantic-structural conception of

---

[3] The semantic or structuralist conception, unlike logical positivists, states that theories are extra-linguistic entities. They are constituted by a set-theoretical predicate –the structure– that is satisfied by a model of the predicate in the real world. This approach is sometimes called semantic insofar as it construes theories as what their formulations refer to when their formulations are given a formal semantic interpretation. However, structuralists argue that the *mapping* relation between the model in the real world and the structure, unlike logical positivist's concept of *correspondence rules*, is not a part of the theory. Another important feature of this approach is that it accepts the thesis of *theory ladenness of observation*: all terms are theory laden, but if they count as theoretical or observational is a context-dependent issue. The same concept can play an observational or a theoretical role depending on what theory it is a part of.

[4] Norton 2008, 34-5.

[5] This is, of course, the 'traditional' view on the matter. F. Muller challenges this view and states that, when originally formulated, matrix mechanics and wave mechanics were not empirically equivalent –and therefore not the same theory. According to Muller, they became empirically and ontologically identical with the work of von Neumann in 1932. See Muller 1997.

[6] Norton 2008, 33-40. For other considerations of theoretical identity between empirically equivalent theories from a semantic-structural point of view, see also French 2011, Yalcin 2001 and Mormann 1995.

theories is adequate enough –an assessment of this issue goes quite beyond the scope of this work. Moreover, Norton's stance is different from the logical positivistic view in an important way. For the latter, the semantic criterion *necessarily* implies that empirically equivalent theories are nothing but alternative formulations of the same theory. In Norton's case, the equivalence only *suggests* that it could be a case of alternative formulations –and it also takes for granted that the extra-structure must be considered as superfluous–. Actually, and in spite of Norton's suggestion, it might turn out that the most relevant cases of EE are indeed cases of two different and rival theories.

However, Norton's argument can be illuminating, if not as a direct solution to the problem, at least as a criterion to show the ineffectiveness of the algorithms to produce empirically equivalent theories given *T*, for he states that the output of algorithms are a case of 'gratuitous impoverishment'. Norton claims that when we have a pair of EE theories in which one of them was algorithmically produced, the 'artificial' member of the pair generally has *less* structure than the other. However, the additional structure in the non-algorithmic theory is essential for the entailment of the observational consequences in *both theories*. For example, there is a proposed algorithm which states that given a theory *T*, we can construct the theory *T'* with identical observational consequences, but with the negation of all the theoretical claims of *T*. Norton's point is that these theoretical claims are essential for the derivation of the observational statements in the *two* theories, but in the artificial one that part of the structure is cut out:

> The natural and cultured pairs of observationally equivalent theories can be construed as variant formulations of the same physical theory if we regard the additional structures of one or other of the pair as physically superfluous [Norton's *natural pairs* are 'real-life' pairs of theories, such as Lorentz's vs. Einstein's; whereas his *cultured pairs* are 'virtual theories' designed by philosophers for specific philosophical ends, such as in the Poincare-Reichenbach *parable*]. The artificial pairs are similar insofar as one member of the pair does accord physical significance to a structure whereas the other member of the pair does not. In these artificial cases, however, I want to urge that the second deprives this additional structure of physical significance improperly. That is, it represents what I shall call a 'gratuitous impoverishment' of the second theory. Unlike the natural and cultured pairs, in artificial pairs is essential to both theories' derivation of their observational consequences and is well confirmed by these observational consequences. Indeed, in most cases the additional structure is fixed by the observational evidence, whereas the superfluous structure of the natural and cultured pairs is usually not. If we discard this additional structure, we lose an essential part of the machinery of both theories and deny something for which we have good evidence –the academic equivalent of burying one's head in the sand[7].

More generally, since the first premise of the argument states that for any theory *T* there is an empirically equivalent rival *T'*, the issue of whether the two theories are really rivals and different becomes relevant. Therefore, a general criterion for identity or non-identity between theories would be most useful. However, it is not currently available. But that does not mean that the problem of EE and UD cannot be posed or that it cannot be solved. The current situation is that whether two EE theories are identical or not must be

[7] Norton 2008, 39.

approached case by case, and there is no reason to presuppose, as the logical positivists did, that all of the cases will be instances of theoretical identity. As P. D. Magnus affirms,

> I do not deny that a criterion of theory identity would be a nice thing to have. Problems of theory individuation, of which the problem of identical rivals is a special case, are interesting in their own right. Resolving them, however, can only come as the result of a careful examination of the history of science –an examination which must be left for some other time. I draw the modest conclusion that this open question need not turn us back from considering underdetermination.[8]

b) Non-empirical virtues: simplicity and explanatory power

A second possibility that I will reject as ineffective to solve the EE and UD problem is to take recourse to theoretical non-empirical virtues of the theories involved, which by definition are not related to their empirical basis. The typical virtues which are discussed in the literature are scope, simplicity and explanatory power. In the case of the problem at stake, the former cannot be considered as a criterion of theory choice, for if two different theories have a different range of phenomena that they explain, I cannot see how they could be EE in the first place. That leaves us the remaining two virtues. I will briefly explain why I think they are not able to provide a solution of the problem at hand.

In the case of simplicity, the criticism is well known. It can be formulated in two levels. On the one hand, it is quite difficult to give a clear definition of the concept; that is, what is for a theory, in objective terms, to be simpler than other? It looks to be a feature which intrinsically depends on subjective considerations, "a person's simplicity is another's person complexity". In general terms, the criterion of simplicity, in order to be asserted, seems to depend on certain previous and independent considerations about the specific case and theory to which it is to be applied. For example, Mario Bunge claims that simplicity can be asserted for a theory from several different perspectives: syntactical, semantic, epistemological and pragmatic; and that these various kinds of simplicity are not compatible with one another[9]. Therefore, the first difficulty to refer to simplicity as a feature which could provide a non-empirical reason to choose between two EE theories lies on the ambiguousness and subjectivity involved in the definition of the concept.

The second level of the criticism is more serious. Even if we take for granted that we have a clear and objective definition of simplicity and a clear and objective criterion to know when to apply it in particular cases, it seems to be that it is not a feature on which one could base a *rational* choice between two EE theories. The argument for simplicity as an epistemic virtue fails to show its connection with a theory's empirical success. First of all, there is no *a priori* reason for the universe to be more likely to be explained by simple theories rather than by complex ones. Moreover, the complex and discarded theory might contain the seed for further empirical enhancement which could explain phenomena that the simpler one cannot.

---

[8] Magnus 2003, 1263. In this paper the author offers interesting analyses and criticisms of possible criteria for theory individuation. Especially interesting is his consideration of the matrix and wave mechanics case from a semantic-structuralist point of view.

[9] Bunge 1961, 122.

In other words, there is nothing in the concept of simplicity which could show that, in the long run, the simple theory will be a better theory, empirically speaking[10].

It is quite obvious that simplicity can indeed be considered as a scientific desideratum. Simpler theories, for example, are more easily testable. However, this cannot count as an indication of a necessary connection with empirical success:

> Simplicity is ambiguous as a term and double-edged as a prescription, and it must be controlled by the symptoms of truth rather than be regarded as a factor of truth. To paraphrase Baltasar Gracián –"*lo bueno, si breve, dos veces bueno*"–. Let us say that a theory, if simple, works twice as well. But this is trivial. If a practical advice is wanted as a corollary, let this be: Ockham's razor –like all razors– must be handled with care to prevent beheading science in the attempt to shave off some of its pilosities. In science, as in the barber shop, better alive and bearded than dead and cleanly shaven.[11]

Turning to explanatory power as a candidate to solve the problem, we have that, *mutatis mutandis*, both levels of the previous criticism hold as well. To make a case for a higher power of explanation of a theory over an empirically equivalent rival obviously presupposes a clear criterion of explanation in general. In this case the problem is not the subjective factors involved, but that the issue of what a scientific explanation is or must be is, to a certain extent, an open philosophical question. Moreover, explanation is sometimes understood, just like simplicity, as context-dependent. For example, van Fraassen considers scientific explanation to be an answer to a why-question asked for a certain phenomenon. Therefore, what is a satisfying explanation depends on what kind of why-question has been put in the first place, so that different questions about the same phenomenon could determine different adscriptions of explanatory power for the same explanation[12].

On the other hand, the link between explanatory power and empirical success is also problematic. Even if we hinder the context-dependence nature of explanation, so that a theory in an EE pair could be regarded as preferable because of its explanatory features,

---

[10] Constructive empiricists propose that both the theories can be accepted as *empirically adequate*. Even though they are rivals, the van Fraassean stand is not committed to the truth or to the trans-empirical theoretical parts of the theories, so there is no problem for them in saying that the theories are equally good in terms of empirical adequacy. However, van Fraassen also asserts that there is more than mere belief in theory acceptance; pragmatic issues –such as decisions regarding which of the contending research programs that the theories are a part of is going to be developed– are involved too. In this pragmatic commitment, pragmatic considerations, such as explanatory power or simplicity of the theories, can be invoked. However, EE and UD between theories, as I will show below, are time-indexed features. Since EE and UD can be broken –and thus one of the theories in the EE pair can become better than its rival from an empirical adequacy point of view–, to decide in terms of simplicity or explanatory virtues becomes a rather risky move, for there are no *a priori* grounds to connect these properties with empirical adequacy.

[11] Bunge 1961, 149.

[12] Van Fraassen 1980, 134-57. He argues that its pragmatic nature is a reason to not consider explanation as one of the central aims of science. The epistemic dimension of science is contained in the factual knowledge that (empirically adequate) theories provide. Dieks & de Regt (2005), on the other hand, assign to explanation and understanding an essential role in the epistemic goals of science, even though they recognize the pragmatic nature of understanding (and, *a fortiori*, of explanation). They claim that a phenomenon *P* is understood if there is an *intelligible* theory about *P*, and a theory *T* is intelligible if scientists are able to recognize qualitatively characteristic consequences of *T* without performing exact calculations. Different 'conceptual toolkits' such as visualization, causal explanations and unifications that different theories can provide –all of which can be considered as explanatory virtues– can work as sources of intelligibility. However, there are no necessary or sufficient conceptual toolkits that can assure intelligibility right away. It depends on a determinate state of science and on contextual matters: "There is no universal tool for understanding, but a variety of 'toolkits', containing particular tools for particular situations. Which tools scientists have at their disposal, depends on the (historical, social, and/or disciplinary) context in which they find themselves. This context-dependence is typical of a meso-level nature, i. e., it is the scientific community that determines what tools are available and which skills are required to achieve understanding", Dieks & de Regt 2005, 158.

there is no *a priori* reason to believe that such a theory is (or will be) superior also from the point of view of its empirical success. Concluding, Bunge's evaluation of simplicity as a desideratum of science, but as detached from any direct or indirect evidential import leading to empirical success, applies to explanatory virtues as well.

Both alternatives that I have examined so far share the feature that they deny the second premise of the problem at hand. They both state that the entailment of observational consequences is not the only source of confirmational support, for they state that non-empirical virtues such as simplicity and explanatory power are also relevant features for the acceptance or rejection of a theory. However, and despite that simplicity and explanatory power are positive and welcome characteristics in a scientific theory, they are not able to ground a final and problem-free decision regarding theory choice. Now I will consider a general argument which challenges not the truth value of any of the premises, but the logical soundness of the argument.

c) Leplin's inconsistency argument

Jarret Leplin offers a solution which consists on showing that EE cannot lead to UD insofar as the EE thesis and the UD thesis are inconsistent with respect to one another. That is, according to Leplin, and for logical reasons, there is no problem at all. His proof of the inconsistency between the theses at issue relies on the fact that in order to determine if two theories are EE, it is needed that the auxiliary assumptions that permit the entailment of the observational consequences may be specified and well established. However, if UD is the case, then the auxiliary hypotheses required are also underdetermined, and the set of the possible and available auxiliaries gets unclear and undetermined. Therefore, there are no clear and well established auxiliaries to perform the entailment of observational consequences from the theories we want to asses. Thus, what are the observational consequences of a theory, if UD is the case, is impossible to determine. If UD is true, EE cannot be decided; and if EE is true, UD cannot be true:

> *The truth of UD would prevent the determination that theories are empirically equivalent in the first place.* Because theories characteristically issue in observationally attestable predictions only in conjunction with further, presupposed background theory, what observational consequences a theory has is relative to what other theories are willing to presuppose. As different presuppositions may yield different consequences, the judgment that they have the same observational consequences –that they are empirically equivalent– depends on somehow fixing the range of further theory available for presupposition. And this underdetermination ultimately disallows.[13]

If the available auxiliary assumptions are not established and justified by a certain criterion, any possible auxiliary could be used. The consequence would be that the auxiliary hypotheses could never be better supported than the theoretical ones that are to be evaluated, leading to a radical holism in which it would be completely impossible to asses any theoretical hypothesis (more or less) directly; a *Duhemian nightmare* would be the case:

[13] Leplin 1997a, 154-5. See also Leplin 1997b.

> Admissible auxiliaries are those independently warranted by empirical evidence. Unless auxiliaries are *better supported* than the theory they are used to obtain predictions from, those predictions cannot be used to test the theory. The significance of their success or failure would be indeterminate as between the theory and the auxiliaries. The result would be a holism that enlarges the possible units of empirical evaluation, and prevents epistemic support from accruing to theories directly. Such is the upshot of the classic theses of Duhem, who stressed the ineliminability of auxiliaries from prediction.

If this were the case, then auxiliaries would be as underdetermined as theoretical hypotheses, this is what a radical holist thesis of UD states. Therefore, there would be no epistemic standard to establish some hypotheses as justified assumptions to be used to derive observational consequences from theoretical hypotheses, "*there will be no fact of the matter as to what the empirical consequences of any theory are*"[14]; and consequently, it cannot be stated whether two theories are empirically equivalent or not.

If we evaluate Leplin's argument that the first premise and the conclusion of our problem are inconsistent, it turns out that it does not work completely. What Leplin shows is that the first premise, the fact that for any theory there is an empirically equivalent rival, is inconsistent with the conclusion of UD only insofar as it is supported by the Duhemian holist thesis, as we saw in the first section of this chapter. If UD is the case in this sense, then it is impossible to determine the set of observational consequences of any theory. However, Leplin concedes that if we grant that such a class is already specified for any theory, and we introduce a certain method to produce an observationally equivalent theory which does not need to determine further classes of observational consequences, then the UD problem raises again. In other words, given a certain theory T whose observational consequences O are established, if there exists an algorithm to be applied to the theoretical hypotheses of T such that a new theory T' results (that includes well established auxiliaries) which also entails O; then the problem of UD comes up anyway:

> The only general strategy for upholding EE that is capable of coping with this difficulty is algorithmic. This strategy exists in many versions. Usually, the empirical-consequence class O of an arbitrary theory T is supposed to be given, and the algorithm operates on T to produce another theory T' whose consequence class is also O.[15]

At best, Leplin's criticism shows that the EE premise and the UD conclusion are inconsistent only if we understand EE as the result of a radical holism. If we understand it as the result of the operations of certain algorithms which presuppose the class of observational statements, the inconsistency is not the case. Now I will offer a further objection intended to show that even this restricted interpretation does not hold.

Let us suppose that there is a well-established epistemic-evidential criterion to determine auxiliary hypotheses to be used in the derivation of observational consequences from new theoretical hypotheses. Let us also suppose that we do not know of any algorithm

[14] Leplin 1997a, 155.

[15] Leplin 1997a, 158. Leplin's proviso to his own argument is somewhat awkward. I understand that he is presupposing that the class of observational consequences O of the theory T is already specified, so that the algorithmically tailored theory T' and its correspondent O' does not require the introduction of new or further auxiliary hypotheses. However, the entailment of O from T does require auxiliary hypotheses!

of the mentioned type. It is also assumed that the theoretical creativity and ingenuity of scientists is an inherent feature of hypothesis-formulation. Consider a new hypothesis *H* which along with the class of well-established auxiliary assumptions (A_1, A_2, A_3) entails *O*; and also consider a second new hypothesis *H′* which together with the class (A_4, A_5, A_6) – which are also well-confirmed– also entails *O*. I referred to the importance of ingenuity and creativity to insure that both theoretical hypotheses are well justified according to theoretical, foundational and philosophical criteria, they are neither *ad-hoc* or bizarre. If this is the case, then obviously *H* and *H′* are underdetermined theories because they are empirically equivalent and not *ad-hoc* or bizarre. What Leplin's argument in this case still shows is only that any further theory which needs either *H* or *H′* to entail observational consequences will face the problem to justifiably determine them[16]. In other words, my criticism is meant to show that algorithms are not necessary for a case in which the observational class *O* is well defined and entailed by two different and rival theories. That is, if we consider a moderate holist thesis to support the EE premise, a holist thesis in which auxiliaries are necessary but in which it is also possible to justify them, then Leplin's inconsistency does not arise[17]. Therefore, his argument is not enough to logically block the problem of UD.[18]

So far I have shown why two attempts to escape from the problem by denying the second premise fail, and why attacking the logical structure of the argument also fails. However, it still remains the possibility to question the first premise of course. Moreover, the recourse to non-empirical virtues as a ground for theory choice is not the only way to jeopardize the second premise. In the next section I will discuss and support the best attempts to solve the problem that go along these paths.

III. Laudan & Leplin's solution

In a very influential paper Jarret Leplin and Larry Laudan undertake an attempt to provide a solution to the problem of EE and UD which I consider essentially correct. Their argument consists in a critical assessment of both the premises that lead to the UD conclusion. If one ponders it in detail, these authors do not completely deny what the premises state, but they show that what the premises claim has a much more reduced scope than what is normally believed. That is, the premises are so weakened that the resulting problem gets diluted. I will provide a careful exposition of each of the steps of their argument, along with the main criticisms it has received. I think that those criticisms are not successful, but to consider them is a good way to precisely understand what has been really achieved.

[16] However, I think it is possible that whether further theories use *H* or *H'* to entail empirical consequences might be indifferent with respect to the content of the class *O*, precisely because *H* and *H'* are empirically equivalent.
[17] Actually, if one pays attention to Duhem's concept of *good sense* as dictating when a falsification or confirmation of a theory has been obtained *in spite of the logical possibility to introduce ad-hoc hypotheses or logical tricks*; it is quite clear that this moderate holism is Duhem's thesis, not the radical one.
[18] For different criticisms of Leplin's argument see Douven 2000 and Sarkar 2000.

1. The first premise

a) EE, observability and auxiliary assumptions

Laudan & Leplin's attack on the claim that *for any theory T there is an alternative rival T' such that both entail exactly the same class O* is twofold. On the one hand, they show that three relatively non-controversial theses regarding the nature of evidential confirmation imply that EE is not a necessary and universal feature in the sense that the first premise claims. Secondly, they attack the algorithms to produce observationally equivalent theories by showing that their output is not really a *rival theory* with respect to the input theory[19].

The three theses they refer to are 1) *the variability of the range of the observable*: "Any circumscription of the range of observable phenomena is relative to the state of scientific knowledge and the technological resources available for observation and detection"[20]. In other words, whether a certain entity or process described by a theory qualifies as observational or not does not depend only on the meaning of the corresponding terms, but also, and crucially, on the available experimental methods and instruments of a certain time[21]. 2) *The need for auxiliaries in prediction*: "theoretical hypotheses typically require supplementation by auxiliary or collateral information for the derivation of observable consequences"[22]. This thesis is nowadays so widely accepted that needs no further comments. Finally, 3) *the instability of auxiliary assumptions*: "auxiliary information providing premises for the derivation of observational consequences from theory is unstable in two respects: it is defeasible and it is augmentable"[23]. This thesis means that, depending on scientific progress (or change), the class of auxiliary assumptions which are suitable for the derivation of empirical consequences from theoretical hypotheses might be enlarged (or reduced) as a consequence of new well-supported theoretical features or novel facts.

Taken together, the effect of these three theses on the claim of EE as a general and necessary condition for any theory whatsoever is clear and profound. If what is observable is variable and depends on the background current knowledge, and if the class of auxiliary assumptions is also so; it follows that whether two theories are observationally equivalent depends on a certain state of scientific knowledge. EE cannot be an absolute feature of

[19] As I mentioned above, Norton shares this position with respect to algorithms. However, he suggests that a *specific kind* of algorithms is ineffective, for they produce theories which are *gratuitously impoverished*. Laudan & Leplin's stance allows a more general assessment of why *different* types of algorithms produce pseudo-theories rather than theories.

[20] Laudan & Leplin 1991, 451.

[21] Laudan & Leplin acknowledge that in spite of being a non-controversial thesis, van Fraassen rejects it. However, they claim that "we reject [van Fraassen's] implicit assumption that conditions of observability are fixed by physiology. Once it is decided what is to count as observing, physiology may determine what is observable. But physiology does not impose or delimit our concept of observation. We could possess the relevant physiological apparatus without possessing a concept of observation at all. The concept we do possess could perfectly well incorporate technological means of detection. In fact, the concept of observation has changed with science, and even to state that the (theory-independent) facts determine what is observable, van Fraassen must use a concept of observation that implicitly appeals to a state of science and technology". *Ibid*, 452, footnote 3. In other words, they state that the fact that observation –in a scientific context– is not an immanent and immutable concept and that it does change along the course of science. Van Fraassen's position, even though it refers to facts of our physiology to define what is observable for us, is linked to a specific state of scientific and technological development. Otherwise, van Fraassen should endorse an immanent and immutable concept of scientific observation. If this is what Laudan & Leplin argue, I totally agree.

[22] *Ibid*, 452.

[23] *Loc. cit.*

theories. That two theories are empirically equivalent today cannot be hold as an indication that they will forever be, since a new instrument or a new theory could perfectly remove the equivalence:

> Therefore, any determination of the empirical consequence class of a theory must be relativized to a particular state of science. We infer that empirical equivalence itself must be so relativized, and, accordingly, that any finding of empirical equivalence is both contextual and defeasible. This contextuality shows that determinations of empirical equivalence are not a purely formal, a priori matter, but must defer, in part, to scientific practice. It undercuts any formalistic program to delimit the scope of scientific knowledge by reason of empirical equivalence, thereby defeating the epistemically otiose morals that empirical equivalence has made to serve.[24]

The sense in which these claims weaken the first premise of the problem is clear. Since the premise implicitly asserts that EE is an everlasting inter-theoretical relation, the UD it helps to entail is also an everlasting problem –and insurmountable by any further development of science. The problem is that there cannot be a rational choice because there will never be a rational criterion to make the choice. Laudan & Leplin claim that if two theories are shown to be observationally equivalent, it can only be stated that they are so *now*, and that it is reasonable to wait for an empirical criterion to distinguish them and make a choice.

Andre Kukla offers a criticism of this view which helps to see what is really achieved by it. He states that we can accept that two theories (T_1, A_t) and (T_2, A_t) can be only considered as empirically equivalent at time *t*. That is, EE is a *time-indexed* relation, and this is all what Laudan & Leplin have shown: "there is nothing in the argument that would force me to give up the view that every indexed theory has empirically equivalent rivals with the same index"[25]. Kukla's assertion means that the EE claim remains ultimate and leading to UD not in a temporal sense, but in an extensional way. That is, the thesis has been temporally relativized, but its scope is still universal in the sense that holds for *any* theory whatsoever. If at time *t'* the equivalence between T_1 and T_2 has been removed, the problem pops up once again since there will be a theory (T_3, $A_{t'}$) empirically equivalent to T_1. If this is the case, any theory whatsoever has a rival theory which is empirically equivalent to it, in spite of the time-index. The only difference is that the empirically equivalent rival is not the same along time. In other words, any theory will always be underdetermined, but the source of the UD simply varies along time indexes: "the point is that we know that, whatever our future opinion about auxiliaries will be, there will be timeless rivals to any theory under those auxiliaries"[26].

Laudan & Leplin's first reply was that, as normally expressed, the claim of EE is an atemporal thesis, so that "the only way to make sense of these assertions [Kukla's observations] is to take this indexed version of EE to mean that the *condition* of empirical equivalence, while perhaps temporary to any pair of theories, is permanently guaranteed for

[24] *Ibid*, 454.
[25] Kukla 1996, 142. See also Kukla 1993.
[26] *Ibid*, 142.

any given theory under changes in the choice of rival"[27]. I think this is a correct characterization of what Kukla states. However, to say that it is only the *condition* of EE what he has rescued as a general feature does not preclude a new way in which the problem arises. That is, EE as a trans-temporal *condition* is enough in order to generate UD once again.

The crucial point in Kukla's challenge lies in the universal extension of the scope of the EE premise. It is the universal quantifier of the sentence what is at stake: "for *any* theory *T* there is a (time-indexed) EE rival". In turn, this universal scope is crucially supported by the effectiveness of algorithms to provide an alternative theory which will be empirically equivalent with respect to any actual theory. If algorithms to produce an EE rival given a theory *T* are effective, it does follow that *any* theory *T* has a time-indexed EE rival Therefore, Laudan & Leplin's attack on this kind of algorithms is essential.

b) Algorithms

In their original paper the approach Laudan & Leplin undertake with respect to algorithms is rather general and more suggestive than conclusive. They only consider some results of formal logic that logical positivists used for other purposes[28], but that have also been understood as algorithms to create EE theories. First, they deal with the Lowenheim-Skolem theorem, which can be understood as showing that if any formal theory has a model at all, then it has infinite possible models. So, in principle, and from a semantic-structuralist point of view, a scientific theory *T* has infinite EE rivals. Laudan & Leplin state that this theorem is not effective in generating EE competitors for two main reasons. There is no guarantee that the potential alternative models could be effectively construable; and even if they were, the semantic requirements for the physical meaning of the terms involved in the alternative models are unlikely to be fulfilled. The remaining candidates for an algorithm, namely, Craig's theorem and Ramsey's sentence, are logical tools that provide *T′*, which is a formulation of a theory *T* in which all of the theoretical terms have been excised. The problem now is that it is quite difficult to consider *T′* as a *different* theory with respect to *T*. Actually, in the original problem in which both the Ramsey sentence and Craig´s theorems were considered, their results were understood as an alternative formulation of the *same* theory possessing the advantage of not making any reference to un-observables. In a way, they are nothing but an anti-realist or empiricist formulation of *T*, not a rival theory. Therefore, neither the Lowenheim-Skolem theorem, Craig's theorem nor Ramsey's sentence are logical procedures that can be used as algorithms that generate an EE rival *T′* given a theory *T*. Accordingly, Laudan & Leplin conclude that

> The algorithm does not produce a rival representation of the world from which the same empirical phenomena may be explained and predicted. On the contrary, a theory's instrumentalized version posits nothing not posited by the theory, and its explanations, if any, of empirical phenomena deducible from it are wholly parasitic on the theory's own

[27] Laudan & Leplin 1993, 9. This article is their reply to Kukla 1993, in which the latter first offered the criticism at issue.
[28] To show the eliminability of theoretical terms.

explanations. A theory's instrumentalized version cannot be a rival to it, because it is a logical consequence of the theory and is bound to be endorsed by anyone endorsing the theory.[29]

In his first reply, Kukla challenges Laudan & Leplin's dismissal of algorithms. He offers two examples which according to him are effective in generating an EE competitor. First, "for any theory *T*, construct the rival *T** that asserts the world to be observationally exactly as if *T* were true, but denies the existence of the theoretical entities posited by *T*"[30]. This example is easily objectionable in terms of its parasitism on *T*. If a theory at all, it is difficult to see how *T** can be a rival to *T* rather than just an antirealist version of it. However, the example that Kukla gives more attention to is the following:

> Take any theory *T* with observational consequences *O*, and construct from it the theory *T′* which says that *T* is true of the universe under the initial conditions that the universe is being observed; but when nobody's looking, the universe follows the laws of *T**, where *T** is any theory which is incompatible with *T*. Clearly, one can find such a theory *T′* for any *T*, and just as clearly, *T′* is empirically equivalent to *T*. QED.[31]

This example, according to Kukla, has the advantage of providing an empirically equivalent theory which cannot be attacked in terms of parasitism. He claims that *T′* is no more parasitic on *T* than what *T* is on *T′*. Laudan & Leplin's reaction is twofold. First, they deny that the only restriction on *T** is its incompatibility with *T*. If the algorithm is to deliver an empirically equivalent theory as its output, the laws in *T** must be such that *T′* may be able to duplicate the predictions of *T*, *T** cannot be any incompatible theory. The algorithm does not include any procedure to construe such a *T**, so that, at best, it is a promissory note to EE. On the other hand, the laws contained in the required *T** will be bizarre in the sense they must assert a mechanism or effect which relate the observer and the universe in a way in which no well-confirmed theory so far has ever posited. That is, *T** is quite likely to be disregarded as being an unjustified theory. There is no rationale to postulate such a peculiar kind of theory. Even worse, if a certain feature of the universe or of other theory about the universe is ever to justify *T**, it is not clear at all that its predictions will be exactly the same as the ones related to *T* –the variability of the observable and of auxiliary hypotheses would hold once again–. In other words, the algorithm is not a complete one, it is only a promissory note; and the required *T**, at least in principle, does not have any justification to be introduced as a competitor to *T*.

However, it is still possible to weaken the algorithm and take it just as stating that the empirically equivalent theory *T′* only asserts that *T* holds while we are observing, but that it does not hold when we are not. If this is the case, then *T′* is so obviously bizarre and ungrounded that it will never be considered as a scientific theory in the first place. But it

---

[29] Laudan & Leplin 1991, 456-7.
[30] Kukla 1993, 4. Concerning this kind of algorithm, Norton states that: "if we assume that the algorithm [construct the theory T' with identical observational consequences as T, but with the negation of all of the theoretical claims of theory T] is applied to a well-formulated theory T whose theoretical structure is essential to T's generation of observational consequences, then the construction of T' amounts to a gratuitous impoverishment of theory T, the denial of structures that are essential to the derivation of observational consequences that are well confirmed by them". Norton 2008, 39-40.
[31] Kukla 1993, 4-5.

could still be considered as an instance of the *evil-genius* argument, as an instance of the fact that, from a merely logical point of view, there are many possible hypotheses consistent with the information of our senses which deny them as providing reliable information about reality. If this is so, the "algorithm" can no longer be considered as the source of a problem of UD for scientific theories, but as an instance –not novel at all– of an epistemological or even metaphysical problem.

c) Theoreticity

I think that Laudan & Leplin's criticism of the considered algorithms is sound. However, what is valuable in Kukla's challenge lies in its demand for a general criterion and a general explanation for the disregard of pseudo-theories. If one argues against algorithms by saying that the hypotheses they produce will not be considered as *scientific theories*, one presupposes a general criterion to decide what hypotheses actually qualify as genuine theories. That is, he demands for an explanation of *theoreticity*:

> It seems to me that the whole philosophical dispute between the received-viewers and Laudan and Leplin comes down to the issue of distinguishing genuine theoretical competitors from logico-semantic tricks. Laudan and Leplin represent the issue as being concerned with the existence or nonexistence of empirical equivalents. But it is evident, both from my example as well from the example they reject in a footnote, that there *do* exist empirically equivalent propositional structures to any theory. The only question is whether these structures fail to satisfy some *additional* criteria for genuine theoreticity. The received-viewers are satisfied with their examples of empirical equivalence. The burden is on Laudan and Leplin to explain why empirical equivalence isn't enough.[32]

In a later article, Kukla claims that there are no clear and consistent criteria of theoreticity available[33]. Accordingly, he concludes that philosophers of science, and scientists themselves, do not possess a satisfying and justified ground to disregard the product of algorithms as pseudo-theories. He ponders several candidates that could play the role of determining when a certain hypothesis qualifies as a theory, but concludes that none of them are satisfactory or firmly grounded.

The first possibility he assesses is 'parasitism', a criterion which we saw Laudan & Leplin endorse. He defines the concept thus: "a putative theory is merely a quasitheory if its formulation necessarily involves a reference to another theory. A parasitic reference must presumably be ineliminable, since every theory trivially has logically equivalent reformulations that refer to other theories"[34]. I think this definition is not precise enough, the kind of ineliminable reference to another theory should be specified. That is, in what sense a theory necessarily refers to another one is a relevant feature in order to decide whether we face a case of parasitism or not. In the case of the algorithms criticized by Laudan & Leplin, namely, Craig's theorem and Ramsey's sentence, we have that the relation involved is that of logical entailment (or even logical equivalence); and that seems to be crucial.

[32] Kukla 1993, 5.
[33] Kukla 1996.
[34] Kukla 1996, 148.

Anyway, Kukla's first criticism of the concept of parasitism asserts that Laudan & Leplin have not shown that a theory *T′*, which states that the observable consequences of *T* are true but that *T* itself is false, could not be alternatively formulated as to circumvent any reference to *T*. This is a rather strange criticism; it is obvious that in this case the burden would be on providing an algorithm capable of doing that. The only two serious alternatives proposed so far have been shown ineffective in that sense, Craig's theorem and Ramsey's sentence.

His second attack states that since a structure such as *T′* –the theory that affirms the observable consequences of *T* but that denies *T*– does play a role in actual scientific practice, a notion of theoreticity that forbids its use would clash against a principle of science. Kukla refers to Daniel Dennett as stating that intentional psychology is accepted by a large piece of the cognitive science community. However, its ontology is mostly rejected because it is incompatible with a physicalist framework. This criticism does not work either. I think that Kukla misunderstands the attack on algorithms in terms of parasitism. Laudan and Leplin are not saying that the outcome *T′* of the algorithm must be dismissed from the outset –as a pseudo-theory– because it is a parasite of *T*. Their point is that the parasitic reference of *T′* to *T* means that *T′* is not a genuine *rival* to *T*, but simply its *instrumentalized* or *antirealist* version. The difference between *T* and *T′* boils down to the epistemic stance one takes towards the very same theory. What Kukla shows is simply that, according to Dennett, in the case of psychology the *instrumentalized* attitude with respect to intentional psychology is more appropriate than the realist one.

Next, he attempts a rejection of 'superfluity' as a criterion of theoreticity. Superfluous hypotheses are those which could be dispensed of without loss of empirical content of the theory. He cites once again the algorithm in which theory *T′′* asserts that the world behaves exactly as *T* says when we are observing, but when we are not it behaves according to *T**, and *T′′* is not parasitic on *T*. His point is that if the superfluity of the assumption that the world does not behave consistently across time is the reason for the disregard of *T′′*, then *T* must also be disregarded insofar as its presupposition of the continuity of the laws of nature is also superfluous. This remark is misconceived. Laws which do not refer to any connection between the presence of an observer and the behavior of the universe –like the laws in *T*– do not need an extra assumption about the continuity of those laws, they just state that what they assert is the case, and in this case what they assert does not include any observation dependency –this must not be conflated with the ontological assumption that the world behaves according to laws, which is not what Kukla is referring to (this is a sort of *epistemological presupposition* that any theory must endorse, otherwise it would be absurd to come up with theories in the first place). Now, if *T′′* is simply understood as claiming that the world behaves differently to what *T* dictates when we are not observing, *without any explanation for why this is so*, we already saw that it reduces to an instance of the *evil-genius* arguments. If *T′′* does include such an explanation, but one which is *irrelevant for the derivation of observational consequences*, it is superfluous and can be disregarded as a genuine theory because of it. But then again, *T′* might include a rationale for why the world behaves according to *T** when no one is observing that is indeed relevant for the entailment of observational consequences –and in that case there would be no place for a superfluity accusation[35]–, but that rationale cannot be provided by the application of the algorithm considered, as I showed above[36]..

Concluding, both parasitism and superfluity are features which justify and specify the concept of theoreticity. Parasitic or superfluous hypotheses are not to be considered as genuine rivals or as genuinely scientific theories, respectively. Of course, one could ask why both features are dismissed by science. The answer to this question requires further reflection about the ultimate nature of science; but is quite obvious that, even under a vague concept of its goals, parasitic and superfluous theories will never be able to achieve them. This connects to one final important remark about theoreticity. Kukla's criticism is made from a perspective in which the concept is required to be *a priori*:

> Even though a hypothesis may possess the traditional empirical virtues of having a truth-value, being confirmable and disconfirmable, and generating indefinitely many testable predictions, it might nevertheless be excluded from serious scientific discourse for failing to satisfy an *a priori* constraint on the proper form for a scientific theory. Let us call this hypothesized property by the name of *theoreticity*.[37]

I think this approach is wrong. Theoreticity cannot be an *a priori* concept. First of all, it depends on a more basic conception of the ultimate nature and the goals of science, and this conception is contingent and historical; so that theoreticity inherits both features. If for whatever reason the typical aims of science are reformulated, the concept at issue could radically change.

Furthermore, even if we take the concept of theoreticity with respect to our particular view of what science is, the features that it requires for a hypothesis to count as a theory cannot be decided in an *a priori* way. To illustrate I will refer to two further properties that Laudan & Leplin assert as related to the disregard of a pseudo-hypothesis, namely, testability and explainability[38]. They describe the former in the following way:

> Because the purpose of theorizing is, at least in part, to gain predictive control over the subject matter under investigation, a theory must, at least in principle, be open to test. A 'propositional structure' that is not even in principle confirmable, that could not logically be an object of

---

[35] Actually, in this case *T'* would 'presuppose' the consistency of the behavior of the world across time as much as *T* does, for the connection between the presence of observers and certain physical phenomena is asserted by the laws in *T'*. This means of course that in neither of the two cases we are facing a superfluous hypothesis. Both *T* and *T'* simply assert what they assert –but they do share the epistemological presupposition of the law-likeness of the world.

[36] Curiously, Kukla does not reply to Laudan & Leplin's remark that the algorithm does not include any method to construe *T**. However, he considers two further possibilities to provide a concept of theoreticity. First, he discards a plain reference to what real scientists actually disregard. Even if the scientific community disregards all of the algorithms, this would not be enough to conclude that the theories they produce are epistemically flawed –I think this is right–. Secondly, he rejects the possibility of an intuitive and implicit conception of theoreticity. He introduces yet another algorithm –for any theory *T* there is a theory *A(T)* which asserts that what we call the universe is a computer simulation wherein events are programmed in order to follow the rules of *T*– and claims that a specific theory of this kind is intuitively intelligible as a genuine one at least to himself and the astronomer John Barrow, the author of the theory–. I have no comments with respect to this argument. See Kukla 1996, 153-9.

[37] Kukla 1996, 146.

[38] The authors do not *label* these properties, they only provide a rough description; but I think that 'testability' and 'plausibility' are accurate nouns for them. They do not offer any reflection about their non-*a priori* nature either; they just introduce what I call 'testability' and 'plausibility' as an answer to Kukla's challenge to clarify the concept of theoreticity.

> epistemic evaluation, is not a theory; for it could not in principle impart understanding nor advance practical interests[39].

If one connects this description with their remark about the instability of the observable and of the auxiliary hypotheses, it follows that for a hypothesis to be testable is a contingent matter, and necessarily connected to a particular state of the development of science. It is true that there are certain hypotheses that not even in principle could be tested; metaphysical statements, for example. However, there are other hypotheses which in spite of a *prima facie* radical non-testability, one could forfeit that in a very different scientific context they might become testable. A mechanism which connects the behavior of the world to the presence or absence of observers in a radical way –as Kukla's T''– is not an *a priori* non-testable hypothesis. A different state of science is conceivable, such that it may be able to predict (or to contribute to predict) observable consequences along with auxiliary hypothesis which have not even been conceived so far[40].

With respect to the property that I call 'plausibility', Laudan & Leplin describe it along the following lines:

> Provisions that fly in the face of what we have good empirical reason to assume must claim some offsetting rationale if they are to be admitted as part of a theory. It would be different if the course of nature were known to exhibit such vast and mysterious ruptures or bifurcations as *T'* envisions [the hypothesis that the behavior of the world changes when we are not looking at it], if natural law did not exhibit isometry, at least.[41]

In order to be considered as a genuinely scientific one, a hypothesis must either be coherent with what our currently accepted and well-confirmed theories state as the regular behavior of the world; or –in the case in which the content of the hypothesis 'flies in the face of what we have good empirical reasons to accept'– be able to be explained or justified by at least some part of the accepted scientific background. This justification or explanation does not need to be a complete and detailed one. It suffices with an offsetting rationale for its tenability in spite of its empirical weirdness; especially if the hypothesis is testable.

It is quite clear that what I remarked about the contingency and variability of what is testable is also the case for plausibility. For a hypothesis to be plausible means that some part of the scientific background of a certain time can provide a minimum rationale for its plausibility. The scientific background knowledge changes as time goes by, sometimes radically and dramatically. Therefore, if a hypothesis is or is not plausible is not an *a priori* matter. Once again, a mechanism stating a connection between the presence of observers and the behavior of the world, in a way that flies in the face of our regular empirically-grounded

[39] Laudan & Leplin 1993, 13.
[40] There is a sense in which *testability* could be an *a priori* requirement in the sense that science, by its very basic definition, demands that property from hypotheses –however, the way in which this property became an essential part of science was contingent, of course. In any case, it is rather clear that if a given hypothesis is testable or not is –in general– a contingent matter.
[41] *Ibid.*, 14.

views, is not an *a priori* unjustifiable hypothesis. Suitable scientific background knowledge could supply a minimum degree of plausibility for it[42].

d) The first premise restated

Now that we are armed with the variability of the observable and of the auxiliary assumptions; plus an analysis of the meaning and cogency of the criteria of theoreticity; it is time to reassess the status of the first premise which leads to the problem under investigation. Its original formulation states that *For any theory T which entails the class of observational consequences O, there is a rival theory T′ that also entails O*. Laudan & Leplin have shown that this is a false statement, and in order to be corrected it must be severely weakened. First, the empirical equivalence relation must be indexed to a specific state of the development of science and technology. Second, since the proposed algorithms to generate an empirically equivalent theory were demonstrated to be defective, the statement cannot be considered neither as universal in its scope nor as assertoric in content. That is, there is no reason to assert that for *any* theory there actually *is* an empirically equivalent competitor. The conditions of theoreticity block the outcome of algorithms because they do not include any procedure to provide testability, plausibility or non-parasitism to the hypotheses they deliver.

Conditions of theoreticity also block Duhem's holist thesis as a basis for universal and actual EE. It is true that, logically speaking, it is always possible to introduce auxiliary hypotheses or sematic tricks to create an empirically equivalent theory, but this logical possibility does not assure at all that the application of these maneuvers will be successful in satisfying the conditions of theoreticity examined[43]. In one of the early reactions to the connection between holism and underdetermination, Adolf Grünbaum stated that neither the holist thesis "nor other logical considerations can *guarantee* the deducibility of *O′* [*O* in our case] from an *explanans* constituted by the conjunction of *H* and some *non-trivial* revised set of A′ of the auxiliary assumptions which is logically incompatible with A *under the hypothesis H*"[44]. Where Grünbaum asserts that there is no guarantee for the entailability of *O′* from suitable *non-trivial* auxiliary assumptions –and we can also add suitable modifications of the hypothesis which result in a rival alternative–; we can understand that nothing guarantees that such maneuvers will be able to fulfill the requirements for theoreticity.

What is crucial to be noticed is that none of Laudan & Leplin's views about the matter preclude that empirical equivalence *could* be the case between two scientific theories. Even though algorithms cannot be invoked, and even though conditions of theoreticity impose constraints that make it a rather unlikely scenario; their arguments do not show that a theory

[42] What I said about the justification of parasitism and superfluity as criteria for scientific disregard of hypotheses – in terms of their relation to the most basic and general goals of science– holds, I think, as clearly in the case of testability and plausibility.
[43] As I mentioned above, Duhem did not connect his holist thesis to UD. Actually, the concept of theoreticity can be considered as a development of his concept of *good sense*. Duhem only introduced his thesis in order to show the fallible nature of falsification. See Laudan 1965.
[44] Grünbaum 1960, 118

T′ which does accomplish these conditions may not be such that it is empirically equivalent with respect to *T* –indexed to a particular state of the development of science and technology, of course. This remark lies at the basis of the reevaluation of the whole problem that I will suggest as my main thesis in this work.

2. The Second Premise

a) Empirical evidence and logical entailment

As I showed above, EE is not enough in itself to prove that UD is the case. It is also needed that *the entailment of evidence is the only epistemic constraint for the confirmation of a theory*. Only if this is the case –and if two theories are equivalent with respect to the class of observational consequences they entail– it follows that the two theories are equally confirmed by the empirical evidence and the problem of a criterion to choose comes up. As I will show now, this second premise is also false and needs to be weakened and corrected.

John Norton offers a very general attack on it[45]. He states that the second premise relies on a hypothetic-deductive (*h-d*) account of confirmation and inductive inference which is quite simplistic and flawed. He offers a classification of theories about confirmation and inductive inference which shows that none of them provides any rationale for the *h-d* view that underlies the problem of EE and UD.

Norton first considers *inductive generalization*, the view that an instance confirms its generalization: the fact that some *A*s are *B* confirms that all *A*s are *B*. One can easily see that this view allows confirmation in cases where the evidence is not entailed by the hypothesis: that this crow is black counts as evidence that the next crow we will observe and that we have not yet observed will be also black; and the latter statement does not entail the former.

The second class of theories of confirmation Norton examines is *hypothetical induction*. This approach is indeed committed to the idea that a hypothesis is confirmed by its consequences. However, this view is too permissive, and Norton shows that many constraints regarding the kinds of entailments to be considered as evidential must be introduced. For example, Norton requires "exclusionary accounts", meaning that *H* not only must entail *E*, it must also be shown that if *H* were false *E* would be very unlikely to be obtained. A second example of the constraints Norton refers to is connected with the conditions of theoreticity. Trivial or superfluous cases of entailments, which can be schematically illustrated with suitable conjunctions, should not be allowed in spite of being a case of logical entailment. It is quite easy to construct an empirically equivalent alternative to theory *T* by defining *T′* as *T & X*, where *X* is any hypothesis no matter how bizarre or unjustified. That is, a bare and rough conception of the *h-d* model of confirmation is completely unsatisfactory, but is this conception underlying the problem of EE and UD:

[45] Norton 2008.

> One sees immediately that this account of confirmation is troubled by an excessive permissiveness […]. It is exactly this permissiveness that renders impoverished hypothetico-deductivism an uninteresting option in scientific practice and in the induction literature. Yet is exactly this permissiveness that the arguments for the underdetermination thesis seek to exploit.[46]

Finally, Norton considers *probabilistic accounts* of confirmation and induction. These views are based on the idea that scientists assign degrees of belief according to probabilistic rules of calculus. The most common and illustrative example is Bayesianism. This line of thought states that the degree of confirmation and belief of a theory depends on the outcome of an algorithm in which the prior conditional probability of the hypothesis –along with the likelihood of the evidence under the assumption of the hypothesis, and the probability of the evidence– works as the input; and the algorithm then delivers the posterior conditional probability of the hypothesis given the evidence. What is important here is that nothing determines that the prior probability assigned to our equivalent rivals *T* and *T′* must be the same. The values of the prior probabilities can be different, and in that case, the evidential import of the same evidence will deliver different values for the posterior probability of the theories –in spite of their empirical equivalence. If for whatever reason a higher prior probability is assigned to *T′*, evidence *e* will offer a higher degree of posterior probability to *T′* than to *T*.

Norton's point does not consist in favoring any particular conception of confirmation. His aim is only to show that under any of the current views about this matter proposed in the philosophy of science, the statement *two theories with the same observational consequences are equally confirmed by any evidence* is false. The main reason is that the entailment of observational statements does not *define* confirmation of a hypothesis, as the presupposed and oversimplified *h-d* view holds.

Laudan & Leplin offer a more detailed and specific argument showing the inadequacy of the conception of confirmation underlying the UD argument. Their general point is that even if we accept that EE holds between two rival theories "the relative degree of evidential support for theories is not fixed by their empirical equivalence". More precisely, they show that "significant evidential support may be provided a theory by results that are not empirical consequences of the theory"[47]. Consequently, even if two theories are empirically equivalent in the sense that they entail exactly the same class of observational consequences, it does not follow that they will be equally confirmed by the evidence.

The first example they mention is just the same case already noted by Norton. Instances of a generalization support each other in spite of not being consequences of one another. More specifically, the underlying principle is that if a theory *T* entails statements *e* and *s* –where *e* is observational and *s* is observational or theoretical–, and if these two statements are logically independent; then if *e* is true it counts as evidence for *T* and also for

[46] Norton 2008, 27.
[47] Laudan & Leplin 1991, 460.

s[48]. Laudan & Leplin observe that one could say that this evidential relationship is based on an entailment anyway: the non-consequential confirmation of *s* by *e* happens *via* a general statement that entails both *s* and *e*. Even if this is conceded, the underlying principle is enough to undermine a simplistic *h-d* view of confirmation:

> Allowing to a statement to accrue indirect evidential support in this fashion already undermines the claim that statements are confirmable only by their empirical consequences. This result alone suffices to establish that the class of empirical consequences of a statement and the class of its prospective confirming instances are distinct.[49]

The second example they provide refers to continental drift theory. They describe it as being committed to the following two hypotheses: H_1) there has been significant climatic variation through the earth's history, so that the climate of all regions differ from their respective climate conditions in former times; and H_2) the current alignment with the earth's magnetic pole of the magnetism of iron bearing-rock in any given region of the earth differs significantly from the alignment of the region's magnetic rocks from earlier periods. This example is illuminating insofar as the huge amount of evidence for H_2 during the 1950's and 60's clearly support H_1 as well, even though they are not logical consequences of each other. The point is that this evidential support occurs because both H_1 and H_2 are consequences of a more general theory, continental drift[50]. The underlying principle of confirmation in this case is simply that if a theory *T* entails two logically independent theoretical hypotheses H_1 and H_2, and if in turn, these hypotheses entail classes of observational consequences E_1 and E_2, then the truth of any member of E_1 will support H_1, *T*, and also H_2 in spite of not being consequentially related to it –the same holds, *mutatis mutandis*, for any member of E_2, of course–[51]. Once again, this principle implies that the class of entailed observational statements of a theory does not define the class of observational statements that can count as evidence for that theory.

The important connection between this analysis of the nature of evidential confirmation with the problem of EE as leading to UD is clear:

> Theoretical hypotheses H_1 and H_2 are empirically equivalent but conceptually distinct. H_1, but not H_2, is derivable from a more general theory *T*, which also entails another hypothesis *H*. An empirical consequence *e* of H is obtained. *e* supports H and thereby *T*. Thus, *e* provides indirect evidential warrant for H_1, of which it is not a consequence, without affecting the credentials of H_2. Thus, one of two empirically equivalent hypotheses or theories can be evidentially

[48] Note that this principle coincides with Hempel's *consequence condition* and *special consequence condition*: "(8.2) *Consequence condition*. If an observation report confirms every one of a class *K* of sentences, then it also confirms any sentence which is a logical consequence of *K*. If (8.2) is satisfied, then the same is true of [...] (8.21) *Special consequence condition*. If an observation report confirms a hypothesis *H*, then it also confirms every consequence of *H*". Hempel 1965, 35.

[49] Laudan & Leplin 1991, 461.

[50] This clearly illustrates Laudan & Leplin's observation that the confirmational support that E_1 gives to H_2 is *indirect* and *essentially depends on* E_1*'s logical relation to T via* H_1. If the theory *T* would not exist, there would be no such confirmation of H_2 by means of E_1, of course. I think that this is what one should understand when Laudan & Leplin state that E_1 confirms H_2, even though it is not one of H_2's logical consequences. What is *directly* confirmed is *T*, and *T*'s confirmational support *flows* to the hypotheses it entails.

[51] This principle coincides with Hempel's *converse consequence condition*: "any prediction obtainable by means of *H* can obviously be established by means of any hypothesis which is stronger than *H*, i.e., which logically entails *H*. Thus while the consequence condition stipulates in effect that whatever confirms a given hypothesis also confirms any weaker hypothesis, the relation of confirmation defined in terms of successful prediction would satisfy the condition that whatever confirms a given hypothesis also confirms a stronger one". Hempel 1965, 36.

> supported to the exclusion of the other by being incorporated into an independently supported, more general theory that does not support the other, although it predicts all the empirical consequences of the other.[52]

According to this line of thought, EE among two theories does not imply that they are necessarily underdetermined by the evidence. The fact that one of them might be connected as a special case or as a logical consequence to a more general theory, whereas the equivalent rival cannot, results in that the evidential support *e* of the general theory also supports the special-case theory –even though *e* is not entailed by the latter–; but it does not support the non-connected rival.

I use the expression *special case*, in addition to a connection in terms of *logical consequence* with respect to the more general theory, in order to make sense of an important remark that Laudan & Leplin make. It is not necessary that the general theory may be a fully developed and very precisely formulated one; it is enough if it is at least a plausible and well-grounded program which has not yet been completely established. In this case, it follows that the connection of the general theory to the more specific one needs not be as strong as logical entailment –moreover, strict logical entailment is hardly, if ever, the case even when a more specific theory is said to be reduced to a larger one–:

> The more general theory via which the evidence supports a hypothesis of which it is not a consequence need not be very precise or specific. For example, the statistical mechanics that Brownian motion supported was more a program for interpreting phenomenological thermodynamics probabilistically than a developed theory.[53]

Before I turn to a revision and evaluation of relevant criticisms against the views just presented, I will briefly refer to a proposal by Richard Boyd that relies on the grounds of the two main arguments that Laudan & Leplin endorse: the holistic nature of confirmation-falsification –along with the variability of the auxiliary hypotheses available–; and the fact that the class of statements that can confirm or falsify a theory is larger than the class of its empirical consequences. His proposal relies on a consideration of the famous argument that Reichenbach offered for the conventional character of the geometry of the world. Boyd's version of the argument states that *F & G* is a theory that claims that the world is governed by a class of forces *F*, and that its spatial features are described by a geometry *G*; whereas *F′ & G′* is a rival theory that states that the world is governed by a class of forces *F′* –a class that contains all the forces in *F* plus a universal force *f′*–, and that its spatial features are described by the geometry *G′* –and the theories are empirically equivalent, of course. This equivalence is supposed to entail that to determine what is the real geometry of the world is a matter of convention and pragmatics, but the argument has also been used as an example of a radical UD produced by the EE between two theories, and this is the sense in which Boyd considers it.

His proposal relies on the *Duhemian* remarks about the entailment of observational consequences, both *F & G* and *F′ & G′* necessarily require auxiliary hypotheses and

[52] Laudan & Leplin 1991, 464.
[53] *Ibid*, 463.

assumptions in order to be able to predict empirical consequences, and those auxiliary statements belong to the rest of the scientific knowledge available. Boyd's point is that this inter-theoretic connection of both the theories with the rest of the corpus of scientific knowledge can break the UD of the choice, even though the empirical equivalence remains:

> Even though "*F & G*" and "*F′ & G′* " have the same observational consequences (in the light of currently accepted theories), they are not equally supported or disconfirmed by any possible experimental evidence. Indeed, *nothing* could count as experimental evidence for "*F′ & G′* " in the light of current knowledge. This is so because the force *f′* required by *F′* is dramatically unlike those forces about which we now know –for instance, it fails to arise as the resultant of fields originating in matter or in the motions of matter. Therefore, it is, in the light of current knowledge, highly implausible that such a force as *f′* exists.
> Furthermore, this estimate of the implausibility of "*F′& G′* "reflects *experimental* evidence against "*F′ & G′* ", even though this theory has no falsified observational consequences.[54]

Boyd connects this argument with a realist defense of 'the consistence of explanations with the rest of the background scientific knowledge' as a methodological and universal rule that leads to higher degrees of verisimilitude in new theories. I think that such a stance goes too far. However, a more modest and appealing conclusion can be extracted here. Given two EE theories, the background knowledge available might be such that it is at odds with some of the core hypotheses in one of the equivalent theories, but completely coherent with the other one. Even though the background knowledge does not break the predictive equivalence between the competing theories, the UD can be broken through the inter-theoretic connections anyway. As Boyd states, the unease between the rest of the well-confirmed theories and components in one of the equivalent theories can work as *indirect empirical evidence* to reject such component. This is the way in which I claimed that Boyd's view relies in Laudan & Leplin's assessment that the class of confirmational statements of a theory is not equivalent to the class of its empirical consequences.

Boyd's proposal is also connected to Laudan & Leplin's view that the variability of auxiliary hypothesis available is an essential feature for the problem of EE and UD. Imagine that the equivalent theories at issue are such that they are both equally coherent with the rest of the scientific corpus. In that case the inter-theoretic connections would not be enough to break the UD of the choice to be made. As science develops, new well confirmed theories will be available, and it might be the case that none of those new theories is able to break the predictive equivalence anyway. However, some of those new well-confirmed theories could be at odds with some core elements of one of the theories, but totally coherent with the other one. If the conflict is big and deep enough, then it could work as indirect empirical evidence to reject the problematic theory and accept the other one. Boyd's view shows that the UD of the choice to be made between EE theories, just as EE itself, is also a time-indexed feature – even if the predictive equivalence remains[55].

[54] Boyd 1973, 7-8.
[55] Notice that a clear connection of Boyd's view and the 'plausibility' requirement of theoreticity can be argued. A theory that becomes problematic with respect to the background knowledge becomes problematic in terms of its plausibility. Moreover, what I just claimed about the possibility of a hypothesis in a theory to *become* at odds with the rest of scientific knowledge –even though there was no such conflict before–, is yet another indication that plausibility is not an *a priori* and everlasting feature of scientific hypotheses.

These observations about the nature of evidential confirmation operate as breaking the link between EE and UD. I think they are quite cogent and convincing. However, it is useful, once again, to pay attention to some criticisms that have been put forward in order to obtain a more clear understanding of what has really been achieved.

b) Hempel's problem

The first criticism I will consider was introduced by Samir Okasha. He notes that Laudan & Leplin's view is committed to two principles of confirmation: 1) if evidence confirms a hypothesis, it also confirms any statement that entails that hypothesis; and 2) if evidence confirms a hypothesis, it also confirms any statement that is entailed by it. These two principles correspond to Hempel's converse consequence condition and special consequence condition, respectively[56]. Okasha claims that, just as Hempel observed, simultaneous commitment to both principles leads to a problematic situation:

> Hempel demonstrated that one cannot, on pain of absurdity, maintain both the special and the converse consequence conditions as ubiquitous constraints on confirmation. The absurdity that results is this: every statement confirms any other one. For consider any statement *S*. Every statement confirms itself, so *S* confirms *S*. By converse consequence, *S* confirms (*S* & *T*), since (*S* & *T*)→*S*. By special consequence, *S* confirms *T*, since (*S* & *T*)→*T*. This result holds for arbitrary *T*, and must therefore be regarded as a *reduction ad absurdum* of the simultaneous use of the special and converse consequence conditions.[57]

Okasha also claims that since this is a result which depends only on the form of the sentences involved, Laudan & Leplin's principle cannot be introduced as a general pattern of confirmation. It could only be argued for by means of historical cases in which the particular circumstances may justify its use. But this is not Laudan & Leplin's approach. They do propose it as a general pattern of the nature of evidential support.

---

[56] As Okasha asserts ([1997], p. 254), Laudan and Leplin's argument can be schematized this way:

i) H_1 and H_2 are EE
ii) $T \Rightarrow H_1$
iii) $T \nRightarrow H_2$
iv) $T \Rightarrow H$
v) $H \Rightarrow e$
vi) $H_1 \nRightarrow e$,
vii) $H_2 \nRightarrow e$,
viii) e,

therefore; *ix)* e confirms T (this requires the *special consequence condition*), and then *x)* e confirms H_1(this requires the *converse consequence condition*); but e does not confirm H_2.

[57] Okasha 1997, 253. Hempel's own formulation of the problem is the following: "But this "converse consequence condition" as it might be called, not reasonable enough, indeed should be it not be included among our standards of adequacy for the definition of confirmation? The second of these suggestions can be readily disposed of: the adoption of the new condition in addition to (8.1) and (8.2) [entailment and consequence conditions], would have the consequence that any observation report *B* would confirm any hypothesis whatsoever. Thus, e.g., if *B* is the report '*a* is a raven' and *H* is Hooke's law, then according to (8.1), *B* confirms the sentence '*a* is a raven'; hence *B* would, according to the converse consequence condition, confirm the stronger sentence '*a* is a raven, and Hooke's law holds'; and finally, by virtue of (8.2), *B* would confirm *H*, which is a consequence of the last sentence. Obviously, the same type of argument can be applied in all other cases". His solution lies on not accepting the converse consequence condition as a general pattern of confirmation, but only in certain cases: only when the relation between the stronger and weaker hypothesis is such that the latter is "essentially a substitution instance of the stronger one". For example, Galileo's law of free falling is a substitution instance of Newton's gravitation law. Hempel 1965, 36-7.

I do not know of any reply to this objection in the literature. However, I think that Laudan & Leplin's view provides the means to solve Hempel's problem. The relevant concept is, once again, theoreticity. We saw that testability and non-superfluity are cogent requirements for the theoreticity of a hypothesis; and it is easy to see that Hempel's problem arises only if we do not take into account constraints like these. An observational statement, once both the special and converse consequence conditions are accepted, confirms any sentence whatsoever only if we allow superfluous or non-testable hypotheses as constituents of a scientific theory. In Okasha's formulation of the problem, T can be an arbitrary statement only if we allow it to be superfluous. If we require that T has to be relevant and necessary for the entailment of observational consequences of the theory it is a part of, then the fact that the evidence confirms T is not problematic at all. Hempel's two conditions, once the theoreticity constraints are considered, do not entail that an observational report confirms any sentence whatsoever, because arbitrary Ts are not allowed to play the game of confirmation in the first place; they must be non-superfluous and testable, for instance.

It could still be argued that the problem holds by means of Hempel's own example: take T to be (H_1 & H_2), where the first conjunct is 'All ravens are black' and the second is 'Hooke's law'. The observational consequences of T would thus be O_T which is defined as (O_1 & O_2), where the first conjunct is the class of observational consequences of H_1 and the second is the corresponding class of H_2. In that case we could say that both hypotheses are testable and non-superfluous –for both H_1 and H_2 are relevant for the entailment of statements in O_T. And if we notice that a theory could be defined as the conjunction of any two theories whatsoever, the problem arises again in the form that any evidence can count as confirmation not for any hypothesis, but at least to any *scientific* hypothesis.

This is true, but I think that this possibility allows us to discover a place in which yet another sound and cogent theoreticity constraint must be introduced. Two different theories are to be conjuncted to form a single one only if by so doing known phenomena are explicated or if novel phenomena are predicted which could not be predicted or explained by means of any of the conjuncted theories on its own. Otherwise, it could be accepted as an admitted maneuver to conjoin string theory with genetics in one single theory, for example, and then facts confirming genetic theory would also confirm string theory. If this was an accepted scientific maneuver, it would follow that discovery of a new gene, associated with a certain behavior or physiological trait of certain organisms, could count as evidence that space-time has 11 dimensions, or something like that. However, it is pretty clear that we gain nothing by creating a theory that conjoins genetics and string theory, for no novel predicted or explained facts are obtained by doing so. Once again, it is a basic conception of the goals of science that supports this requirement of theoreticity: it makes sense to create a theory T by conjoining two previously existing theories t_1 and t_2 only if T explains and/or predicts more than the sum of the predictions/explanations of t_1 and t_2.

Yet another connection between confirmation issues and the concept of theoreticity can be noticed if we consider the following statement that Laudan & Leplin defend: there are cases in which observational consequences entailed by a hypothesis cannot count as evidence for it. Consider the hypothesis proposed by a televangelist consisting in that systematic reading of the scriptures induces puberty in young males. Suppose also a study made on 1000 young males which from the age of 7 were obliged to scripture reading during 9 years, and it was found that in all of them puberty, around the age of 16, did happen. The hypothesis, of course, entails the results of the study. However, nobody would say that this result counts as evidence for it:

> No philosopher of science is willing to grant evidential status to a result *e* with respect to a hypothesis *H* just because *e* is a consequence of *H*. That is the point of two centuries of debate over such issues as the independence of *e*, the purpose for which *H* was introduced, the additional uses to which *H* may be put, the relation of *H* to other theories, and so forth.[58]

I think that a better assessment of this feature consist in that in order to be allowed to play the game of confirmation, a hypothesis must first accomplish the theoreticity constraints; rather than saying that there are cases in which observational consequences do not count as evidence for a hypothesis. Cases like the televangelist are better understood as situations in which the hypothesis is not genuinely scientific. But once a hypothesis has been shown to be genuinely scientific, the truth of its observational consequences will necessarily count as evidence for it In general words, if a hypothesis does not accomplish this set of requirements, it is not allowed to play the game of evidence. After all, if you do not bring a bat, you cannot play baseball.

c) General-encompassing theories

Returning to Laudan & Leplin's principle of evidential support, a second important criticism has been proposed by Sorin Bangu. As we saw, the principle is that given two equivalent rivals we can find a more general theory which encompasses only one of them, and which has its own evidential support. This evidential support is inherited by the encompassed theory but not by the non-encompassed one. However, Bangu notes that

> But this will not do. The supporter of underdetermination can reply that nothing rules out the possibility that another theory T^* exists, such that $T^* \rightarrow H_2$ [H_2 being the equivalent rival which was not originally encompassed by a larger theory]. Moreover, it is possible that T^* is supported by evidence *e* as well.[59]

Since it is possible to find a more general theory T^* to reduce H_2, and since it is possible that T^* may be supported by the evidence *e*, support that the general theory *T* provided to H_1 but not to H_2; then the evidence *e*, which originally operated as giving larger evidential support to H_1 over H_2, could also flow to H_2 via T^* -and then the evidential equality among the hypotheses gets restored. It is true that T^* has to be able to pass all of the theoreticity requirements, but a genuine scientific theory T^* *is* possible:

[58] Laudan & Leplin 1991, 466.
[59] Bangu 2006, 273-4.

> The only constraint imposed on the relation between *T* and *T** is that they behave differently with respect to H_2: *T** entails it, while *T* does not. What evidence supports each of these theories is another matter. So, can two different theories, each entailing different hypotheses, be supported by the same evidence? This is trivially true.[60]

There are interesting subtleties in this argument which clarify what is the correct scope of Laudan & Leplin's solution. First, the fact the both *T* and *T** are possible, but not necessary, has to be remarked. Bangu's objection is not algorithmic, but is quite effective in showing that Laudan & Leplin's solution is not algorithmic either. What I mean is that these authors' view is something like a receipt to find a solution –if EE is the case– for a possible situation of UD. That is, they offer something like the following: if you have two EE theories, *look for a more general theory to satisfactorily and fruitfully encompass one of the two theories, or check if available more general theories in related fields can actually reduce one of them; so that the evidential support of the larger theory can remove the evidential equality*.

Bangu's criticism shows that this method is not a general algorithm, it is just a receipt. For even if its application were successful, *T** is still a possible scenario. A second subtlety in this statement is also interesting and important. If *T** is the case, we have two possible situations regarding UD. First, *T* and *T** are not empirically equivalent, for nothing dictates that the classes of observational statements that might support them must be equal, they only need to be overlapping with respect to *e*. In this case, the solution is simple, the outcome of the competition between *T* and *T** will automatically determine a choice between H_1 and H_2, in spite of their empirical equivalence. Second, *T* and *T** are empirically equivalent, for nothing precludes that their evidential statements could determine equal classes. If this is the case, we have an UD problem again. However, the method of encompassing in a more general theory can be applied once again to solve the issue, not directly between H_1 and H_2, but between *T* and *T**. If it can be so solved, the competition between H_1 and H_2 will also have a winner.

d) The second premise restated

These remarks are useful for a conclusive reassessment of Laudan & Leplin's attack on the second premise of the problem. What they have achieved is to prove that the problem is not definitive. There are features of evidential confirmation of theories which allow us to avoid UD consequences even in contexts where EE is indeed the case. The *h-d* model of confirmation and evidence that the argument presupposes is too simplistic to be adequate, and it is not sensitive to subtleties which are relevant for our problem. The most important one is that general-encompassing theories can provide evidential support by means of true observational statements which are not logical consequences of one of the equivalent theories at issue. This feature of empirical evidence and confirmation makes it possible that the support of two empirically equivalent theories may be different.

---

[60] *Ibid*, 274.

However, and as I showed by examining Bangu's criticism, Laudan & Leplin's solution is not definitive either. That is, even though the evidential issues they refer to can be effective to dissolve UD under certain circumstances, they are not effective in closing, in advance, any possible door for UD to come into science. There is nothing in Laudan & Leplin's argument that precludes the possibility of alternative general theories which can reintroduce the EE between the theories. EE as leading to UD remains a possible scenario.

IV. Remaining challenges

Before I offer a general appraisal of the precise nature and meaning of the solution that Laudan & Leplin's arguments provide, I will tackle three further arguments which some authors have introduced to establish a link between EE and UD. They are not, *prima facie*, grounded on algorithms or in a radical conception of holism regarding confirmation and falsification. Two of them are presented as specific cases of EE which result in UD, while the third refers to *global theories* as a case in which Laudan & Leplin's arguments cannot work. I will show that none of these attempts works either.

a) Van Fraassen's formulations of Newton's theory

In *The Scientific Image*, Bas van Fraassen offers an argument for his constructive empiricism related to EE and UD. He asserts that since Newton's theory can be formulated in many different alternative ways –and such that they are rivals to each other and empirically equivalent–; it follows that there are no empirically-grounded reason to conceive one of them as true. This shows that the reasonable stand is simply to accept that all of those theories are empirically adequate; and if we want to choose one of them the decision can only be made in terms of pragmatic considerations. His own statement of the case is the following:

> Let us call Newton's theory (mechanics and gravitation) *TN*, and *TN(v)* the theory *TN* plus the postulate that the centre of gravity of the solar system has constant absolute velocity *v*. by Newton's own account, he claims empirical adequacy for *TN(0)*; and also that, if *TN(0)* is empirically adequate, then so are all the theories *TN(v)*.[61]

As I said above, I take the problem not as connected to the debate about the plausibility of realism. It is the rational choice between the alternatives what is here at stake[62]. Van Fraassen claims that asserting a specific value to the absolute velocity of the gravitational center of the solar system results in an alternative formulation of Newton´s theory, so that assigning different values to that velocity results in empirically equivalent rivals –and we would then have a case of EE as leading to UD of the choice to be made. A possible solution relies on Laudan & Leplin's attack on the atemporality of the EE thesis:

[61] Van Fraassen 1980, 46.
[62] Anyway, as I comented in footnote 1 following J. Busch, van Fraassen's view is vulnerable to this problem as well.

further development of science, say, electromagnetic phenomena, could lead to diverging predictions when attached to the alternative formulations. Van Fraasen anticipates this reply, and states that the assumption of absolute velocity is so central to *TN* that any electromagnetic feature could be suitably included in any of the possible alternatives so that EE is retained. Laudan & Leplin briefly comment on van Fraasen's position and argue that to do so –to suitably include any electromagnetic features in the alternative Newtonian theories– means to conceive the relativity of motion as a conventional statement rather than as an empirical discovery of Newtonian (and some pre-Newtonian) physics. They prefer the second view, in which further diverging predictions are possible[63].

I think, though, that the right answer to van Fraasen's challenge lies in more basic considerations. Once again, theoreticity requirements, specifically non-superfluity and testability, come to the rescue. It is true that in Newton's own formulation *TN* takes the form of *TN(0)*, for he conceived the solar system in absolute rest. This hypothesis, he recognized, is not testable by any means. It is also true that absolute space is a necessary assumption of *TN*, so that absolute velocity is an intelligible concept. However, the fact that absolute velocity is an intelligible and meaningful concept does not mean that a fully developed formulation of *TN must* include a specific value for it –in the case of the solar system or in the case of any object whatsoever. That is, there is no reason to consider *TN(v)* as a fully satisfactory formulation of *TN*. If that were the case, there would have been a big debate about the value of *(v)* within the scientific community, but that never happened. What did happen was a philosophical debate about the meaningfulness or meaninglessness of *(v)*, but this is a whole different issue which is not related to EE and UD, but to the ontology of space.

The explanation of why *TN(v)* is not a satisfactory formulation of *TN* is that it contains a superfluous and non-testable assumption: a value for absolute velocity. This assumption is not necessary for the prediction of observable consequences –otherwise the theories would not be empirically equivalent. Moreover, as Newton and van Fraasen themselves stated, it is untestable. Actually, 20th century formulations of *TN*, which made use of the concept of space-time, were introduced so that the very concept of absolute velocity was undefined. The *substantivalist* flavor of *TN* with regard to space was contained in the notion of absolute acceleration, which is defined in terms of inertial frames in the space-time formulations of the theory (absolute space *itself* is not needed to define absolute acceleration)[64]. My main point is that a fully developed formulation of *TN* which accomplishes all the requirements of theoreticity does not make any reference to a value of *(v)*, or even to the very concept of *(v)*. Therefore, formulations of *TN* which include reference to it are not genuine rivals of the Newtonian theory. They could be considered as empirically equivalent hypotheses, but not as genuine scientific theories. As Kyle Stanford puts it, the underlying principle in the generation of a theory like *TN(v)*,

> trades in underdetermination for a long-standing philosophical problem, this time in the theory of confirmation: if true empirical consequences of a theory are all that matters to its

[63] See van Fraassen 1980, 47-51; and Laudan & Leplin 1991, 457-9.
[64] Neo-newtonian space-time is described in Sklar 1974, 202-6.

confirmation, then evidence E confirming theory T will equally well confirm theory T+C (where C is any further claim that does not undermine T's implication of E), thus offering spurious confirmation to C itself.[65]

b) The Poincare-Reichenbach argument

Yet another concrete example that has been considered to illustrate the link between EE and UD is given by Poincare's famous argument for the conventional character of the geometry of the world, which was in turn refined by Reichenbach. The reasoning they offered states that the same physical reality can be equally described by two alternative theories which assign different metric structures to it. Theory T is defined as $G+F_0$, and T' is defined as $G'+F_n$, where G is Euclidian geometry, G' a non-euclidian geometry, and F is a *universal force* which contracts all physical objects –in T its value is 0, but in T' is greater than 0–. Further details in both theories can be settled so that they may be empirically equivalent. The point is that if T holds for the world, so does T'. In general words, given a certain physical theory of the world that includes a metric structure, there exists a rival empirically equivalent theory with an alternative metric and a suitable universal force that deforms physical objects and so determines the outcome of any possible spatial measurements.

This argument is relevant in this context given the empirical equivalence between $G+F_0$ and $G'+F_n$. One possible reply to the fact that such equivalence leads to UD is that the argument does not preclude that further auxiliary assumptions added to the theories might remove the equivalence –F might prove to be a detectable non-universal force, for example. That is, it is yet another example that empirical equivalence must be considered as time-indexed. However, I think that even conceding that the equivalence will be everlasting, this example can be discarded as a case for UD.

If the argument is considered as an example, it can be easily attacked. The rivals it considers are not *real* scientific theories. Consider first Poincare's version. It refers to a *two-dimensional* circular world in which temperature takes the place of F. Flatlander's measurement devices get shrunk as they approach the boundaries of their world (and they are unable to measure the corresponding force or to know that it is at work), so that they will measure it as infinite and will think they live in a Lobachevskian plane. Of course, with F set to a suitable value, an alternative theory will indicate that they live in a Euclidean finite disk. Poincare's argument is not an example of two empirically equivalent theories about *our* world, it is a sort of *parable* or a thought experiment designed to prove the conventionality of geometry. The world that the 'theories' describe in the parable is a mental construction, so it can be manipulated in a way such that both $G+F_0$ and $G'+F_n$ correctly describe it, and in a way such that both the 'theories' are in equal standing regarding theoreticity conditions. That is, Poincare's parable is not an example of *real* science. The argument is not enough as to show that *our* world is suitably described by two theories like $G+F_0$ and $G'+F_n$.

[65] Stanford 2001, S5.

If we turn to Reichenbach's refined formulation, the same holds. It is not an example of actual empirical equivalence between two theories. Rather, it is a general epistemological analysis provided to make a similar point:

> *Theorem Θ*: "Given a geometry *G'* to which the measuring instruments conform, we can imagine a universal force *F* which affects the instruments in such a way that the actual geometry is an arbitrary geometry *G*, while the observed deviation from *G* is due to a universal deformation of the measuring instruments." […]
> The theorem asserts that Euclidean geometry is not preferable on epistemological grounds. Theorem Θ shows all geometries to be equivalent; it formulates the *principle of the relativity of geometry*. It follows that it is meaningless to speak about one geometry as the *true* geometry. We obtain a statement about physical reality only if in addition to the geometry *G* of the space its universal field of force *F* is specified. Only the combination *G*+*F* is a testable statement.[66]

Considered with respect to its original target neither Reichenbach or Poincare ever used their arguments in connection with UD. Actually, considered as an explication of the epistemology of geometry –which was both Poincare and Reichenbach's original goal–, the argument does not need to be stated as an example of two competing *real* theories, the mere *logical possibility* of the theories is enough. However, if it is intended to show the connection between EE and UD, it must be considered as an actual example. It is quite clear that neither Poincare's nor Reichenbach's version can be understood in this way.

The only way to relate the argument to our problem is to understand it as an algorithm to generate an empirically equivalent theory. As such, it is immediately clear that it will fall prey to all the criticisms provided to the algorithmic approach. The algorithm would be something like this: *given a space-time theory T, it is possible to construct a rival theory T' by introducing a universal force F with a suitable value as to assign a different metric to the world, while conserving all of T's empirical consequences*. According to the general criticism of algorithms explained above, it is clear that this procedure is not enough to provide a genuine scientific theory. More precisely, *F* cannot be any force whatsoever, it must show a minimum degree of plausibility and it must be testable and relevant to entail consequences, but the algorithm says nothing about how to achieve that. For example, consider the following passage in Stanford's paper:

> While Eddington, Reichenbach, Schlick and others have famously agreed that General Relativity is empirically equivalent to a Newtonian gravitational theory with compensating "universal forces", the Newtonian variant has never been given a precise mathematical formulation (the talk of universal forces is invariably left as a promissory note), and it is not at all clear that it can be given one. (David Malament has made this point to me in conversation). The "forces" in question would have to act in ways no ordinary forces act (including gravitation) or any forces could act insofar as they bear even a family resemblance to ordinary ones; in the end, such "forces" are no better than "phantom effects" and we are left with just another skeptical fantasy. At a minimum, defenders of this example have not done the work needed to show that we are faced with a credible case of non-skeptical empirical equivalence.[67]

This criticism, though, must not be understood as an *a priori* and ultimate rejection of the very possibility of a theory with universal forces and alternative metric with respect to

[66] Reichenbach 1958, 33.
[67] Stanford 2001, S6 (footnote 6).

some actual theory *T*. I think there is no way to *a priori* show that a certain hypothesis will never be able to accomplish the basic theoreticity requirements; but the crucial point is that no algorithm can be introduced to achieve this.

Concluding, the Poincare-Reichenbach argument is not an example of two empirically equivalent theories. The only way to refer to it as a case for EE and UD is to consider it as an algorithm. But this does not work either, for it is an algorithm with no procedure to accomplish theoreticity requirements; so that its output cannot be considered as a genuine theory from the outset.

c) "Total theories"

One last argument in order to establish an intrinsic connection between EE and UD lies in turning the attention from typical theories to what has been labeled as *total theories* or *systems of the world*. If the theories that are considered as empirically equivalent are of this kind, recourse to the variability of the stock of auxiliary assumptions, or to possible encompassing in more general theories, will not be possible in order to break the equivalence or to ground a different degree of confirmation, respectively.

A total theory, insofar as I can deduce from the literature –the characterizations given are not totally clear and coherent with each other–, is a theory that *i)* encompasses all of the scientific knowledge accepted up to the date; and *ii)* is valid for all of the world, that is, its scope ranges along all possible realm of phenomena:

> The thesis of underdetermination of theory by evidence is about empirically adequate total science; it is a thesis about what Quine calls "systems of the world" –theories that comprehensively account for all observations– past, present and future. It is a thesis about theories that entail all and only the true observational conditionals, all the empirical regularities already confirmed by observation and experiment.[68]

EE between systems of the world is considered to be a special challenge because Laudan & Leplin's arguments do not work in this case. Recourse to further scientific development as a possible source of empirical consequences is not possible, since by its very definition a total theory includes the class of all possible auxiliary assumptions. Total theories do not need auxiliary assumptions to entail its observational consequences. By the same token, the totality of the theory implies that the variation of the range of the observable becomes futile in order to remove the equivalence between two total theories. Hoefer and Rosenberg clearly show that the three theses that Laudan & Leplin invoke, namely, the variability of the observable, the intrinsic necessity of auxiliaries in prediction, and the instability of the class of auxiliary assumptions; do not operate as principles that entail the time dependency of EE between theories like these:

> Once we have acquired empirically adequate systems of the world, that is, theories that account for all observable events, any such variability [of the observable] becomes *ex hypothesi* moot.

[68] Hoefer & Rosenberg 1994, 594.

> Either this variability has ceased, or if some remains, this variation is assumed not to lead to conflict with the system of the world in question.[69]

> Given a purported system of the world, no such auxiliaries are either available or needed. The theory will *ex hypothesi* include all the resources needed to derive observations, and no auxiliary theory could be added to increase its observable consequences.[70]

> [The instability of auxiliary assumptions] is inoperative if we consider systems of the world which need no external theoretical auxiliaries. If Laudan and Leplin mean to include factual or initial data here, then of course it can be argued that the fallibility of observation means that instability of presumed initial data is unavoidable. But as we noted above, when one is considering empirically adequate total theories, such potential problems are *ex hypothesi* ruled out.[71]

On the other hand, Laudan & Leplin's second argument concerning the nature of confirmation –that by encompassing a theory into a more general one the class of evidential statements of the former gets enlarged by statements which are not logically entailed by it– cannot work for total theories either, for by definition there are no, and there cannot be, further more-general theories. Therefore, if Laudan & Leplin's arguments fail, the solution of the problem must be looked for somewhere else. One could say that two empirically equivalent systems of the world is a rather unlikely scenario. The requirements of theoreticity remain at work, so that the constraints for a genuine total theory to be empirically equivalent to an already given system of the world are quite severe. However, this unlikely scenario is still possible, so that in spite of theoreticity requirements the problem is nevertheless philosophically relevant.

I think it is true that EE between total theories posits a problem which is not solvable in the same way as EE for local theories. It seems to be the case that if two systems of the world are in fact empirically equivalent, they will be underdetermined. However, total theories are not a real case for UD. The problem lies in more basic considerations: the very concept of a total theory is intrinsically problematic. First of all one might ask whether or not the world admits to be described by at least one total theory. If the answer is no, the problem dissolves. If it is yes, one might now ask whether human science is able to provide at least one total theory or not. If the answer is yes, one could ask if it is possible for us to recognize that a certain theory is indeed a total one or not. I think that the answer to this later question is negative, for it presupposes a definite solution for the problem of induction. We could at least know that a given theory has the *form* of a system of the world, but the problem of induction condemns us to never be able to know whether the theory is empirically adequate –as Hoefer & Rosenberg stated in their definition, a total theory is understood to be empirically adequate for all past, present and *future* phenomena. The last condition is something that we can never know. Even if all the previous questions about the possibility of a system of the world are solved, we can never know of a specific theory if it is total or not. Moreover, I suggest that this last remark preclude us to be able to find an answer, either negative or affirmative, for the previous questions.

---

[69] *Ibid*, 597.
[70] *Loc. cit.*
[71] *Ibid*, 598.

A more precise formulation of the intrinsic problems of the concept of a system of the world has been offered by Samir Okasha. His argument departs from a well known thesis of post-logical-positivistic philosophy of science, namely, that as a consequence of the theory-ladenness of observation, the distinction between theory and observation cannot be traced in absolute terms. The distinction is still cogent, otherwise it would be meaningless to talk about the observational consequences of a theory or of its theoretical terms. But the crucial point is that the distinction is essentially context-dependent. A term might belong to the realm of the theoretical when considered as a part of one theory, but might be understood as observational when in the context of a different theory.

This pragmatic distinction is all what is needed for an argument connecting EE between local or partial theories and UD. But when the argument shifts to the level of total theories, the context-dependency of the distinction undermines its coherency. As we saw, two empirically equivalent systems of the world –insofar as Laudan & Leplin's arguments are ineffective in this level– are to be considered as underdetermined with respect to any possible empirical data. The problem with the argument is that it presupposes an absolute notion of observation according to which 'empirical data' gets defined and determined:

> If this suggestion is to make sense, it must be possible to say of any true statement whether it belongs on the 'theory' side or whether it describes one of the 'phenomena' which have to be saved. Therefore a context-relative theory/observation distinction will clearly not do so. For, as I have shown, it permits one and the same statement to count as theoretical at some times, as observational at others. To give content to the idea that the global theory of the world might be underdetermined by the totality of the empirical data, an absolute theory/data distinction is a conceptual pre-requisite.[72]

The context-relativity of the distinction at issue is not a problem for the partition of the statements of a local theory into observational and theoretical. In the case of global theories, however, there is no further context to make sense of the distinction, so that whether a statement belongs to its observational or to its theoretical part is simply undefined –unless we may have a well-grounded absolute distinction. And if the class of the observational sentences of a global theory is something that cannot be determined, the very notion of a global theory becomes incoherent, let alone the view that it is underdetermined by the observational data:

> If we are even to understand this suggestion, let alone endorse it, we must have a criterion for deciding which side of the divide an arbitrarily chosen statement falls on. But such a criterion is precisely what the minimal, context-relative theory/data distinction does not give us. If that distinction is all we have to go on, we can get no grip on what it means for our 'global theory' to be underdetermined by 'the empirical data', nor indeed on what a 'global theory' is even supposed to be.[73]

---

[72] Okasha 2002, 317.
[73] *Ibid*, 318.

V. Reassessing the solution

Now that an analysis of the most successful solution of our problem has been offered, it is necessary to reassess it considering its effective scope. I will now introduce a precise formulation of what is the status of EE and UD once Laudan & Leplin's arguments have been revised. As I showed, their attack on the first premise, the EE thesis, is twofold. First, by referring to the variability of the observable and of the auxiliary assumptions, they show that EE has to be considered as relative to a specific stage of science, and that it can be broken by later stages. Second, they show that algorithmic procedures to generate EE are ineffective, insofar as their outputs cannot be considered from the outset as genuinely scientific.

Their attack on the second premise, the thesis that observational logical consequences are the only possible source of confirmation for a theory, is false: the class of statements which can count as evidence for a theory is not identical to the class of observational statements entailed by the theory. In the specific context of the problem of EE and UD, the most important instance of this view is the fact that given two empirically equivalent theories, if one of them can be encompassed in a more general theory while the other cannot, the evidential support of the general theory flows only to the encompassed one, so that, in spite of EE, one of the theories receives a larger confirmational support than its rival.

Boyd's argument can be used to find yet another way out of the problem. The inter-theoretic connections of the theories plus the variability of the background knowledge determine that the UD may get broken even though the predictive equivalence remains. Even if the rival theories still predict exactly the same consequences, concepts or entities included in one of them might get at odds with the (new) background knowledge, and the conflict can be serious enough as to ground a decision.

The general evaluation that Laudan & Leplin make is that their arguments are enough to definitively solve the problem at issue: "The thesis of underdetermination, at least in so far as it is founded on presumptions about the possibility of empirical equivalence for theories, stands refuted"[74]. I think that this appraisal goes too far. It is true that they propose an effective solution of the problem; but I think that the nature of the solution provided is not absolute. More correctly stated, they have made explicit that the methodology of science is capable of dealing with cases of EE, so that the resulting UD can be overcome and a basis for a rational choice becomes available.

Concerning the first half of their argument, its result is not that EE cannot be the case. They show that its possibility is restricted, and more importantly, that it cannot be considered as a condition that *any* given theory faces, for the algorithms are not effective. I showed that by a deeper consideration of the conditions of theoreticity, it becomes clear that the constraints for the possibility of EE between two theories are quite strict; and added to the fact that EE must be considered as essentially time-indexed, it is a feature which will count more as an exception than as a rule in the development of science. That is, even after Laudan & Leplin's argument, EE remains a possible scenario.

[74] Laudan & Leplin 1991, 466.

Something similar holds for the second half of the solution. What it really shows is that UD is not a necessary consequence of EE between two given theories. There are principles of confirmation that determine a differential degree of evidence even when EE is indeed the case. However, this argument is not enough to prove that UD cannot be the case. Bangu's criticism shows that in spite of the principle of confirmation Laudan & Leplin invoke UD can happen anyway. The second half of their argument works as a way out of UD, but not as an absolute rejection of its possibility.

However, a correct evaluation of the solution they offer is not that it is nothing but a remark of how unlikely the problem is, even though possible. Their argument contains an implicit recursive methodology that can be applied to solve the problem even when EE and UD are the case. This recursive methodology is something like the following receipt:

> 1) If two theories are empirically equivalent, the scientific community can focus on developing new techniques or instruments, and on research in theoretical fields that could provide new auxiliary assumptions; such that these methods, instruments or auxiliary assumptions could break the equivalence. And (following Boyd) even if the equivalence is not broken, new scientific findings might support one of the theories and reject the other one via inter-theoretic connections.
>
> 2) If two theories are empirically equivalent, the scientific community can focus on the formulation of a more general theory to encompass one of them, so that the evidential tie can be broken.

These two methodological principles are enough to show that even if the possibility of EE and UD does become real, science has resources to face it and to look for a solution which permits a rational choice. It is true that the application of these two principles does not assure success, but this shortcoming gets balanced when one takes into account their recursive nature. That is, if the result of their application does not succeed in breaking the equivalence or the evidential tie, they can nevertheless be applied over and over again. EE as leading to UD is just another challenge for science, but a challenge that it can face and that it has tools to solve, even though success is not guaranteed. In this sense the problem at stake is not a special one, is just another of its endeavors. With this philosophical evaluation of the issue, I will now turn to consider how it works when a real-life case is considered.

CHAPTER 2

Lorentz's Ether Theory and Special Relativity

This chapter is divided in five sections. In the first one I provide a schematic outline of the scientific context which motivated Hendrik Antoon Lorentz to invent his ether theory. This overview helps to grasp a better understanding of Lorentz's scientific work, which I present in section two –from a chronological point of view, and paying special attention to Poincare's amendments and contributions to the ether theory. The third section is devoted to a concise exposition of Einstein's special relativity. In the fourth section I show that the theories at issue are indeed predictively equivalent –but only if the crucial work of Poincare is considered– and that the theories are different and contenders. In the final section I explain and evaluate the reasons that can be invoked to make a choice between Einstein's and Lorentz's theories.

I. Historical-scientific context: the quest for the ether[1]

The historical appraisal of the origin of Einstein's special relativity theory (SR) is a field in which very different interpretations have been provided. The 'textbook view' commonly suggests a close connection between Einstein's motivation to create his theory and what Tetu Hirosige labels as 'the ether problem'[2]. More specifically, the fact that most of the textbook expositions of SR refer to the negative results of the Michelson-Morley experiment of 1887 suggests that Einstein's theory was the final solution to the ether problem by showing its superfluity. Beginning in the 1960's, historians of physics such as Hirosige, Holton, Schaffner and Miller compellingly showed that this is not an adequate historical claim. Einstein's motivations for SR were not intrinsically linked to the ether problem. However, they have also shown that in order to understand –even by contrasting– all of the relevant issues concerning the rise of SR, it is necessary to take a close look at the development of the ether problem. For example, this view allows one to clearly see that Lorentz's theory, in his 1904 version, was a satisfactory solution of the problem. From these general remarks it is obvious that Lorentz's theory is one of the final stages of the quest of the ether, so that an adequate historical and conceptual understanding of the former requires an examination of the latter, examination which I will now undertake.

a) Stellar aberration and the nature of light

[1] This section is based on Hirosige 1976, Schaffner 1972, Darrigol 2005 and Janssen & Stachel 2004.
[2] Hirosige 1976, 3-6.

During the 1720's James Bradley performed astronomical observations set out in order to find stellar parallax. Since the Earth changes its position along its translational motion, the distant stars should change their apparent position in the course of a year. This effect is a function of the ratio between the diameter of the translational motion of the Earth and the distance to the star considered. The last quantity is too big for the parallax effect to be detected by the experimental equipment available to Bradley. However, he observed another kind of systematic change in the apparent position of the star he was looking at. He noticed it could not be stellar parallax since the pattern of this change was a function of the velocity variations rather than positional shifts. More precisely, the effect he observed was proportional to the ratio between the velocity of the Earth in its orbit around the Sun and the velocity of light, that is, proportional to v/c.

This effect was readily explainable in terms of the then prevailing particle-emission theory of light. If on a rainy windless day a person walks covered by an umbrella, the way in which she should hold it in order to stay dry depends on how fast she is walking, on the direction that she is moving, and on the velocity of the falling raindrops. More precisely, the apparent direction of the falling rain depends on the ratio between the person's velocity and the velocity of the raindrops –both velocities considered with respect to the Earth. Analogously, a telescope set out to look at a star has to be tilt, even if the star is right overhead, as an effect of the Earth's velocity along its orbit and the velocity of the light particles entering the telescope.

However, there was a sense in which Bradley's discovery put a challenge for the particle-emission theory of light. The calculations of the velocity of light based on stellar aberration were consistent with the measurements made by Ole Römer in 1670 based on the changes in the periods between successive eclipses of Jupiter's moon Io. He explained those changes in terms of the time it takes for the light to travel from Jupiter to the Earth. The consistency between the value that Römer obtained and the value derived from Bradley's observation of the aberration effect suggested that there is something like *the* velocity of light. This concept is problematic within a particle theory of light, for the measured velocity of the light-particles depends on the velocity of the source with respect to the measuring receiver. Moreover, there was no reason to think that an emitting object only emits streams of light particles with one single velocity. Therefore, the most natural assumption in a particle-emission theory was that light-particles coming from the distant stars should be received on Earth within a wide range of different velocities. From the point of view of a wave theory of light, however, the concept of *the* velocity of light was quite natural, for the velocity of waves only depends on the properties of the medium they move in, not in the state of motion of the emitting body –though the relative motion of the *receiver* of the wave with respect to the medium should affect the value of the velocity measured.

In 1810 Francois Arago tested this assumption. He covered half of a telescope with an achromatic prism to refract the light coming from a star, and aimed it to a star on different dates along a year in order to be sure that the velocity of the star with respect to the Earth was different every time due to the translational motion. He found no associated changes in the patterns of refraction –the rays and their refraction always respected Snell's law, what indicates that the velocity of the incoming light was always the same (with the angle of the light calculated with respect to the apparent position of the star, not with respect to its real position with the aberration effect corrected). Arago concluded, in order to save the phenomena from the view of a particle theory, that stars do emit light with different velocities, but that in order to be perceived by an observer, the ratio between the relative velocities of the source and the receiver must lie within a specific range. This explanation was considered even by Arago himself as highly implausible, so he looked for further opinions.

b) Fresnel vs. Stokes

Arago turned his attention to the wave theory of light to try to find a more suitable explanation for stellar aberration and the results obtained with respect to the velocity of light. By 1815, he knew that Augustine Fresnel was working in that field, so he encouraged him to develop an explanation. In 1818 Fresnel provided a theory based on two main assumptions. The first was the 'immobile ether hypothesis', i.e., that the Earth moves through a stationary ether without 'carrying' it along. This assumption offers a simple explanation in terms of a wave conception of light. If the Earth were to drag some amount of ether along its orbital motion, the consequent motion of the ether would affect the path of light coming from the distant stars, so that the stellar aberration effect would not be expected. But if the ether stays still in spite of the motion of the Earth across it, the light emitted by the stars would follow a rectilinear path, so that the aberration effect follows quite naturally. This simple explanation in terms of a stationary ether had been already introduced by Thomas Young in 1804.

However, Arago's observations showed that the motion of a transparent dense medium did not affect the refraction pattern of light coming from a star when it enters this medium. In other words, his findings imply that the glass lenses of telescopes directed to a star do not alter the path of the incoming light, but it was known that transparent dense mediums, such as glass, refract light and alter its path in a specific angle. Therefore, the aberration effect should be affected depending on the state of motion of the Earth: the aberration pattern should be different if the incoming starlight entered the telescope in different phases of the Earth's translational motion. Arago's experiment showed that this alteration of the aberration effect did not occur. Consequently, Young's explanation only works if we suppose that the telescopes used are hollow. The assumption of the stationary ether was not enough by itself to explain stellar aberration from the perspective of a wave theory of light.

Fresnel's second assumption comes to solve this problem. It states that transparent dense mediums such as glass *drag* a part of the ether within them when moving across it. The glass of a telescope picks up a fraction of the light's velocity coming into it, so that the expected alteration of the aberration effect gets canceled. The quantitative expression for Fresnel's dragging coefficient *f* is $1 - {}^{1}/_{n^2}$, where *n* is the refraction index of the transparent medium.

The physical interpretation that Fresnel proposed for his dragging coefficient was that a moving transparent body, with a refraction index greater than 1, drags along the excess of ether inside it with respect to the density of the ether outside it:

> Following Young, Fresnel assumed that the ether density in a transparent medium was proportional to the square of the medium's index of refraction. For any classical wave, the speed of propagation is given by $\sqrt{T/\rho}$, where T is the tension and ρ is the density [of the medium]. If the tension is assumed to be constant, as Fresnel did, the velocity *c/n* is proportional to $1/\sqrt{\rho}$. Hence, $\rho \propto n^2$. Fresnel further assumed that, in optically dense media, only the ether density in excess of that pervading all space would be carried along by the medium. Let the density outside the medium be ρ and let the density inside be $\rho' = n^2\rho$. On average the ether inside the medium moving through the ether with velocity $\mathbf{v}$ will then move with velocity
> $$\left(\frac{\rho' - \rho}{\rho'}\right)\boldsymbol{v} = \left(1 - \frac{\rho}{\rho'}\right)\boldsymbol{v} = \left(1 - \frac{1}{n^2}\right)\boldsymbol{v}$$ [3]

The introduction of Fresnel's coefficient was very successful in explaining both stellar aberration and Arago's experiment from the standpoint of a wave theory of light. Moreover, it was also capable to explain further phenomena in the context of experiments performed with terrestrial sources of light. Without the coefficient, laboratory experiments on refraction should yield deviations of the order *v/c* with respect to Snell's law, and that deviation should be interpreted as a function of the motion of the Earth with respect to the ether. The introduction of Fresnel's coefficient precludes that deviation, and therefore, any possibility of detecting the motion of the Earth with respect to the ether by this kind of experiments. Many tests of this sort were carried out, and the results were always consistent with Fresnel's theory[4]. Yet another source of empirical support for Fresnel's coefficient was a prediction he made as early as 1818. The coefficient holds for any medium with a value for its refraction index *n* greater than 1. Therefore, if the observations of a star are made with a telescope filled with water, Fresnel's coefficient corresponding to water would cancel the effect of refraction of this element. That is, the water in the telescope should not affect the measured angle of aberration. In experiments carried out in the early 1870s, George Airy confirmed this prediction.

However, the physical interpretation provided by Fresnel himself was not quite satisfactory. Many objections were made against it. Maybe the simplest and deepest one was given by Willhelm Veltmann's experimental results during the early 1870s. It was originally assumed that Fresnel's coefficient presupposed a refraction index *n* as referring to an average frequency of light, but Veltmann found out that the coefficient should be applied individually to each frequency. Since the index depends on the specific color-frequency of light, then in Fresnel's view transparent bodies should drag different amounts of excess-

[3] Janssen & Stachel 2004, 13-4.
[4] See Janssen & Stachel 2004, 12-3.

ether for every color. Objections like this determined the attitude of the scientific community towards Fresnel's theory. Its huge empirical success grounded the view that any optical theory committed to a stationary ether should include the coefficient. All refraction experiments showed that optical phenomena followed the same laws as if the Earth were still with respect to the ether –at least up to first order of v/c. Nevertheless, the true nature of the physical mechanism underlying Fresnel's coefficient was highly dubious. As we will later see, the claim for a satisfactory account had to wait until Lorentz's work.

Beyond the empirical success of Fresnel's theory, there was yet another problematic feature in it which led to the formulation of a rival theory. By the 1840s it was already known, by means of experiments showing polarization effects in light, that the ether should be an elastic solid medium of high rigidity. If light is a polarized wave it has to be a transverse one, and then its propagation medium cannot be a gas or a fluid, for these can only carry longitudinal waves. On the other hand, the extremely high value of the speed of light required the medium to possess a high degree of rigidity. Then, if the ether must have the features stated, how could it be conceived that a massive object such as the Earth moves across it without altering it at all?

George Gabriel Stokes, in 1845-6, proposed an alternative theory which was able to avoid this problem. The main assumption was simply that the Earth drags along the ether that surrounds it as it moves in its orbit. However, his theory demanded a complex explanation for the behavior of light approaching the Earth, for any motion of the ether should affect the path of light, and then the aberration effect could not be explained without further considerations. Following a hydrodynamic analogy, Stokes described the ether as behaving as a rigid solid for high-frequency waves –such as light– but as a fluid for relatively slow massive objects moving across it –such as the Earth. Stokes' ether was a sort of fluid with large viscosity. Being a fluid, it allows the Earth to move through it –but dragging the part which surrounds the planet. Its large viscosity makes it to behave just like a solid body with small shear elasticity and large plasticity. This feature explains, according to Stokes, why it is able to behave as a rigid solid with respect to high frequency waves. This description gives a more realistic model of the ether and of the nature of its interaction with the Earth, but in order to explain stellar aberration Stokes had to assume a complex description of what happens at the border between the immobile ether far from the Earth and the mobile ether which is dragged by the Earth's motion: the ether has to be an incompressible fluid in irrotational motion with a velocity potential with respect to the motion of the Earth, so that the bending of the wave fronts of starlight –when they cross the border between the immobile ether and the dragged one– produces the observed pattern of aberration:

> In Stokes' view, the ether was a jelly-like substance that behaved as an incompressible fluid under the slow motion of immersed bodies but had rigidity under the very fast vibrations implied in the propagation of light. In particular, he identified the motion of the ether around the earth with that of a perfect liquid. From Lagrange, he knew that the flow induced by a moving solid (starting from rest) in a perfect liquid is such that a potential exists for the velocity field. From his recent derivation of the Navier-Stokes equation, he also knew that this property

> was equivalent to the absence of instantaneous rotation of the fluid elements. Consequently, the propagation of light remains rectilinear in the flowing ether, and the apparent position of stars in the sky is that given by the usual theory of aberration[5].

At this point it is relevant to pay attention to yet another empirical test of Fresnel's coefficient. I already mentioned that in most of these tests the coefficient operates as a canceling certain optical features which otherwise would be obtained. That is, its confirmation was mainly related to 'negative' facts. One important exception came along with Hypolite Fizeau's experiment of 1851. Its main objective was to measure the value of the speed of light in the laboratory, rather than by means of astronomical observations. The experiment consisted in a device made out of two connected tubes in which water was made to flow in opposite directions. Fizeau examined the effect of the flowing water on light that was made to pass through it: he wanted to find out what happens when light emitted from the same source was made to pass through water flowing in opposite directions. The effect he found was a shift on the interference pattern of the light rays after passing through the water, a shift whose value was quite consistent with what should be expected on the assumption of Fresnel's coefficient. Fizeau's experiment was thus considered as a more direct and successful test of the coefficient than the ones based on the null effect of the motion of the Earth through the ether in refraction experiments.

This test had effects on Stokes' theory. In order to account for Fizeau's experiment, it needed to include Fresnel's coefficient. But, as we saw, one of its main attractions was that it did not need it in order to explain stellar aberration and the absence of ether-wind effects in refraction experiments carried on terrestrial labs. There were no deviations from Snell's law because the ether surrounding the Earth was dragged, so that they were at rest with respect to each other. Now, in spite of the relative rest between the Earth and the surrounding ether Fresnel's coefficient had to be considered anyway. According to this, Fizeau interpreted his experiment as supporting Fresnel's immobile ether theory over Stokes'. In any case, the latter theory still had the advantage of providing a more reasonable account of the interaction between massive objects, light waves, and a solid ether.

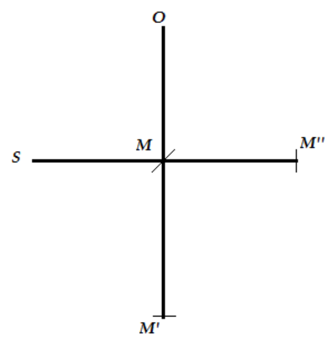
c) The Michelson-Morley experiment

The situation ca. 1860 was then that of a hard competition between Fresnel's and Stokes' theories with respect to the problem of the ether. Fizeau's experiment turned the balance somewhat in Fresnel's favor. However, Stokes' theory had the attraction mentioned in the previous paragraph which compensated its complex explanation of stellar aberration. Besides, Fresnel's theory was empirically very successful, but the mechanism underlying the partial drag coefficient was quite unclear. Therefore, an experiment capable to decide between the theories in a more definitive way was to be most welcome. J. C. Maxwell, in the entry for *Ether* in an edition of the *Encyclopaedia Britannica*, made a suggestion for an

[5] Darrigol 2005, 4-5.

experiment to measure the velocity of the Earth with respect to the ether in a terrestrial laboratory that consisted in looking for variations in the speed of light travelling back and forth between two mirrors. He noticed, however, that the related effects were too small to be measured –of the order of v^2/c^2–, and the alternative method he suggested in order to obtain expected effects of first order of v/c required astronomical data about the periods between eclipses of Jupiter's moons along 12 years, data which by the time were not available with the precision required.

Albert Michelson took the challenge set by Maxwell's suggestion. He designed an 'interferometer', a device two 'arms' perpendicularly connected at *M*. In one of the ends of the arms, there is a source of light *S*, and in the opposite end *M''* of the same arm there is a mirror. In the other arm, in one of the ends there is an 'observer' *O* which measures interference patterns produced by two light beams, whereas in the opposite end *M'* there is also a mirror. In *M* there is yet another mirror, a 'beam-splitter' placed at a suitable angle that partly reflects and partly transmits light. Finally, the distance *MM''*=*MM'*=*l*. If a light beam is emitted in *S* and it is split in *M*, a reflected beam travels back and forth along *MM'*, and another transmitted beam travels back and forth along *MM''*. The two beams meet again at *M* and are transmitted and travel together along *MO*, where the pattern of interference they create is measured:



Suppose that the ether is moving with respect to the interferometer with velocity v parallel to *OMM'* –or that the earth is moving with velocity v across the ether in the opposite direction, of course. In this case the time it takes for the beam traveling along *MM'*, i.e., in the direction parallel to v, is $\frac{l}{c+v}+\frac{l}{c-v}=\frac{2lc}{c^2-v^2}\approx\frac{2l}{c}\left(1+\frac{v^2}{c^2}\right)$. In the case of the light beam travelling perpendicularly to v along *MM''*, its travel time is given by $\frac{2l}{\sqrt{c^2-v^2}}\approx\frac{2l}{c}\left(1+\frac{1}{2}\frac{v^2}{c^2}\right)$. From the comparison between the two expressions it follows that the time it takes the beam to travel in the parallel direction to the relative motion of the ether and the Earth (*MM'M*) is larger than the speed of the beam traveling perpendicularly to that motion (*MM''M*). The value of the time difference is approximately $\frac{l}{c}\frac{v^2}{c^2}$. This time difference multiplied by the frequency *f* of the light used gives the phase difference of the beams which determines the interference pattern measured in *O*. If the frequency *f* is expressed as c/λ –where λ is the wavelength– and is so multiplied by the time difference expression, one obtains $\frac{l}{\lambda}\frac{v^2}{c^2}$ for the phase difference. Even though the quantity v^2/c^2 is minute, the ratio l/λ between the length of the arms and the wavelength of the light used can be made very large. This is why Michelson's interferometer is able to measure an effect of second order.

Of course, the experiment cannot assume what is the direction of motion of the Earth across the ether. Moreover, only *changes* in the phase difference can be observed as changes in the interference pattern at O. For these reasons Michelson's interferometer was designed to be rotated. By rotating it in 90 the roles of the arms get inverted, so that the change in phase difference and the corresponding interference pattern to be observed is twice the amount given in the expression above.

One final important remark about the design of the experiment is that Michelson made a considerable mistake. He calculated a time for the travel of the beam perpendicular to v of $2l/c$, just as if the interferometer were at rest with respect to the ether, instead of $\frac{2l}{\sqrt{c^2-v^2}}$, which is the value that does consider the relative motion between the interferometer and the ether. The result of this mistake was that he miscalculated the time difference between the two trips by a factor of 2, and this overestimation reflected in the value of the interference pattern shift he expected.

Michelson carried out the experiment in Potsdam in 1881. The result he observed was by far within the range of expected disturbances due to the ambient. Michelson himself interpreted it as a refutation of a theory of an immobile ether, for no effect of the motion of Earth was detected, even in the order of v^2/c^2. He also suggested that his experiment could be interpreted as a crucial one and favoring Stokes over Fresnel.

At this point is where Hendrik Antoon Lorentz gets involved in the quest for the ether. In 1886 he published a paper in which he deeply and compellingly criticized the foundations of Stokes' theory. He showed that the assumption of an ether in irrotational motion and the assumption that the ether surrounding the Earth is fully dragged are inconsistent, and he proposed a theory that mixed elements of both Fresnel's and Stokes':

> He made the following assumptions: first, that the ether surrounding the earth is in motion and that this ether has a velocity potential; second, that the motions of the ether and the earth can be different from each other at the earth's surface; third, that when the ether moves through a transparent body, the elementary waves of light in this body are dragged along the direction of the relative motion of the body with respect to the ether with the velocity kv. Here v denotes the relative velocity of the body to the ether, and $k = 1 - 1/n^2$, n being the refractive index of the body. Finally, Lorentz made no assumptions about opaque bodies. With these assumptions and neglecting terms higher than the first order of v/c, Lorentz examined the path of light rays with regard to the earth –the relative rays, as he called them– to show that all phenomena occur as if the earth were at rest and the relative rays followed the path of light rays with regard to the ether, that is, the path of the absolute rays. In other words; except for the Doppler effect, there is no detectable effect of the motion of the earth upon optical phenomena […].
> [Lorentz claims that] If we regard the atoms of matter as a local modification of the ether, we may expect that the ether freely penetrates material bodies however thick they might be. Lorentz considered this problem so important that he urged physicists not to be content with considerations of probability or simplicity, but to decide on the basis of experiment whether the ether at the surface of the earth is at rest or in motion.[6]

[6] Hirosige 1976, 26-7.

Hirosige's summary of Lorentz's 1886 theory (or better, Lorentz's *sketch* of a theory) illuminates some important features. First, it contained an account of Fresnel's coefficient in terms of light rays being carried rather than excess-ether, Lorentz's view provided a new rationale for the physical process underlying the coefficient. I will later show how this was done in terms of electromagnetic considerations. Second, assumptions 1 and 2 imply that Lorentz's definitive response to the question of the ultimate state of motion of the ether remains open and waiting for empirical testing –even though he favored an immobile ether theory– but at the same time the theory offers an account of the absence of observed effects related to this issue. Finally, he offers a rationale for the problem that motivated Stokes' theory, namely, the interaction between massive bodies and the ether. If atoms are considered as a sort of state of the ether, then it is quite natural to suppose that they will move through it without disturbing it (in terms of motion). In this view one can find a seed of the 'electromagnetic view of nature' that Lorentz later endorsed.

Lorentz also put special attention on Michelson's mistake. He stated that the 1881 experiment could not at all be considered as refuting Fresnel and supporting Stokes. Therefore, his 1886 work operated as one of the motivations to repeat the interferometer experiment with increased accuracy. He expected that it would finally show the empirical success of Fresnel's view, fulfilling his will to have an empirically based decision on the state of motion of the ether.

Alfred Potier also drew attention on Michelson's mistake, so that the inconclusiveness of his experiment became blatantly apparent. On the other hand, Lord Rayleigh and William Thomson encouraged Michelson to repeat it, but this time first performing Fizeau's experiment with a higher degree of accuracy. He followed the advice and in 1886, in collaboration with Edward Morley, carried out an improved version of it. The results they obtained strongly confirmed Fresnel's coefficient, and they even took it as a confirmation of the immobile ether hypothesis –the opposite conclusion to the one Michelson had obtained in 1881.

Their next step was, of course, to repeat the interferometer experiment. This time they got the right calculations and designed a much more sensitive and reliable device: it was capable to be rotated in a much smoother way and the light beams were sent back and forth many times along their paths, so that the ratio between l and λ got largely augmented and the expected shift in the interference pattern to be measured increased tenfold. Once again, the result was negative. They repeated the experiment some months later in order to discard the almost fantastic possibility that at the first time the overall velocity of the Earth with respect to the ether had been quite small. But the result was negative as well.

The resulting situation was thus quite dramatic. Stokes' theory had been severely undermined by Lorentz's criticism. Moreover, also in 1887, Hertz succeeded in detecting the electromagnetic waves predicted by Maxwell. This discovery led to the inclusion of optics into electrodynamics, and it turned out that it was very difficult –if possible at all– to incorporate any ether drag in Maxwell's theory and at the same time to have an account of phenomena such as aberration and Fizeau's effect. On the other hand, Fizeau's experiment as carried out by Michelson and Morley strongly confirmed the reality of Fresnel's coefficient, but its physical explanation could not be that of an inner partial ether drag, because of the very same reasons just outlined. Furthermore, the immobile ether thesis to which Fresnel's theory was committed was deeply threatened by the negative result of the Michelson-Morley experiment. Hence, none of the two available alternatives was able to successfully face the radical problem of the ether.

The depth and difficulty of the problem immediately underscores how important the solution that Lorentz later provided was. Totally committed to the unification of optics and electromagnetism that Maxwell's electrodynamics brought, he faced the task of creating a theory under the assumption of an immobile ether capable to offer an account for all the negative results of the experiments so far performed in order to measure the motion of the Earth through the ether. I now turn to this subject.

II. Lorentz's Theory[7]

What at the time was called Lorent'z *Theory of the Electron*, was the outcome of a scientific enterprise that Lorentz started in 1892 (but with its basic roots settled in 1886), and finished in 1904 –though he made later important remarks and revisions until 1916. Therefore, different and progressive stages of its development can be distinguished, and this distinction is quite useful in order to understand his work in a deeper and more accurate way. The stages which I will differentiate in this work (closely following the 'standard view' of the historians who have written about the subject) are two: from the seeds of the theory of 1886 up to the *Versuch* of 1895; and from the formulation of what Janssen calls the 'generalized contraction hypothesis' in 1899, up to its definitive inclusion as a part of the theory of 1904. In between both periods, and after the second, important criticisms and reinterpretations introduced by Henri Poincare must be considered if one is to consider Lorentz's work as predictively equivalent to Einstein's SR of 1905.

a) Stage one: 1886-1895

Hendrik Lorentz's first major scientific work was his doctoral dissertation of 1875. In it he tackled a problem of optics which was first acknowledged by Helmholtz in 1870. Once the luminiferous ether had been depicted as an elastic solid medium, and under Maxwell's analogy between motions in a dielectric and motions in the ether, Helmholtz noticed that the assumption of an ether with those properties implied that, at the limit between two transparent media, the boundary conditions needed to explain reflection and refraction of light were inconsistent to each other. Fresnel's theory, for example, gave the correct formulas at the price of overlooking this problem. Lorentz attempted the task of solving this difficulty from a point of view in which he flirted with the 'action at a distance' approach for charges

[7] My presentation of Lorentz's theory is based mainly on Miller 1998 [1981] and Janssen 1995.

and currents that constituted the mainstream view in continental Europe at the time. The relevance of this early work is that in it, and in his following published paper of 1878 where he dealt with an electromagnetic explanation of dispersion, the roots of the basic ontology of his ether theory got settled: the divorce of ether from matter. That is, since the very beginning of his career, Lorentz was committed to a dualist ontology in which the optic (and electromagnetic) ether was a substance of an essentially different kind than 'regular' matter[8].

After 1878 Lorentz turned to problems of kinetic theory and thermodynamics, but in 1886 he returned to electrodynamics and optic issues. As I mentioned above, in that year he published his *On the Influence of the motion of the Earth on Light Phenomena* –published in Dutch and the following year in French– where he made two very important remarks: that Stokes' theory was founded on inconsistent assumptions, and that Michelson's mistake of 1881 made the experiment completely inconclusive. Therefore, he concluded that the latter was not at all a reliable source of empirical support for Stokes' theory. In that same work he sketched the outlines of a hybrid theory which assumed features both form Stokes' and Fresnel's. However, it was clear that he favored a plan for a definitive theory in which the ether was immobile and fully transparent to the motion of massive bodies: "It seems to me that the latter view [immobile transparent ether] is at least as simple as the former, if not simpler. It may be that what we call an atom is nothing but a modification of the state of this medium; then one could understand that an atom could move without dragging the ether"[9].

Notice that in the last quote a subtle twist in Lorentz basic ontological view is contained. Ether and matter are divorced entities; however, when he writes that *an atom is a modification of the ether*. Here one can already recognize his commitment to an electromagnetic view of nature: the hypothesis that the ultimate nature of reality is electromagnetic. Charges, the ether and electromagnetic forces are the main constituents of physical reality and from them mechanical features emerge.

Yet another relevant feature in Lorentz's publication of 1886 was the derivation of Fresnel's coefficient and a more satisfactory rationale for its underlying physical mechanism. He claimed that the coefficient was not connected to an ether drag, rather, it was the outcome of the interaction between the molecules of the transparent body in which the light entered and the ether surrounding them. That is, his explanation was purely electromagnetic and permitted a conception in which the ether is completely immobile: there is no theoretical necessity for any excess-ether drag or of any kind of partial drag –even though Lorentz's final position about the state of motion of the ether was still open in 1886.[10]

---

[8] For Lorentz's concept of a purely electromagnetic ether, see Nersessiann 1984 and McCormmach 1970b. For Lorentz's work before 1886, see Darrigol 1994.

[9] From Lorentz 1886, quoted in Darrigol 1994, 274-5.

[10] Arthur Miller offers an explanation of the matter and a very interesting remark connected to Einstein: "Lorentz (1886) used Huygens' principle [Miller depicts a nice figure explaining the operation of Huygens' principle in Lorentz's reasoning] and Fresnel's hypothesis to deduce the velocity u_r of light that traversed a medium of refractive index N that was at rest on the earth as $\boldsymbol{u}_r = {}^{c}/_{N} - {}^{v}/_{N^2}$ (Eq. 1.17), where the source could have been either on the earth or in the ether. For N=1, Eq. (1.17) reduced to Eq. (1.13) [$\boldsymbol{u}_r = \boldsymbol{c} - \boldsymbol{v}$, which is the regular explanation for aberration without considering any refraction of light, with $\boldsymbol{c}$ being the velocity of light and $\boldsymbol{v}$ that of the Earth], and for $N \neq 1$, Eq. (1.17) explained Arago's experiment and an equivalent one by George Bidell Airy [see above, p. 41]. Lorentz (1886) continued by noting that from the viewpoint of the geocentric system we could say that 'the waves are entrained by the ether' according to the amount $-\mathbf{v}/N^2$ [...]. On the other hand, an observer at rest in the ether

As it can be noticed, Lorentz work of 1886 was a sort of first draft of a consistent theory that was able to account for up to first order optical phenomena –mostly 'negative' ones– which merged elements both from Fresnel and Stokes. An essential question that he was not yet able to conclusively answer was that of the true state of motion of the ether: "to what degree the ether participates in the motion of bodies that traverse it … is of interest not only for the theory of light. It has acquired a more general importance since the ether probably plays a role in electric and magnetic phenomena"[11]. In spite of his clear sympathy for a completely immobile and transparent ether, which is apparent in his new explanation of Fresnel's coefficient, he considered that this issue was still open and needed to be decided on the basis of empirical data: "In my opinion we cannot permit ourselves to be guided in such an important problem by considerations concerning the degree of probability or simplicity of one hypothesis or the other, but to address ourselves to experiment in order to ascertain the state of rest or motion of the ether at the earth's surface"[12].

The last two quotes clearly explain the motivation of Lorentz's subsequent work. From the first one, one can see that his approach will be that of Maxwell's unification of optics and electromagnetism. From the second one, one can see that the negative result of the Michelson-Morley experiment implied a huge new challenge for the project of an electrodynamical theory able to explain all of the observations regarding the relative motion of the Earth and the ether: the theory should also be able to explain negative results for an experiment of the second order of v/c.

In 1892 Lorentz published two works where he attempted both tasks. The first, *The Electromagnetic Theory of Maxwell and its Application to Moving Bodies* –originally in French– was a big study on Maxwell's theory that included a new term –he later dubbed it as 'local time'– that enabled him to predict negative results for *any* kind of experiments to measure the relative motion between the Earth and the ether, not only for refraction effects tests –up to first order of v/c, though. The negative result of the Michelson-Morley experiment was of course out of the scope of this explanation. Lorentz coped with the special case of second order experiments in a paper entitled *The Relative Motion of the Earth and the Ether*. In it he introduced yet another concept, the hypothesis of a 'length contraction' of bodies moving across the ether that precluded the measurement of second order effects of ether-wind. Later on, in 1895, he published his famous *Attempt of a Theory of Electric and Optic Phenomena in Moving Bodies* –in German– commonly known as the *Versuch*. In it he presented in a more systematic way the results of both the mentioned works of 1892. For simplicity, and

measured the velocity of the light that was propagating through the medium at rest on the moving earth to be $\boldsymbol{c}' = \boldsymbol{u}_r + \boldsymbol{v}$ (1.18). Lorentz (1886) noted that the ether-fixed observer could interpret Eq. (1.18) as the 'entrainment of the light waves by the ponderable matter'. Consequently, although the phenomenon of stellar aberration depended on only the relative velocity between the earth and the star, ether-based theories of optics described it in two different ways depending on whether the source or observer were in motion. […] Einstein considered redundancies of this sort as 'asymmetries which do not appear to be inherent in the phenomena'. In summary Fresnel's hypothesis of a dragging coefficient explained: 1) the dependence of the velocity of light on the velocity of the medium through which it propagated; 2) to first order in v/c, where v was the earth's velocity relative to the ether; optical phenomena were unaffected by the earth's motion". Miller 1998 [1981], 18-20.

[11] From Lorentz 1886, quoted in Miller 1998 [1981], 18.

[12] *Ibid*, 21-3

following Janssen 1995, I will immediately refer to the *Versuch* as the next step in the first stage of Lorentz's theory.

The basic framework on which Lorentz developed his theory in the *Versuch*, and also in his previous work of 1892 on Maxwell's theory, was the assumption that the sources of electromagnetic disturbances in the immobile, non-mechanical, and purely electromagnetic ether, were microscopic charged particles able to freely move through it. On this assumption he presented the following Maxwell equations for the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$:

$$div_0\boldsymbol{E} = \rho/\varepsilon 0, \qquad div_0\boldsymbol{B} = 0, \qquad curl_0\boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t_0}, \qquad curl_0\boldsymbol{B} = \mu_0\rho\boldsymbol{u}_0 + \frac{1}{c^2}\frac{\partial \boldsymbol{E}}{\partial t_0};$$

where $\boldsymbol{E}$, $\boldsymbol{B}$, the charge density ρ, and the current density $\rho\mathbf{u}_0$, are all quantities that are functions of the spatial and time coordinates ($\mathbf{x}_0$, t_0) of a reference system S_0 which is *at rest with respect to the ether* (the subscript 0 in $\mathbf{u}_0$ indicates that $\mathbf{u}$ is a velocity with respect to S_0). Then Lorentz shows that if the Galilean transformations ($\boldsymbol{x} = \boldsymbol{x}_0 - \boldsymbol{v}t_0$; $t = t_0$) are applied to these equations in order to obtain the ones that hold for a system *S in motion with respect to the ether*, then the formulas for $\boldsymbol{E}$ and $\boldsymbol{B}$ in *S* become:

$$div\,\boldsymbol{E} = \rho/\varepsilon 0, \qquad div\,\boldsymbol{B} = 0, \qquad curl\,\boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t_0} + v\frac{\partial \boldsymbol{B}}{\partial x},$$

$$curl\,\boldsymbol{B} = \mu_0\rho(\boldsymbol{u}+\boldsymbol{v}) + \frac{1}{c^2}\left(\frac{\partial \boldsymbol{E}}{\partial t} - v\frac{\partial \boldsymbol{E}}{\partial x}\right).$$ [13]

It is quite apparent that these field equations for a moving frame are not Maxwell's equations for a frame at rest in the ether. This is just another way of saying that the motion of the Earth across the ether should yield observable effects. We saw that Fresnel's coefficient was an explanation for the non-existence of these effects, but only in the case of refraction phenomena, and up to first order of v/c. This is the context and motivation on which Lorentz introduces his famous auxiliary quantity of 'local time'. With the help of this *mathematical tool* –for he did not assign any physical meaning to it– he was able to introduce a new set of transformations for the system in motion with respect to the ether, such that the resulting equations have the same form of Maxwell's equations for the ether-rest frame –if terms of second and higher orders of v/c are neglected. That is, the transformations are such that Maxwell equations become *Lorentz-invariant*.

It is important to remark that Lorentz line of thought involves three different reference frames: S_0, for which Maxwell equations hold; *S*, whose field equations are not Maxwell's; and S', the *auxiliary* frame in which the equations obtained through the Lorentzian transformations hold. This point illustrates more clearly that Lorentz did not assign any kind of physical meaning to his transformations. S' is an auxiliary frame which

[13] The time derivative $\partial/\partial t_0$ of the first set of equations has been replaced by the differential operator $\partial/\partial t - v\,\partial/\partial x$, and in which the velocity $\mathbf{u}_0$ of the first set of equations has been replaced by $\mathbf{u}$+$\mathbf{v}$, where $\mathbf{u}$ is a velocity with respect to *S* –and with v being the velocity of *S* with respect to S_0 in both replacements.

does not reflect measured quantities. Considering this remark, the structure of Lorentz reasoning is thus: on the field equations valid for S_0, Galilean transformations are applied so that the equations for S are obtained; and then Lorentz transformations are applied to the latter so that the Maxwell's Lorentz-invariant equations (up to first order) valid for the frame S' are obtained. The quantitative expression of the transformations is

$$\boldsymbol{x}' = \boldsymbol{x} = \boldsymbol{x}_0 - \boldsymbol{v}t_0, \qquad t' = t - (v/c^2)x = t_0 - (v/c^2)x$$ [14]

As Janssen points out, from a modern point of view the derivation of the field equations for S' is simply a part of a proof that, to first order, Maxwell's equations are invariant under the transformation that Lorentz obtained. From the modern perspective, the quantities $\mathbf{x}'$, t', $\boldsymbol{E}'$, and $\boldsymbol{B}'$; belong to the *Lorentzian* frame S' moving with velocity v with respect to S_0, just as the corresponding unprimed quantities belong to the *Galilean* frame S which moves with velocity v with respect to S_0.

Armed with his 'local time' and the new transformations it permits, Lorentz formulates his famous *theorem of corresponding states*, which explicitly states that the transformations constitute a general proof that, to fist order, no effects of the relative motion among the Earth and the ether will be observed. That is, Lorentz 1895 version of the theorem provides a first order solution for the problem of the ether:

> If there is a solution of the source free Maxwell equations in which the real field $\mathbf{E}$ and $\mathbf{B}$ are certain functions of $\mathbf{x}_0$ and t_0, the coordinates of S_0 and the real Newtonian time, then, if we ignore terms of order v^2/c^2 and smaller, there is another solution of the source free Maxwell equations in which the fictitious field $\mathbf{E}'$ and $\mathbf{B}'$ are those same functions of $\mathbf{x}'$ and t', the coordinates of S and the local time in S.[15]

Since from a modern standpoint, as Janssen points out, what Lorentz did in 1895 is understood in a quasi-relativistic way, one should be careful and subtle about the differences between Lorentz's and the modern conception of the matter. Two remarks show that Lorentz's work was not at all a relativistic theory in the modern sense. First, the fact that the frame S' in which the Lorentz-invariant equations hold is *auxiliary* implies that the explanation of the negative results of the experiments which aim to measure the effects of the relative motion between the Earth and the ether cannot be given in terms of the *measured quantities* in S'. S' is not a physically *real* frame; it results as an auxiliary one from the application of the transformations that, in turn, are mathematical tools. For example, many experiments of the mentioned kind are based on observation of interference patterns –in how they change or in the fact that they do not. The explanation that the Lorentz transformations give, Lorentz-like interpreted, is not that in the moving frame the measurements of the time for the trips of light beams which produce the interference patterns will be the same as the corresponding measurements in the ether-rest frame; but that the *structure* of the patterns in

[14] Janssen 1995, 3.1.1, points out that Lorentz also introduced the fictitious fields $\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}$, and $\boldsymbol{B}' = \boldsymbol{B} - \frac{1}{c^2}\boldsymbol{v} \times \boldsymbol{E}$. The adjective *fictitious* is yet another indication that he considered his new transformation as a mathematical tool devoid of any physical meaning.

[15] This is Janssen's paraphrase of Lorentz's formulation in his 1995, section 3.1.1.

both frames are the same: if at a certain place in one of them there is darkness in the pattern, there will be darkness in the corresponding state of the other frame. In other words, the fact that the theorem of corresponding states implies that *the patterns of light and darkness in the corresponding frames will be the same* is not equivalent to the statement that *the value of each of the measured quantities involved will be the same*. The difference in meaning between these two statements indicates the nature of the difference between the Lorentzian and the relativistic approach:

> Consider some field configuration In S_0 and its corresponding state in S. according to the theorem of corresponding states, the same functions that give the real fields **E** and **B** as a function of the real coordinates $\mathbf{x}_0$ of S_0 and the real time t_0 for the configuration in S_0 will give the fictitious fields **E′** and **B′** as a function of the coordinates $\mathbf{x}=\mathbf{x}'$ of S and the local time t' for the corresponding state of that configuration in S. Suppose the configuration in S_0 is such that at a point P with coordinates $\mathbf{x}_0=\mathbf{a}$ it is dark. That means that the fields **E** and **B** vanish at this point, not just at one instant, but over a stretch of time that is long compared to the period of the light waves described by the fields **E** and **B**. It follows that the fictitious fields **E′** and **B′** [see note 14] will vanish at $\mathbf{x}=\mathbf{a}$ in the corresponding state in S. Since the relation between the real and the fictitious field is linear, this means that the real fields **E** and **B** in the corresponding state will also vanish at $\mathbf{x}=\mathbf{a}$. It follows that the patterns of light and darkness in the moving frame and the patterns of light and darkness in the frame at rest are the same.[16]

The modern reader might complain that the relativity of simultaneity plays a role in the issue and that it defies the possibility of this Lorenztian explanation. It is true that, in a relativistic explanation, the relativity of simultaneity underlies the possibility of an explanation in terms of the identity of the measured quantities involved. However, the patterns of light and darkness are such that they are meaningful from the perspective of time periods which are large compared to the periods of the waves used. Therefore, the fact that the sets of simultaneous events in both frames are different is harmless for the Lorentzian explanation[17].

The second important remark about the difference between the relativistic and the Lorentzian standpoint is that Lorentz did not conceive his transformations as symmetric operators. That is, to 'go back' from S' to S_0, the transformation to be applied is not a symmetrical Lorentz transformation, but its 'inverse function'. This is yet another indication that the real measured quantities belong to the system S_0, and that the coordinates and fields in S′ are mere auxiliary quantities which are the outcome of the application of a

[16] Janssen 1995, 3.1.2. This is the only place in the literature that I went through in which this subtle and important remark is underlined. Janssen refers to brief hints of it in McCormmach 1970b (p. 471) and in Darrigol 1994 (p. 288). In the first case it is just a sentence that can be so interpreted after acknowledging the issue. In the case of Darrigol, he gives an explanation of the null results of the experiments which is quite similar to the one that Janssen offers. However, he does not stress the subtle but important difference with respect to a relativistic point of view.

[17] "the stationary nature of patterns of light and darkness plays a crucial role in this argument. Without this property, the *x*-dependence of local time would lead to serious complications. Suppose that in two points P_0 and Q_0 of S_0, the fields vanish with respect to the real Newtonian time. In the corresponding points P and Q of the moving frame S, the field then will vanish simultaneously with respect to the local time. Since the local time depends on x, this means that they will *not* vanish simultaneously with respect to the real Newtonian time. This would invalidate Lorentz's conclusion with regard to patterns of light and darkness. Fortunately, patterns of light and darkness, by their very nature, are stationary situations. The concepts of light and darkness only have meaning on time scales that are large compared to the periods of the light waves used. So, when at P and Q it is dark at the same instant in local time, it will also be dark at both points at the same instant in real time". Janssen 1995, 3.1.2.

mathematical tool. I will show below that it was Poincare, and of course Einstein, who introduced the right relativistic interpretation.

Coming back to Lorentz theory itself, we have that yet another feature which illustrates its big importance, in the context of unified electrodynamics and the problem of the ether, is given by the formal and general derivation of Fresnel's coefficient. Consider a medium with refractive index n at rest in the ether in which a plane wave propagates with velocity c/n along the x-axis of a frame which is also at rest with respect to the ether. The components of the fields which describe this wave depend on x and t through the expression governing the phase of the wave $t - \frac{x}{c/n}$. Therefore, in its corresponding state, i. e., in a system in motion across the ether with velocity v in the x-direction, the components of the auxiliary fields which constitute the wave depend on the expression $t' - \frac{x}{c/n}$. Considering the Lorentz transformation for time $t' = t - (v/c^2)x$, then the auxiliary fields depend on t via:

$$t - (v/c^2)x - \frac{x}{c/n} = t - (v/c^2 + n/c)x$$

from this expression it can be inferred that the velocity of the wave with respect to the moving medium is:

$$\frac{1}{\frac{v}{c^2} + \frac{n}{c}} = \frac{\frac{c}{n}}{\frac{v}{nc} + 1} \approx \frac{c}{n}\left(1 - \frac{v}{nc}\right) = \frac{c}{n} - \frac{v}{n^2}$$

and in order to obtain the formula for the velocity of the wave, *with respect to the rest frame in the ether*, the velocity v of the moving frame must be added, so that:

$$\frac{c}{n} - \frac{v}{n^2} + v = \frac{c}{n} + v\left(1 - \frac{1}{n^2}\right)$$

The velocity of the wave with respect to the ether is then ${}^{c}\!/_{n} + v\left(1 - {}^{1}\!/_{n^2}\right)$, in agreement with Fresnel's coefficient $\left(1 - {}^{1}\!/_{n^2}\right)$ [18]. It is important to compare this derivation with the electromagnetic treatment that Lorentz gave to the issue in 1886. This time the derivation does not refer to any specific electromagnetic assumption, and this is why it can be considered as more general. The crucial factor is now of course *local time*. That the Fresnel coefficient can be obtained in this purely mathematical way from the Lorentz transformations is yet another case in which it is apparent that the theorem of corresponding states is a fundamental (first order) solution for the problem of the ether. The fact that refractive phenomena will not produce any observable features revealing the relative motion of the Earth and the ether is just a specific consequence of the theorem. This remark clearly

[18] This derivation is from Janssen & Stachel 2004, 25. Janssen also offers a slightly different one in his 1995, section 3.1.3, along with a derivation of the classical expression for the Doppler effect and of the classic formula for the aberration effect.

indicates the generality and unification-power of the theory that Lorentz was attempting. Once again, and in modern terms, it was a theory whose aim was to obtain Lorentz-invariance for Maxwell equations. Lorentz's *Versuch* of a theory of 1895 achieved this goal up to first order of v/c.

Unfortunately for Lorentz, the Michelson-Morley experiment was designed to measure effects of second order. Therefore, it was out of the scope of his theorem of corresponding states of 1895. That part of his theory was incapable of providing an explanation of it. With this in mind[19], Lorentz in his *Versuch* returned to the length contraction hypothesis that he had introduced in 1892. We already saw in the analysis of the Michelson-Morley experiment that the time required for a light ray to travel back and forth along one of the arms of the interferometer in the direction parallel to the direction of its motion through the ether is $\frac{l}{c+v} + \frac{l}{c-v} = \frac{2lc}{c^2-v^2}$, whereas the corresponding time required for the ray traveling perpendicularly to the interferometer's motion through the ether is $\frac{2l}{\sqrt{c^2-v^2}}$; and the different value of these expressions yielded a change in the interference pattern produced as the device was rotated. Lorentz assumed that as a body moves through the ether it gets contracted by a factor $\sqrt{1-v^2/c^2}$, so that in the first expression for the travel-time of the light ray the length l becomes $l\sqrt{1-v^2/c^2}$, and therefore the whole expression becomes $\frac{l\sqrt{1-v^2/c^2}}{c+v} + \frac{l\sqrt{1-v^2/c^2}}{c-v} = \frac{2l\sqrt{c^2-v^2}}{c^2-v^2} = \frac{2l}{\sqrt{c^2-v^2}}$. It is clear that Lorentz's contraction hypothesis implies that the travel time for both rays is the same, and in this case the Michelson-Morley experiment yields a null result.

This hypothesis was clearly introduced in order to account for one particular experiment, but at least Lorentz provided an argument to make it plausible. In 1892 he had already shown that the electromagnetic forces $\boldsymbol{F}'$ in a frame in motion with respect to the ether and the electromagnetic forces in a rest system with respect to the ether are related in the following way:

$$\boldsymbol{F}'_{x'} = \boldsymbol{F}_x \qquad \boldsymbol{F}'_{y'} = \frac{\boldsymbol{F}_y}{\sqrt{1-v^2/c^2}} \qquad \boldsymbol{F}'_{z'} = \frac{\boldsymbol{F}_z}{\sqrt{1-v^2/c^2}};$$ [20]

where the primed coordinates belong to the moving frame across the ether and the unprimed ones belong to the frame at rest in the ether. This result can be interpreted as stating that if a system at rest in the ether is in equilibrium under a configurations of forces $\boldsymbol{F}$,

[19] The following quote clearly illustrates Lorentz's concern about the issue: "Fresnel's hypothesis, taken conjointly with his coefficient $1-1/N^2$, would serve admirably to account for all the observed phenomena were it not for the interferential experiment of Mr. Michelson, which has, as you know, been repeated after I published my remarks on its original form, and which seems decidedly to contradict Fresnel's views. I am totally at a loss to clear away this contradiction, and yet I believe if we were to abandon Fresnel's theory, we should have no adequate theory at all, the conditions which Mr. Stokes has imposed on the movement of aether being irreconcilable to each other. Can there be some point in the theory of Mr. Michelson's experiment which has as yet been overlooked?" From a letter to Lord Rayleigh dated August 18, 1892 –shortly after finishing *The Electromagnetic Theory of Maxwell and its Application to Moving Bodies*. Quoted in Miller 1998 [1981], 27-8.

[20] Actually, Lorentz's expression was $\boldsymbol{F}'_{y'} = \boldsymbol{F}_y(1+p^2/2V^2)$, where p=v and V=c; and to first order of v/c, $(1+p^2/2V^2) \cong 1/\sqrt{1-v^2/c^2}$. The same holds for $\mathbf{F}'_{z'}$, of course. A review of how Lorentz derived this result is presented in Janssen 1995, section 3.2.5. See also Miller 1998 [1981], 26-9.

then that same system, if in motion across the ether, is in equilibrium under a configuration of forces ***F'***.

He then assumed that what he called 'molecular forces' determine the shape and length of a body, and that these forces act by intervention of the ether; but he also stated that the nature of molecular forces was totally unknown, so that his assumption was not directly assessable. However, if it is assumed that the molecular forces behave just as the electromagnetic forces do, then the length-contraction obtains. That is, if a system at rest in the ether is in its 'equilibrium shape' under a configuration of molecular forces $\boldsymbol{F}_m$, then the same system when in motion across the ether will be in 'equilibrium shape' under a configuration of molecular forces $\boldsymbol{F}_m'$. And if the transformation from $\boldsymbol{F}_m$ to $\boldsymbol{F}_m'$ is the same as the transformation for electromagnetic forces, then when at motion in the ether the system gets contracted[21]. Lorentz did not offer this reasoning as a *proof* of the contraction hypothesis, but only as a plausibility argument for it: "one may not of course attach much importance to this result; the application to molecular forces of what was found to hold for electric forces is too venturesome for that"[22].

The hypothesis of length contraction as an explanation for the null result of the Michelson-Morley experiment had been already introduced by G. F. Fitzgerald in 1889. It is interesting to take a look at the way in which he justified the hypothesis. The plausibility argument that Fitzgerald offered was based upon Oliver Heaviside's 1888 discovery that the electromagnetic field around a moving charge gets shrunk along its direction of motion. The quantitative expressions that he determined for this effect were:

$$E = \frac{q}{r^2}\frac{(1-v^2/c^2)}{(1-v^2/c^2\, sin^2\theta)^{3/2}} \qquad\qquad H = Ev\sin\theta,$$

where E is the electric field directed radially outward of the charge, H the magnetic field in circles centered around the line of motion, q is the charge, v is the velocity through the ether, c the speed of light, r the distance of the charge to a point, and θ the angle to the line of motion. With respect to this result Hunt comments:

---

21 "Let A be a system of material points carrying certain electric charges and at rest with respect to the ether; B the system of the same points while moving in the direction of the x-axis with the common velocity p through the ether. From the equations developed by me, one can deduce which forces the particles in system B exert on one another. The simplest way to do this is to introduce still a third system C, which just as A, is at rest but differs from the latter as regards the location of the points. System C, namely, can be obtained from a system A by a simple extension by which all dimensions in the direction of the x-axis are multiplied by the factor $(1 + p^2/2V^2)$ [see note 20] and all dimensions perpendicular to it remain unaltered.
Now the connection between the forces in B and C amounts to this, that the x-components in C are equal to those in B whereas the components at right angles to the x-axis are $1 + p^2/2V^2$ times larger than in B.
We will apply this to molecular forces. Let us imagine a solid body to be a system of material points kept in equilibrium by their mutual attractions and repulsions and let system B represent such a body whilst moving through the ether. The forces acting on any of the material points of B must in that case neutralize. From the above, it follows that the same cannot then be the case for the system A whereas for system C it can; for even though a transition from B to C is accompanied by a change in all forces at right angles to the axis, this cannot disturb the equilibrium, because they are all changed in the same proportion. In this way it appears that if B represents the state of equilibrium of the body during a shift through the ether then C must be the state of equilibrium when there is no shift. But the dimensions of B in the direction of the x-axis are $(1 - p^2/2V^2)$ times the corresponding dimensions of C whereas the dimensions along right angles to the x-axis are the same in both systems. One obtains, therefore, exactly an influence of the motion on the dimensions equal to the one in which, as appeared above, is required to explain Michelson's experiment". From *The Relative Motion of the Earth and the Ether* (1892), quoted in Janssen 1995, section 3.2.6.

22 From *The Relative Motion of the Earth and the Ether* (1892), quoted in Miller 1998 [1981], 29.

> Note especially the $(1 - v^2/c^2)$ factor and the way the field lines bunch up around the 'equator' as the speed increases. This compressed field is in fact the same as the Fitzgerald-Lorentz contraction of the electrostatic field of a charge at rest, in exact accordance with Einstein's theory of relativity. The surface of electrical equilibrium, called a Heaviside ellipsoid, is an oblate spheroid contracted along the line of motion by a factor of $\sqrt{1 - v^2/c^2}$, although it was not until 1892 that this fact was fully clarified by Heaviside's friend G. F. C. Searle. All of this follows directly from Maxwell's equations and shows quite clearly that 'relativistic' effects were already implicit in Maxwell's theory. […]
> Fitzgerald replied [to Heaviside] that he was 'very glad to hear that you have solved completely the problem of the moving sphere' and remarked that, as the formula suggested, the velocity of light might be a physical limit to speed. He also mentioned the possible application of Heaviside's work to 'a theory of the forces between molecules', indicating that Fitzgerald already thought that intermolecular forces might be essentially electromagnetic. Indeed, since he believed that all physical forces, as well as matter itself, arose from the various motions of a single ether, Fitzgerald regarded Heaviside's formula for how electromagnetic forces varied with a velocity as a valuable guide to how other forces were likely to be affected by motion through the medium.[23]

In 1889 Fitzgerald made his insight concrete and proposed a length contraction factor of $\sqrt{1 - v^2/c^2}$ in a paper published in the journal *Science*[24]. This journal was a rather obscure one by the time, so the hypothesis did not have an immediate impact in the community. It only became more prominent when Lodge referred to it in his 1892-3 publications. It was via these works that Lorentz got acquainted with it, and in the *Versuch* he mentioned that Fitzgerald had independently arrived at the same result he obtained. The fact that both Lorentz and Fitzgerald introduced the same length contraction hypothesis, and the fact that they both justified it in the same way reinforce Lorentz's view (and also Fitzgerald's, of course) that from an electromagnetic view it was a rather plausible physical feature.

A second important remark about the Lorentz-Fitzgerald contraction consists in that Lorentz immediately noticed that a longitudinal contraction was not the only possible dynamical explanation for the null result of the Michelson-Morley experiment. A transverse dilation of bodies when moving across the ether in the suitable amount, or a combination of both effects, would also do[25]. However, by 1904 he got committed to a purely longitudinal effect for reasons connected to his model of the electron[26]. Moreover, in 1905-6 Poincare stated that for theoretical reasons –consistency with his 'relativity principle' and the mathematical properties of the Lorentz transformations– the effect should be a purely longitudinal contraction. I will return to this issue below.

Finally, it is important to underscore a feature of Lorentz's theory which can be a source of confusion. In his *The Electromagnetic Theory of Maxwell and its Application to Moving Bodies* of 1892, Lorentz considered a set of coordinate transformations that included a spatial one in addition to the temporal one that in 1895 he called 'local time'. However, it is quite clear that the former transformation was not an expression of the length contraction

---

[23] Hunt 1988, 71-2. He explicitly argues that Fitzgerald might have arrived to the contraction hypothesis even without knowing about the Michelson-Morley experiment, or even if it had never been performed. I remain neutral about this thesis, but his review of how Fitzgerald conceived the hypothesis is quite interesting in connection with Lorentz.

[24] "The length of material bodies changes according as they are moving through the ether or across it by an amount depending on the square of the ratio of their velocities to that of light. We know that electric forces are affected by the motion of the electrified bodies relative to the ether, and it seems a not improbable assumption that the molecular forces are affected by the motion and that the size of bodies alters consequently". From Fitzgerald's *The Ether and the Earth's Atmosphere*, quoted in Hunt 1988, 75.

[25] On this issue, see Brown 2001.

[26] By 1895, it was clear that Lorentz sympathized with an only-longitudinal-contraction effect, but he left open the possibility of the mentioned alternatives. In 1904 his position became more definite and definitive.

hypothesis, but a purely mathematical tool connected to his quest of invariance for the Maxwell's equations, in the same sense that 'local time' was. Miller is very clear in this respect: after applying the Galilean transformation to the Maxwellian wave equations, Lorentz remarked that they no longer had the same form, so that in order to obtain invariance, he

> proposed an additional coordinate transformation on the inertial coordinates (x_r, y_r, z_r, t_r) in order that the Eq. (1.42) [the wave equation for the system S_r in motion through the ether which results of the application of the Galilean transformations] possessed the proper form of a wave equation […]:
>
> $$x' = \gamma x_r \qquad y' = y_r \qquad z' = z_r \qquad t' = t - (v/c^2)\gamma^2 x_r \,;$$
>
> where $\gamma = 1/\sqrt{1 - v^2/c^2}$. (I called the primed reference system Q')[27]. Lorentz considered the transformation from S_r to Q' as a purely mathematical coordinate transformation –for example, he introduced x' as a "new independent variable", and similarly for t'.[28]

It is clear that in this context γ has nothing to do with length contraction. Moreover, these transformations, including γ, were used by Lorentz in his derivation of the electromagnetic forces that hold for a frame in motion with respect to the ether and that underlie his plausibility argument for the length contraction hypothesis. The question is then why γ was not included in the transformations that ground the theorem of corresponding states of 1895. Miller suggests that:

> Whereas a Galilean transformation from S to S_r failed to yield a proper wave equation, a further transformation from S_r to Q' resulted in a wave equation for a disturbance that depended on the emitter's motion, thereby violating an ether-based wave theory of light. Although Lorentz did not comment explicitly on this result for Q', we can assume that he noticed it because he wrote that calculations in the remainder of (1892a) [*The Electromagnetic Theory…*] were only to first-order accuracy in v/c, because this approximation facilitated further calculations, and it led to a "*theorème générale*." To first order in v/c the equations for the electromagnetic field quantities of the molecules constituting matter had the same form in S as in a reference system connected with S_r through the equations:
>
> $$x' = x_r \qquad y' = y_r \qquad z' = z_r \qquad t' = t - (v/c^2)x_r$$
>
> […] Hence, to order v/c the mathematical coordinate system Q' becomes in its spatial coordinates identical with the spatial Galilean coordinates, and the time coordinate mixes the Galilean absolute time $t_r(=t)$ with the Galilean spatial coordinate x_r.[29]

In other words, the coordinate transformation for x that included γ, considered only as a mathematical tool, yielded a wave equation dependent on the state of motion of the emitter. That problem would be solved by adding the length contraction hypothesis to the γ-including transformation for x, but Lorentz took that step only in 1899, as I will show below.

I said that this feature of Lorentz's work of 1892 can be confusing because of the example of Zahar 1973[30]. He seems to understand that the γ of *The Electromagnetic Theory* is interpreted as the contraction factor in 1895 –and also in *The Relative Motion of the Earth and*

---

[27] Notice that Lorentz's 'three-step method' is at work. S is the rest ether frame in which Maxwell's equations hold, S_r is a frame in motion with respect to the ether and in which the equations that hold is the result of the application of Galilean transformations, which are not invariant. Q' is the auxiliary frame in which the equations of S_r have been Lorentz-transformed and that are Lorentz-invariant.
[28] Miller 1998 [1981], 26.
[29] *Ibid*, 27.
[30] See pages 111-2.

the Ether of 1892. This leads to an interpretation in which γ and the contraction hypothesis get conflated; but I think that Miller is very clear in that they are two very different things: the former is a mathematical tool, a coordinate transformation; whereas the latter is a physical hypothesis. They will only get more closely connected by Lorentz in 1899 and 1904, even though remaining logically independent. Actually, the fact that the *Versuch* offered two different and disconnected explanations for the null result of ether wind experiments of first and second order of v/c was the aim of a criticism that Poincare made about Lorentz's theory. Now I turn to it and to some other observations that the French scientist and epistemologist introduced with respect to Lorentz's work.

b) Interlude: Poincare

Henri Poincare got involved in the development of Lorentz's work by underscoring that the explanations for the negative results of ether-wind experiments of first and second order it provided were two different and disconnected parts of the theory. His dissatisfaction about it was grounded on his view that physical science should be built upon certain principles that might be respected. In this case the relevant one is his *principle of relativity*,

> according to which the laws of physical phenomena must be the same for a stationary observer as for an observer carried along in a uniform motion of translation; so that we have not and cannot have any means of discerning whether or not we are carried along in such a motion.[31]

Based on this principle, Poincare believed that the result of any ether-wind experiment should be negative –a result that he did not qualify as surprising– and that the explanation for it must be based on the very core of a physical theory rather than on a compilation of different hypotheses and assumptions. It was in this sense that he criticized the structure of Lorentz's theory:

> I must explain why I do not believe, in spite of Lorentz, that more exact observations will ever make evident anything else but the relative displacements of material bodies. Experiments have been made that should have disclosed the terms of the first order; but the results were nugatory. Could that have been by chance? No one has admitted this; a general explanation was sought, and Lorentz found it. He showed that the terms of the first order should cancel each other, but not the terms of the second order. Then more exact experiments were made, which were also negative; neither could this be the result of chance. An explanation was necessary and was forthcoming; they always are; hypotheses are what we lack the least. But this is not enough. Who is there who does not think that this leaves to chance that this singular concurrence should cause a certain circumstance to destroy the terms of the first order, and that a totally different but very opportune circumstance should cause those of the second order to vanish? No; the same explanation must be found for the two cases, and everything tends to

[31] Poincare 1958, 94. This formulation is from an article entitled *L'État Actuel et l'Avenir de la Physique Mathématique* that he originally published in 1904. Charles Scribner shows that this view can be traced in Poincare as early as 1895: "Experiment has revealed a multitude of facts which can be summed up in the following statement: it is impossible to detect the absolute motion of matter, or rather the relative motion of ponderable matter with respect to the ether; all that one can exhibit is the motion of ponderable matter with respect to ponderable matter". From *L'Éclairage Électrique*, quoted in Scribner 1964, 673. It was only in 1904, in the passage that I quoted, when he first dubbed his principle as 'the principle of relativity'. Notice that in the passage of 1895 it is quite clear that the status of this principle is not *a priori* or ultimate, in the sense that it does not require further explanation; the principle is grounded on experience Poincare never quit to this view.

show that this explanation would serve equally well for the terms of the higher order and that the mutual destruction of these terms will be rigorous and absolute.[32]

A second important issue in which Poincare was important in the development of Lorentz' theory consists in his interpretation of 'local time'. The analysis that he provides of this concept is such that, unlike Lorentz's, it has a definite physical meaning: local time is a *measured* quantity in a frame in motion with respect to the ether, whereas the real time can only be measured in the ether-rest frame. He provides his analysis by means of the case in that time measurements and the determination of simultaneity are established through the interchange of light signals between two observers. At time t_A=0 in his watch observer A sends a light signal to observer B, and when the latter receives the signal at time t_B $B's$ clock must be set to AB/c in order to get synchronized with $A's$ –where AB is the distance between the observers and c is the speed of the light signal. If at $t_b = AB/c$ B sends back a light signal to A, then A will receive it a time $2AB/c$. At this point is when Poincare introduces his relevant observation:

> In fact they mark the same hour at the same physical instant, but on the one condition, that the two stations are fixed. Otherwise the duration of the transmission will not be the same in the two senses, since the station A, for example, moves forward to meet the optical perturbation emanating from B, whereas the station B flees before the perturbation emanating from A. The watches adjusted in that way will not mark, therefore, the true time; they will mark what may be called the *local time*, so that one of them will gain on the other. It matters little, since we have no means to perceive it. All the phenomena which happen at A, for example, will be late, but all will be equally so, and the observer will not perceive it, since his watch is slow; so, as the principle of relativity would have it, he will have no means of knowing whether he is at rest or in absolute motion.[33]

Poincare does not explicitly mention the ether as the referential 'object' with respect to which the light does travel with the same velocity in all directions[34]. However, if one reads his writings it is clear that he is consistent in using the expression 'absolute motion' as meaning 'motion with respect to the ether'. Therefore, the *true* time is *measured* only in the ether-rest frame, whereas any motion with respect to it determines that the time to be *measured* will be the *local* one. Thus, Lorentz's auxiliary quantity in his 1895 coordinate transformations offers an explanation, up to first order, of why the observers do not notice their motion across the ether:

---

[32] Poincare 1952a, 172. From *Sur les Rapports de la Physique Expérimentale et de la Physique Mathématique*, originally published in 1900.
[33] Poincare 1958, 99. From *L'État Actuel*... (1904).
[34] One must be careful about this point. In an article of 1898, Poincare states that the assumption of the light speed being the same in all directions –in the ether-rest frame– is a *convention* that cannot be verified by any experiment. See *The Measure of Time*, the English translation of that article, in Poincare 1958, 27-36.
With respect to the ether and the speed of light, the following passage is maybe more clear:
"If they are carried along in common motion... [Suppose] now that A, for example were overtaking the light that went to B, while B receded from the light that went to A. if the observers are thus carried along in a common translation and they do not suspect it, their regulation [of their clocks] will be defective; their clocks will not indicate the same time; each of them will indicate the *local time* proper to the place where they find themselves.
The two observers will have no means of perceiving if the stationary ether always transmits the advancing light signals with the same velocity. ... The phenomena that each of them would observe would be either advanced or retarded; they would not occur at the same moment as if the translation did not exist, but as if when one were to observe a badly regulated clock, one could not perceive [the motion]... . The appearances would not be altered". Quoted in Goldberg 1967, 940.

> The proof goes as follows. When B receives the signal from A, he sets his watch to zero (for example), and immediately sends back a signal to A. when A receives the latter signal, he notes the time τ that has elapsed since he sent his own signal, and sets his watch to the time $\tau/2$. By doing so he commits an error $\tau/2 - t_-$, where t_- is the time that light really takes to travel from B to A. This time and that of the reciprocal travel are given by $t_- = AB/(c+u)$ and $t_+ = AB/(c-u)$, since the velocity of light is c with respect to the ether. The time τ is the sum of these two traveling times. Therefore, to first order in u/c, the error committed in setting the watch A is $\tau/2 - t = (t_+ - t_-)/2 = uAB/c^2$. At a given instant of the true time, the times indicated by the two clocks differ by uAB/c^2, in conformity with Lorentz's expression of the local time.[35]

That is, according to Poincare, Lorentz's local time does not only explain optical experiments designed to measure ether-wind, but also why this *time measuring* effect occurs. The latter phenomenon was not envisioned in the scope of Lorentz's mathematical interpretation of local time.

One last reference to Poincare's reception of Lorentz's theory that I will address has to do with yet another criticism he put forward. Since the theory conceives a purely electromagnetic ether that affects ponderable matter in electrodynamic terms, but which in turn is not affected by the latter –the most apparent example being that its motion has no consequences at all on the ether–; it implies a violation of Newton's third law, the principle of action and reaction. Lorentz had seen that point, but unlike Poincare, and based on the empirical success and the very wide scope of his theory, he simply concluded that the principle had to be considered in a more modest way:

> It is true that this conception [the immobile ether] would violate the principle of the equality of action and reaction –because we do not have grounds for saying that the ether *exerts* forces on ponderable matter– but nothing, as far as I can see, forces us to elevate that principle to the rank of a fundamental law of nature.[36]

Poincare rejected this attitude towards the principle because its violation gets associated with violations of other important and central mechanical laws, namely, the conservation of momentum and the center-of-mass theorem[37]. He illustrated his point with an example:

> Imagine, for example, a Hertzian oscillator, like those used in wireless telegraphy; it sends out energy in every direction; but we can provide it with a parabolic mirror, as Hertz did with his smallest oscillators, so as to send all the energy produced in a single direction. What happens then according to the theory? The apparatus recoils as if it were a cannon and the projected energy a ball; and that is contrary to the principle of Newton since our projectile here has no mass it is not matter, it is energy.[38]

The 'antenna', by emitting energy in one direction, recoils; so that its velocity changes. The recoil has no reaction associated, so the change of velocity implies a change in

[35] Darrigol 2005, 10. u is of course the velocity of A and B with respect to the ether.
[36] From Lorentz's *Versuch*, quoted in Janssen 2003, 34.
[37] This theorem affirms that in an isolated system, a system in which no external forces act, no process can alter the state of motion of its center of mass. It is quite obvious that it is closely connected to Newton's third and first laws.
[38] Poincare 1958, 101. Originally from *L'État Actuel…* (1904). His original treatment of this issue appeared in *La Théorie de Lorentz et le Principe de Réaction*, included in a collective volume celebrating the 25th anniversary of Lorentz doctorate, and published in 1900. In this work Poincare also referred to his interpretation of local time and to his criticism of the structure of Lorentz's theory.

momentum that is not compensated in any other part of the system. Moreover, the recoil also implies that the center of mass of the whole system moves, thus violating the corresponding theorem.

To rescue the violated principles of classical mechanics, Poincare proposed that the action and reaction law should not be interpreted as being valid only for matter with mass, but that an electromagnetic momentum G, defined as $G = \int E \times B \, dV$ –where V is the volume and E and B are the electric and magnetic fields– had to be considered as well. This electromagnetic momentum plays the role of the compensation for the recoil of the antenna, so that the conservation of momentum law survives, but now expressed as $\sum mv + \int E \times B \, dV = constant$.

The electromagnetic momentum introduced by Poincare was conceived as carried by a *fictitious fluid* in the ether, whose mass density M is given by $M = J/c^2$, where J is the electromagnetic energy of the fluid. The recoil of the example is no longer a violation of the momentum conservation law, but it still seems to imply a violation of the center-of-mass theorem. Poincare's way out to this remaining problem consisted in his specific interpretation of the energy carried by the fictitious fluid. According to it, the possibility of the transformation of electromagnetic energy in other forms of energy within the system implies that it is not necessarily a *closed* system; therefore, there is no reason for the theorem to hold in this case:

> In Newtonian mechanics, the constancy of the quantity of motion means that the center of gravity of the system is uniform and rectilinear. But this condition about the center of gravity is not justified here because the fictitious fluid is not indestructible. If there is neither creation nor destruction of electromagnetic energy, the center of gravity of the entire system would behave as in Newtonian mechanics; but suppose that, at certain points, electromagnetic energy were to be converted into other forms of energy. Then it would be necessary to consider not only the motion of the ponderable material and the motion of the electromagnetic energy as represented by the fluid, but also the motion of the nonelectromagnetic energy arising from the conversion. This nonelectromagnetic energy would not necessarily be moving with the system in question. On the other hand, in the case of the creation of electromagnetic energy from other forms, the fluid which would be created at any point could, at first, appear without any velocity. It would then have to receive its velocity from fluid already in existence and therefore the velocity of the entire ensemble would diminish unless some outside agent intervened to hold the velocity constant.[39]

Two remarks are interesting with respect to the electromagnetic momentum and the fictitious fluid introduced by Poincare. First, they reinforce the role of the ether in the theory, for both are clearly and intrinsically related to it; one might even say that they are determinations of the ether. Even though Poincare finally accepted that Lorentz's theory implied violations or deep reinterpretations of central principles of classical mechanics, the introduction of concepts and physical features as grounded on determinations of the ether was a consistent practice both in Lorentz and Poincare. This is an indication of how essential the ether was in the ontology of the theory; and this feature is, I think, a definitive argument against a view in which Lorentz's theory (with all the amendments introduced by Poincare) and SR are not only predictively equivalent, but also conceptually and ontologically

[39] Goldberg 1967, 941.

equivalent –that is, two different formulations of the same theory. By the same token, it is also a strong argument against the view that Poincare should be considered as the author of the SR theory alongside Einstein.

Second, it is remarkable that Poincare's expression for the energy of the fictitious fluid, in a way, contains Einstein's famous equation $E = mc^2$. However, he did not interpret the former as stating that energy, in and by itself, has inertia; or that the inertial mass of a material body can vary according to its energy content[40]. His analysis of the example is quite clear about it. He understood his fictitious fluid as an *explanation* for the recoil, whereas had he understood its energy expression in a general and 'relativistic' way he would have seen that both the momentum conservation law and the center-of-mass theorem are respected. Yet another curious feature of this story is that Einstein's second derivation of his famous equation, in 1906, was based on the validity of the center-of-mass theorem, and he explicitly stated that his treatment of the issue was similar to the one undertaken by Poincare that I just sketched. Regarding this, Janssen shows that in order to make the theories fully equivalent, Lorentz's must *borrow* $E = mc^2$ from SR. This is certainly right, especially from a historical point of view. However, from a conceptual standpoint, it must be acknowledged that the famous equation was there, in Lorentz's theory, 'waiting to be discovered'. It is true that it was not, and that Lorentz only saw it after Einstein's work; but all the conceptual machinery needed to formulate it was already present in Lorentz's theory. I will return to this issue.

Summarizing, ca. 1900, Poincare got crucially involved in the development of Lorentz's theory. He criticized its structure, and as I will now show, Lorentz's reaction greatly improved its foundations. On the other hand, he was able to see that there was a physical meaning contained in 'local time', and he also noticed that the theory had implications on momentum conservation and the center-of-mass theorem that, if properly considered, would lead to the energy-mass relation equation. Both these features are crucial in a case for the predictive equivalence of the theories.

c) Second stage: 1899-1904

Lorentz's definitive formulation of his theory was presented in his *Simplified Theory of Electrical and Optical Phenomena in Moving Bodies* (1899), and in *Electromagnetic Phenomena in Systems Moving with any Velocity Less than that of Light* (1904). In the first work he gave a unified and exact formulation of the theorem of corresponding states that gets closely connected to the hypothesis of length contraction. In the second, he added his very important and famous model of the electron. I will now offer an exposition of both issues in turn.

[40] Even though Lorentz found out that the inertial mass of a body is proportional to its velocity, and therefore to its *kinetic* energy, neither he nor Poincare initially interpreted this result as an instance of a general relation between *energy*, without any last name, and mass.

The definitive formulation of the corresponding states theorem was given by a modification of the coordinate transformations he had introduced in 1895. The new transformations are:

$$x' = l\gamma x, \qquad y' = ly, \qquad z' = lz, \qquad t' = l[t/\gamma - \gamma(v/c^2)x],$$

with $= \frac{1}{\sqrt{1-v^2/c^2}}$. The term l can differ from 1 only by an amount in the order of v^2/c^2 and Lorentz left it undetermined in 1899, but in 1904 set it to 1 –for reasons that I will refer to below. In any case, the presentation of the final theorem gets harmlessly simpler if it is set to 1 right away.

It is important to remember that Lorentz is using his three-step method, so the coordinates in S_0 at rest in the ether convert to the coordinates of S, in motion with respect to the ether, by means of the Galilean transformations. Finally, the S coordinates convert in the coordinates in the auxiliary frame S' through the Lorentz transformations. Therefore, combining the Galilean transformations from S_0 to S with the transformations from S to S', we have that the transformations from S_0 to S', with l=1, are:

$$x' = \gamma(x_0 - vt_0), \qquad y' = y_0, \qquad z' = z_0, \qquad t' = \gamma[t_0 - (v/c^2)x_0]$$ [41]

By means of these new transformations, Maxwell's field equations become invariant without neglecting terms of any order of v/c. That is, the auxiliary fields in the system S', considered as functions of the auxiliary coordinates in S', satisfy the same equations as the real fields considered as functions of the real coordinates in S_0. Thus Lorentz's new formulation of the theorem of corresponding states consists in that

> If there is a solution of the source free Maxwell equations in which the real fields **E** and **B** are certain functions of $\mathbf{x}_0$ and t_0, the coordinates of S_0 and the real Newtonian time, then there is another solution of the source free Maxwell equations in which the fictitious fields **E**′ and **B**′ are those exact functions of **x**′ and t', the coordinates of S and the local time in S.[42]

As I mentioned above, one must be careful and not to conclude right away that the factor γ expresses the length contraction factor. Actually, in 1892 he used it as a mere mathematical tool, as I showed above. Janssen is very clear that the same precaution must be taken here. This final formulation of the theorem of corresponding states does not logically entail the contraction; this is a further physical assumption. However, it is quite clear that one of the main motivations underlying Lorentz's work of 1899 and on was to merge the theorem with the contraction, in order to provide a unified and general explanation for why none of the optic experiments set out to find ether-wind effects had obtained positive results –that is, to fulfill Poincare's demand. The accomplishment of this goal was given by a specific interpretation of the theorem under what Janssen dubs the *generalized contraction hypothesis*:

---

[41] The derivation of the time coordinate transformation goes as follows:
$t' = t/\gamma - \gamma(v/c^2)x = t_0/\gamma - \gamma(v/c^2)(x_0 - vt_0) = \gamma(t_0[1/\gamma^2 + v^2c^2] - (v/c^2)x_0)$; and since $(1/\gamma^2 + v^2c^2) = 1$, then $t' = \gamma[t_0 - (v/c^2)x_0]$.

[42] Janssen 1995, section 3.3.3.

> If a material system, i.e., a configuration of particles, with a charge distribution that generates a particular electromagnetic field configuration in S_0, a frame at rest in the ether, is given the velocity **v** of a Galilean frame S in uniform motion through the ether, *it will rearrange itself* so as to produce the configuration of particles with a charge distribution that generates the electromagnetic field configuration in S that is the corresponding state of the original electromagnetic configuration in S_0.[43]

To clearly see the difference between the theorem of corresponding states, as a mathematical tool, and the generalized contraction hypothesis, as a physical assumption, it is enough to pay attention to the fact that the former establishes a relation between –on the one hand– two *real* frames S_0 and S, and –on the other hand– an *auxiliary* frame S' through the application of the coordinate transformations; whereas the physical length contraction assumption establishes a relation between the *real* frames S_0 and S. That is, the field configuration in S_0 *physically* transforms in its corresponding state configuration in the frame S. Notice also that the generalized contraction hypothesis can be understood as a generalization of the plausibility argument that Lorentz offered for the contraction hypothesis in 1892-5.

Janssen, in order to clarify this point, shows that the theorem can be applied to obtain an explanation of the Michelson-Morley experiment *with* and *without* the generalized contraction hypothesis. Without it, the corresponding state of the interferometer in S is a 'stretched out' interferometer in S_0, so that the latter

> will change its shape as the moving interferometer is rotated. This means that whether it is dark or light at P' (and thereby at P) will depend on the orientation of the moving interferometer. Without the contraction hypothesis, Lorentz's theory therefore predicts a positive result in the Michelson-Morley experiment.

On the other hand, with the generalized contraction hypothesis,

> The corresponding state of the moving *contracted* interferometer is simply the uncontracted interferometer at rest in the ether. So, the shape of the corresponding state will not depend on the orientation of the moving interferometer with respect to its velocity. As a consequence, we now expect negative results.[44]

Let us remind the general explanation of the optic experiments given by the 1895 version of the theorem. It affirmed that the *structure* of a pattern of light and darkness in S_0 is the same as the one in S. This time it must be added that the pattern of light and darkness in S differs from the one in S_0 in that the former is contracted by a factor γ^{-1}. This *real* difference yields that the experiments will not have positive results even for the second order of v/c.[45]

[43] Janssen 1995, section 3.3.3, my italics.

[44] *Ibid.* The explanation *without* the contraction hypothesis would correspond to Lorentz's conception of γ in 1892 that I explained above.

[45] *Ibid.* Yet another concrete case in which is clear that the theorem and the physical assumption are logically independent, though they get merged by Lorentz in 1899, is mentioned by Janssen and Stachel: "What prompted Lorentz's new more general theory was in fact a variant of the Michelson-Morley experiment proposed in 1898 by Alfred Liénard. Liénard wanted to repeat the Michelson-Morley experiment with some transparent medium in the arms of the interferometer. In that case, the Lorentz-Fitzgerald contraction would no longer ensure that the travel time in an arm of an interferometer is independent of whether the arm is parallel or perpendicular to the ether drift. Liénard did not actually perform the experiment, but both he and Lorentz strongly suspected that the outcome, as the outcome of so many experiments before, would be negative. As Lorentz emphasized in his 1899 paper, his new theory could account for such a negative result". Janssen & Stachel 2004, 28.

Summarizing, we have that the theorem and the physical hypothesis are logically independent. However, Lorentz's view of the subject consisted in that the theorem had to be interpreted and understood under the physical assumption. He explicitly stated this view in his 1899 work:

> We shall not only suppose that the system S_0 *may* be changed in this way into an imaginary system S[46], but that, as soon as the translation is given to it, the transformation *really* takes place, of itself, i.e., by the action of the forces acting between the particles of the system, and the aether. Thus, after all, S will be the *same* material system as S [this clearly should be S_0].
> The transformation of which I have spoken, is precisely such a one as is required in my explication of Michelson's experiment.[47]

This important subtlety has been many times unnoticed and has led to some confusion, for it can suggest a wrong interpretation of Lorentz's theory –in the sense that it is understood in a *modern relativistic* way that is not justified. For example, Janssen quotes a passage of Pais' famous scientific biography of Einstein in which it is affirmed that "the reduction of the Lorentz-Fitzgerald contraction to a consequence of Lorentz transformation is a product of the nineteenth century"[48]. I think that Janssen is quite clear and right that there is no such reduction, and that Lorentz was conscious of it.

The revised and improved version of the theorem of corresponding states, understood under the generalized contraction hypothesis, entails a very surprising result from the point of view of classical physics. In the ether-rest frame, Newton's 2nd law has the form $F = ma$, whereas in the frame S' in motion along the x-axis of S_0, it is $F' = ma'$; for according to Newtonian physics the inertial mass of a body is an absolute invariant quantity. This cannot be the case in Lorentz's theory. The line of thought that led Lorentz to this result was based on the case of an oscillating electron in the ether-rest frame which generates an electromagnetic wave, wave whose oscillation satisfies Newton's law. Then he considered that same electron in a frame moving through the ether with velocity v. Its motion in the corresponding state, determined by the auxiliary quantities in the coordinate transformations, is the same as its motion in the ether-rest frame. This implies that, in terms of the *real quantities*, Newton's law holds only if the mass of the electron depends on its velocity.

Remember that Lorentz assumed that all forces transform as electromagnetic ones do, so that in a frame in motion with respect to the ether the force F' is given by

$$F'_{x'} = F_x \,, \qquad F'_{y'} = \gamma F_y = \frac{F_y}{\sqrt{1-v^2/c^2}}, \qquad F'_{z'} = \gamma F_z = \frac{F_z}{\sqrt{1-v^2/c^2}} \,.$$

[46] Janssen claims that this system S referred by Lorentz does correspond to the real system S, and that use of the adjective 'imaginary' is simply based on the fact that the state of the system S is described by the fictitious *imaginary* primed coordinates. It might be so, but I think that the meaning of the passage gets even clearer if one substitutes S' for S. I completely agree in that in the second case he is clearly referring to S_0.

[47] From Lorentz's *Simplified Theory...* (1899), quoted in Janssen 1995, section 3.3.4.

[48] *Ibid.*

On the other hand, the relation between the acceleration $\boldsymbol{a}$ in the ether-rest frame and the acceleration $\boldsymbol{a}'$ in the auxiliary frame –if the velocity and amplitude of the electron's oscillation are small enough as to be neglected– is given by:

$$a'_{x'} = \gamma^3 a_x, \qquad a'_{y'} = \gamma^2 a_y, \qquad a'_{z'} = \gamma^2 a_z \ .$$

With these expressions, Newton's 2nd law can be formulated, in terms of the *real quantities*, as

$$F_x = m\gamma^3 a_x \ , \qquad F_y = m\gamma a_y \ , \qquad F_z = m\gamma a_z$$ [49].

From these expressions it follows that if Newton's law of motion is to hold in the moving frame then it cannot be the case that inertial mass is an absolute quantity. In this theory mass becomes a velocity-dependent property. To state the precise formula for this dependence, it must be noticed that the first of the expressions just above holds for the acceleration in the direction of motion –assuming that the motion occurs along the x-axis, of course–, whereas the other two hold for the accelerations perpendicular to the direction of motion. Considering this we have that in Lorentz's theory inertial mass is given by

$$m_L = \gamma^3 m_0 \ , \qquad m_T = \gamma m_0 \ ;$$

where m_L is the *longitudinal mass*, m_T is the *transverse mass*, and m_0 is the *ether-rest mass*.

This is a very surprising result from the point of view of Newtonian Mechanics, so Lorentz somehow felt a necessity to provide an explanation of it. He did so in his 1904 work, from the point of view of his model of the electron, model that was closely connected to what is commonly known as the *electromagnetic view of nature*. I now turn to a brief exposition of both subjects[50].

All the difficulties that classical mechanics had faced at the end of the 19th century –the ether problem and Lorentz's solution for it, for instance– led to the formulation of a new program for physical science. In 1900 Wilhelm Wien published a sort of *manifesto* whose main tenets were the assumption of a basic ontology determined by negative and positive charged particles that constitute all ponderable matter, and the view that the inertial mass of those particles is of electromagnetic origin. That is, the inertial mass of any object is the outcome of the interaction between the charged particles, their electromagnetic fields, and the ether. The explanation that Lorentz provided for his derivation of the velocity dependence of mass was grounded on these tenets. He actually claimed that the intrinsic relation between mass and velocity-across-the-ether was not that surprising after all, for some years before scientists like Thomson, Heaviside and Searle had already shown that the

---

[49] For $F'_{x'} = F_x$; and $F'_{z'} = \gamma F_z = m\gamma^2 a_z$, so that $F_z = m\gamma a_z$. The last result also holds, *mutatis mutandis*, for $F'_{y'}$.

[50] This brief outline of the electromagnetic view of nature is based on the classic paper on the issue: McCormmach 1970b; Kragh 1999, chapter 8; and Harmann 1982, chapters 4 and 6, also deal with this subject. On Lorentz's electron model, I closely follow Janssen & Mecklenburg 2007 and Janssen 1995, section 3.4. Miller 1998 [1981], 62-80 is also useful.

effective mass of a charged particle was a function of its velocity with respect to the ether: the effective mass of a particle was given by the sum of its *Newtonian* mass and its *electromagnetic* mass, where the latter was velocity-dependent. The electromagnetic view simply took one further step and asserted that *all* the inertial mass was of electromagnetic origin and velocity dependent. This view provided Lorentz with the basis for an explanation of his surprising result. If all inertial mass is a function of the interaction among charged particles and the ether, then its velocity dependence becomes quite a natural and expected feature.

Lorentz's specific position on this issue was formulated in his 1904 *Electromagnetic Phenomena in Systems Moving with any Velocity less than that of Light*, and it was grounded on five main assumptions: *i*) that all forces transform in the same way as electromagnetic forces do; *ii*) that a spherical electron, when moving across the ether, undergoes a physical deformation expressed by the coordinate transformations and becomes ellipsoid, i.e., the Lorentz contraction holds also for electrons themselves; *iii*) that the origin of all of the inertial mass of an electron is of electromagnetic; *iv*) that the deformation occurs only in the longitudinal direction with respect to the motion of the electron; and *v*) that the masses of all bodies, charged or not, vary with motion just as the mass of electrons does. Under these assumptions Lorentz was able to derive, for the electron itself, expressions for the velocity dependence of its mass that were equivalent to the ones he obtained in 1899 without considering a specific model of the electron.

Assumption *iv*) entails that the factor l I mentioned above gets definitively set to 1. Lorentz's reason for this choice was that 1 is the only value for l in which the velocity dependence of mass is consistent with the corresponding states theorem and the generalized contraction hypothesis, for any other value would make it possible to measure some ether-wind effect.

In order to understand more deeply the meaning of Lorentz's electron model, it is useful to make a brief comparison with its rivals. By 1905 there were two other alternatives available. Both Abraham's and Bucherer-Langevin's models –just as Lorentz's– assumed a full electromagnetic origin for the inertial mass of charged particles, and that charged particles were the ultimate constituents of ponderable matter. The differences were that Abraham's electron was rigid and not affected by a contraction when set in motion through the ether; and that the Bucherer-Langevin electron, along with a longitudinal contraction, also suffered a transverse expansion due to motion, the value of the contraction being $\gamma^{2/3}$ and the value of the dilation being $\gamma^{1/3}$ –so that the factor l is equal to $\gamma^{-1/3}$. In other words, Abraham's model denied Lorentz's assumptions *i*) and *ii*), and assumption *iv*) becomes irrelevant; whereas Bucherer-Langevin's denies assumption *iv*); but all three models share assumptions *iii*) and *v*), which constitute the basic tenets of the electromagnetic view of nature.

The line of scientific development from the problem of the ether to the formulation of Lorentz's theory settled the basic groundings for the electromagnetic world view. By the early 1900s, the three mentioned models committed to this view were contending, so the choice to be made was understood as a matter of empirical tests. The interpretation of the data obtained from Kaufmann's experiments performed during 1901-3, in which β-radiation was used in order to measure the precise value for the velocity dependence of the inertia of particles, was the empirical battlefield on which the models –and also SR– competed. It turned out that the technology available was not enough in order to set the experiment in a way that the resulting data could be considered as totally reliable. However, the relevant point is that ca. 1905 the electromagnetic view of nature was hold as a very promising and unifying program for the development of physics, and as a program in that classical mechanics became reduced to electrodynamics.

In spite of the promising path that Lorentz's theory was opening –within the context of the electromagnetic view–, some problems quickly came up. The most relevant one was that of an ambiguity, or even an inconsistency, in the formulation of the longitudinal mass. This problem was first posed by Abraham in 1905: in Lorentz's theory the value for the longitudinal mass in terms of the electromagnetic momentum was not the same as the one expressed in terms of energy.

The Newtonian expression for force can be formulated by means of the rate of change of momentum through time, i.e., $F = \frac{d\boldsymbol{p}}{dt}$. If in this formula, and in the spirit of the electromagnetic view, the ordinary momentum is replaced by the electromagnetic momentum of the electromagnetic field $\boldsymbol{P}_{em}$[51], then it follows that $F = \frac{d\boldsymbol{P}_{em}}{dt} = m_L a + m_T a$ – where m_L and m_T are the longitudinal and transverse mass, respectively[52]. From this formula it follows that $m_L = \frac{d\boldsymbol{P}_{em}}{dv}$, which is the expression for the longitudinal mass *in terms of the electromagnetic momentum*. The formula for the electromagnetic momentum for the field associated to an electron is $\boldsymbol{P}_{em} = \frac{4}{3}\gamma l \left(\frac{U'_{em}}{c^2}\right) v$[53], where U'_{em} is the electromagnetic energy of the electron in a corresponding state of the ether-rest frame, i.e., in a system moving with respect to the ether. By plugging the right side of this formula in the formula for the longitudinal mass, we have that $m_L = \frac{d(\gamma l v)}{dv}\frac{4}{3}\frac{U'_{em}}{c^2}$; and since $\frac{d(\gamma v)}{dv} = \gamma^3$[54] (and with l=1), then $m_L = \gamma^3 \frac{4}{3}\frac{U'_{em}}{c^2}$. If we recall that in 1899 Lorentz obtained $m_L = \gamma^3 m_0$ from the corresponding states theorem and the generalized contraction hypothesis, both expressions become equivalent by defining the rest mass as $\frac{4}{3}\frac{U'_{em}}{c^2}$[55].

---

[51] The definition of the electromagnetic momentum is $\boldsymbol{P}_{em} \equiv \int \varepsilon_0 \boldsymbol{E} \times \boldsymbol{B} d^3x$.

[52] "Proponents of the electromagnetic view of nature took the eq. 18 [$F = \frac{dP_{em}}{dt}$] to be the fundamental equation of motion and derived Newton's law from it by indentifying the ordinary Newtonian mass with the electromagnetic mass m_0 of the relevant system at rest in the ether", Janssen & Mecklenburg 2007, 10.

[53] For the derivation of this formula, see Janssen & Mecklenburg 2007, 19.

[54] See Janssen & Mecklenburg 2007, page 11 and footnote 30.

[55] In this respect, Janssen and Mecklenburg comment: "Lorentz (1904) had thus found a concrete model for the electron with a mass exhibiting exactly the velocity dependence that he had found in 1899. This could hardly be a coincidence. Lorentz concluded that the electron was indeed nothing but a small spherical surface charge distribution, subject to a microscopic version of the Lorentz-Fitzgerald contraction when set in motion, and that its mass was purely electromagnetic, i.e., the result of interaction with its self-field [see Lorentz's five assumptions I mentioned above][...]. So it is indeed no coincidence that Lorentz found these same relations twice, first, in 1899, as a necessary condition for rendering ether-drift unobservable and then, in 1904, as the mass-velocity relations for a concrete Lorentz-invariant model of the electron. But the explanation is not, as Lorentz thought, that his model provides an accurate representation of the real electron; it is simply that the mass of *any* Lorentz-invariant model of *any* particle –whatever its nature and whatever its shape– will exhibit the same velocity dependence", 21.

However, from a modern-relativistic point of view, to define the rest mass as $\frac{4}{3}\frac{U'_{em}}{c^2}$ is rather awkward, for in that case the mass-energy relation becomes $\frac{4}{3}E = mc^2$, instead of $E = mc^2$. This issue is commonly known as the "4/3 puzzle" and it is involved in the problem of the ambiguity of the expression for the longitudinal mass I am considering[56].

Turning now to the expression for longitudinal mass in terms of energy, we have that as an electron moves in the x-direction, and assuming the absence of an external field, the work expended can be expressed as a change in the internal energy of the electron, $dU = -dW$. Since the work is done by the force coming from the self-field for the electron –given the no-external-field assumption–, and identifying the internal energy with the electromagnetic energy, then $dU_{em} = -dW = -\boldsymbol{F}_{self} \cdot d\boldsymbol{x}$. $\boldsymbol{F}_{self}$ is equal to minus the rate of change of electromagnetic momentum over time[57], so that

$$dU_{em} = \frac{d\boldsymbol{P}_{em}}{dt} \cdot d\boldsymbol{x} = m_L a_L \cdot d\boldsymbol{x} = m_L \frac{dv}{dt} dx = m_L v dv;$$

Hence, $dU_{em} = m_L v dv$; and therefore, $m_L = \frac{1}{v}\frac{dU_{em}}{dv}$, which is the expression for the longitudinal mass *in terms of the electron's energy*.

The inconsistency comes up because it can be shown that $U_{em} = l\left(\frac{4\gamma}{3} - \frac{1}{3\gamma}\right) U'_{em}$[58] –as usual, the primed quantity corresponds to the system in motion through the ether, whereas the unprimed corresponds to the ether-rest frame. Plugging the right side of this equation in the expression for the longitudinal mass, one obtains $m_L = \frac{1}{v}\frac{d}{dv}\left(\frac{4\gamma l}{3} - \frac{l}{3\gamma}\right) U'_{em}$. By setting l to unity, and since $\frac{d\gamma}{dv} = \gamma^3 \frac{v}{c^2}$[59]; then

$$m_L = \frac{1}{v}\frac{4}{3}\frac{d\gamma}{dv} U'_{em} - \frac{1}{3v}\frac{d}{dv}\left(\frac{1}{\gamma}\right) U'_{em} = \gamma^3 \frac{4}{3}\frac{U'_{em}}{c^2} - \frac{1}{3v}\frac{d}{dv}\left(\frac{1}{\gamma}\right) U'_{em}.$$

Since the first term in the last expression, namely $\gamma^3 \frac{4}{3}\frac{U'_{em}}{c^2}$, is equal to m_L as expressed *in terms of the electromagnetic momentum*, the presence of the second term, $-\frac{1}{3v}\frac{d}{dv}\left(\frac{1}{\gamma}\right) U'_{em}$, indicates that the value for the longitudinal mass obtained *in terms of the energy* is not the same as the value obtained from electromagnetic momentum. In other words, the problem that Abraham saw in Lorentz's theory was that in it $m_L = \frac{d\boldsymbol{P}_{em}}{dv}$, $m_L = \frac{1}{v}\frac{dU_{em}}{dv}$; but $\frac{d\boldsymbol{P}_{em}}{dv} \neq \frac{1}{v}\frac{dU_{em}}{dv}$.

Abraham pointed out the problem in his 1905 *Theory of Electricity: electromagnetic theory of radiation* –published in German. The interpretation he made of this issue was that

---

56 For a deeper treatment of this issue see Miller 1998 [1981], 72-80; Miller 1986, 266-301 and 309-18; Janssen 1995, section 2.2; Janssen & Mecklenburg 2007, 19-50; and Janssen 2003. I will say some more words about it in the next section.

57 The force of the self-field is equal to minus the time derivative of electromagnetic momentum, that is, $\boldsymbol{F}_{\text{self}} = \int \rho(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) d^3x = -\frac{d\boldsymbol{P}_{em}}{dt}$; where ρ is the density of electron's charge distribution. See Janssen & Mecklenburg 2007, 8.

58 For the derivation of this equation, see Janssen & Mecklenburg, 18.

59 See Janssen & Mecklenburg 2007, page 11 and footnote 30.

the disagreement was grounded in the fact that "the entire energy of Lorentz's electron could not be accounted for by electromagnetic forces alone"[60]. The solution consisted then in including an extra force to account for the total energy. The problem was that Abraham noticed that the introduction of such a force would threaten the purity of the electromagnetic view that Lorentz's theory was endorsing, for the compensating force could not be an electromagnetic one. Moreover, the compensating non-electromagnetic force was necessary to provide stability to Lorentz's electron; otherwise its own Coulomb repulsive forces would make it to explode[61]. The realization of this solution was introduced by Henri Poincare in 1906, along with other interpretations and remarks about Lorentz's theory that are essential to really obtain its predictive equivalence with respect to SR.

d) Poincare, once again

In his 1906 *On the Dynamics of the Electron* –published in French[62]– Poincare introduced a non-electromagnetic quantity in order to solve the problem of stability of the electron and the inconsistency of the expressions for its longitudinal mass. The quantity is commonly known as *Poincare-pressure*. The formula for this pressure is $P_{Poincare} = -\frac{1}{3}\frac{U\prime_{em}}{V_0}$, where V_0 is the volume of the electron at rest[63]. The quantity is negative since it counterbalances Coulomb *repulsive* forces; and it is exerted only on the surface of the electron, so that there is a sudden drop in pressure at its edge. Besides, Poincare-pressure also provides a foundation for Lorentz's assumption that the electron itself gets Lorentz-contracted as it moves across the ether:

> These forces serve two purposes. First they prevent the electron's surface charge distribution from flying apart under the influence of the Coulomb repulsion between its parts. Second, as the region where $P_{Poincare}(\mathbf{x})$ is non-vanishing always coincides with the ellipsoid-shaped region occupied by the moving electron[64], these forces make the electron contract by a factor γ in the direction of motion.[65]

Abraham had pointed out that the total energy of the electron could not be accounted for only by means of its electromagnetic energy. The Poincare-pressure contributes the missing energy in an amount of $\frac{1}{3}\frac{U\prime_{em}}{\gamma}$, which is minus the product of the Poincare-pressure and the volume $V = V_0/\gamma$ of the moving electron. Accordingly, the electron total energy is

[60] Miller 1998 [1981], 72.

[61] Strictly speaking, the problem of the electron's stability held for the three models. However, Abraham's assumption of the rigid electron was 'axiomatic', so that the counterbalance of the Coulomb forces was not a force, but some sort of 'rigid constraints'. Therefore, he could tackle the problem of stability from within the electromagnetic world picture, that is, without introducing non-electromagnetic forces or energy. See Janssen 1995, section 3.4.3; and Janssen & Mecklenburg 2007, 22-5.

[62] A shorter version of the same paper appeared in 1905, with the same title.

[63] For an analysis of Poincare's derivation of this expression, see Janssen & Mecklenburg 2007, 26-31. For a detailed examination of Poincare 1906, see Miller 1986, 29-150.

[64] The Poincare pressure can be written, for an electron moving along the x-direction through the ether, as $P_{Poincare}(\boldsymbol{x}) = -\frac{1}{3}\frac{U_{0em}}{V_0}\vartheta\left(R - \sqrt{\gamma^2 x^2 + y^2 + z^2}\right)$, where ϑ is a function defined as $\vartheta(x)=0$ for x<0 and $\vartheta(x)=1$ for x≥0, and where R is the radius of the electron at rest. This holds for a co-moving frame related to an ether-rest frame by means of the Galilean transformation.

[65] Janssen & Mecklenburg 2007, 31.

the sum of the electromagnetic and the non-electromagnetic energies. Recalling that $U_{em} = l\left(\frac{4\gamma}{3} - \frac{1}{3\gamma}\right)U'_{em}$, and with l set to unity, it follows that $U_{tot} = \frac{4}{3}\gamma U'_{em} - \frac{1}{3}\frac{U'_{em}}{\gamma} + \frac{1}{3}\frac{U'_{em}}{\gamma} = \frac{4}{3}\gamma U'_{em}$. Then, using $m_L = \frac{1}{v}\frac{dU_{tot}}{dv}$ instead of $m_L = \frac{1}{v}\frac{dU_{em}}{dv}$ as the expression for the longitudinal mass of the electron in terms of its energy, it turns out that $m_L = \frac{1}{v}\frac{4}{3}\frac{d\gamma}{dv}U'_{em} = \gamma^3\frac{4}{3}\frac{U'_{em}}{c^2} = \gamma^3 m_0$; and the ambiguity is solved.

Moreover, Poincare-stress also solves the problem –with respect to the modern-relativistic perspective– related to the energy-mass relation. For the ether-rest case the total energy reduces to $U_{tot} = \frac{4}{3}U'_{em}$; so that, with l set to 1, the expression for electromagnetic momentum is $\boldsymbol{P}_{em} = \gamma\frac{4}{3}\left(\frac{U'_{em}}{c^2}\right)v = \gamma\frac{U_{tot}}{c^2}v$; and therefore $m_0 = \frac{U_{tot}}{c^2}$. That is, Poincare unconsciously solved the 4/3 puzzle in Lorentz's theory[66].

Just as Abraham stated, Poincare's solution entails that the goal of the electromagnetic world view was not completely fulfilled by Lorentz's theory. However, Poincare did not hesitate about endorsing this result because Lorentz's theory and its model of the electron was the only available theoretical approach that respected his principle of relativity, and so precluded a positive outcome for any ether-wind experiment. As Janssen poses it:

> There [was] no electron model that [was] both compatible with the electromagnetic view of nature and compatible with the general experimental indication that we will never be able to detect ether drift, and therefore with Einstein's relativity principle. The Lorentz electron [was] incompatible with the electromagnetic world view. The Abraham and Bucherer-Langevin are incompatible with the absence of any signs of ether drift.[67]

Both Lorentz and Poincare accepted to sacrifice the purity of the promising new program for the unification of physics. As we have seen, the solution to the problem of the ether, the quest for the invariance of Maxwell equations, and Poincare's synthesis of all the related issues in his principle of relativity were highly valuable achievements. However, Poincare explicitly remained committed to the view that all the inertial mass of the electron is of electromagnetic origin: "If the inertia of matter is exclusively of electromagnetic origin, as is generally admitted since Kaufmann's experiment, and all forces are of electromagnetic origin (apart from this constant pressure I just mentioned), the postulate of relativity may be established with perfect rigor"[68]. If it also considered that the basic ontology of the theory remained the same after Poincare's amendments –charges, fields and the ether–, it is clear that Lorentz's theory stayed committed to the main tenets of the electromagnetic world view.

[66] All what has been said rests on a reconstruction made with the benefit of hindsight, and its aim is to make a case for the predictive equivalence between Lorentz's theory, as amended by Poincare, and SR. From a historical point of view, things are not so modern, for, as Miller points out: "contrary to what sometimes is attributed to this paper, Poincare never computed the counter term necessary to cancel the second term on the right hand of (48) [the equation for work that leads to the problematic equation of energy], nor did he reduce the factor of 4/3 in *G* [P_{em}] to unity. Rather, he proved the necessity of introducing mechanical stresses into Lorentz's theory to account for the inertia of a deformable electron in a manner consonant with the principle of relativity", Miller 1986, 70.
[67] Janssen 1995, section 3.4.1.
[68] From Poincare *On the Dynamics of the Electron*, quoted in Janssen & Mecklenburg 2007, 34. This passage clearly indicates that Poincare did not see that $U_{tot} = mc^2$, or at least that he did not interpret it in the modern way.

Another important amendment that the French scientist introduced in Lorentz's theory had to do with a correction in the expressions for the velocity and charge density transformations. The problem with this issue in Lorentz's theory was rooted, Poincare noticed, in the three-step method that I mentioned above. Let us recall that this method consisted in that Lorentz first connected a system S_0 at rest in the ether with a system S which moves with respect to S_0 with velocity v by means of the Galilean transformations $x = x_0 - vt_0$, $y = y_0$, $z = z_0$, $t = t_0$. Then the system S is connected to the auxiliary frame S' by means of the Lorentz transformations

$$x' = l\gamma x, \qquad y' = ly, \qquad z' = lz, \qquad t' = l[t/\gamma - \gamma(v/c^2)x].$$

The velocity transformation from $u_x = \frac{dx}{dt}$ to $u'_x = \frac{dx'}{dt'}$ that Lorentz obtained was $u'_x = \gamma^2 u_x$. Its derivation goes like this: from the Lorentz transformations from S to S' we have that $dx' = \gamma l dx$, and that $dt' = l\left[\frac{dt}{\gamma} - \gamma(v/c^2)x\right] = \frac{l}{\gamma}dt(1 - \gamma^2 v u_x)$[69]. By plugging the last two equations in $u'_x = \frac{dx'}{dt'}$ one finally obtains $u'_x = \frac{\gamma^2 u_x}{(1-\gamma^2 v u_x)}$[70]. This formula reduces to $u'_x = \gamma^2 u_x$ only if $u_x \ll 1$. That is, it holds if the velocity of the object moving in the frame S is much smaller than c, but if u_x increases the velocity transformation that Lorentz obtained is no longer correct.

Poincare corrected this problem by means of a maneuver which has important consequences for the meaning of the Lorentz transformations. He simply avoided the three-step method and directly connected S_0 with S' through the Lorentz transformation expressed in their modern way:

$$x' = \gamma(x_0 - vt_0), \qquad y' = y_0, \qquad z' = z_0, \qquad t' = \gamma[t_0 - (v/c^2)x_0].$$

By so doing, the expression for the velocity transformation is simply $u'_x = \frac{\gamma d(x-vt)}{\gamma d(t-vx)} = \frac{dx-vdt}{dt-vdx} = \frac{u_x-v}{1-vu_x}$[71]. This is of course the modern relativistic velocity transformation, which implies that c is the maximum possible velocity in any inertial frame[72].

The deeply important consequence of Poincare's maneuver that I mentioned consists in that, by directly relating S_0 and S', the relevant velocity v which operates in the transformation is, in the end, the *relative* velocity between the frames, not the velocity of S with respect to the *ether* –this is the reason why in the last paragraph I used simply x and t in the transformation that Poincare obtained, instead of x_0 and t_0. That is, the amendment

---

69 By choosing the suitable units such that c=1.

70 $u'_x = \frac{\gamma l dx}{\frac{l}{\gamma}dt(1-y^2 v u_x)}$, and since $u_x = \frac{dx}{dt}$; then $u'_x = \frac{\gamma l u_x}{\frac{l(1-\gamma^2 v u_x)}{\gamma}} = \frac{\gamma^2 u_x}{(1-\gamma^2 v u_x)}$.

71 Once again, under the assumption that c=1.

72 For brevity and simplicity, I only referred to the case of the x-velocity component. The expression for the other two components, in the case of both Lorentz's and Poincare's derivations, follows quite analogously. For the transformation for the charge density, which is corrected by Poincare by means of the same maneuver, see Miller 1986, 47-7 and 72-4.

introduced by Poincare is a step towards an interpretation of the Lorentz transformations in terms of relative velocities between the frames involved.

This feature becomes even more apparent by considering yet another improvement that Poincare made on Lorentz's theory. He also showed that the Lorentz transformations form a *group*. A transformation group is a collection of transformations such that *i*) the transformation obtained through the successive application of two transformations of the collection is also a transformation of the collection; *ii*) the transformations are associative, that is, the transformation obtained by the composition of (AB) and C is equal to the transformation obtained from the composition of A and (BC), where A, B, C are transformations of the collection; *iii*) there is an identity transformation D such that DA=A; and *iv*) there exists an inverse transformation.

By proving that the Lorentz transformations comply with requirements *i*) and *iv*), Poincare took yet another step towards an interpretation only in terms of relative velocities between frames. Consider the transformations from a frame S at rest in the ether to a frame S' moving with velocity v with respect to it:

$$x' = \gamma l(x - vt), \qquad y' = ly, \qquad z' = lz, \qquad t' = \gamma l(t - vx),$$

and then consider a second set of Lorentz transformations in terms of γ', l' and v'; but this time connecting S' to a different frame S'' moving with velocity v' with respect to S':

$$x'' = \gamma' l'(x' - v't'), \quad y'' = l'y', \qquad z'' = l'z', \qquad t'' = \gamma' l'(t' - v'x'),$$

where $\gamma = 1/\sqrt{1 - v^2}$, and $\gamma' = 1/\sqrt{1 - v'^2}$. Through the composition of the two transformations one obtains

$$x'' = \gamma'' l''(x - v''t), \quad y'' = l''y, \qquad z'' = l''z, \qquad t'' = \gamma'' l''(t - v''x),$$

where $v'' = \frac{v - v\prime}{1 - vv'}$, $\gamma'' = \gamma\gamma'(1 - vv') = 1/\sqrt{1 - v''^2}$, and $l'' = l'l$. These transformations, which connect the frame S with the frame S'', are also Lorentz transformations; so that the first requirement is satisfied. Notice that in the step which goes from S' to S'', the Lorentz transformation applied considers a velocity v' which is simply the relative velocity between the frames, the velocity with respect to the ether is not involved[73].

Yet another important result in Poincare's proof of the group property of the Lorentz transformations consists in his reasoning leading to l=1. He considered the case in which the systems S and S' get rotated in 180 about their y-axes. The transformations between the frames so rotated, which must also belong to the group, are

---

[73] "The crucial point is that if one considers three coordinate systems S, S', S" where S' and S" move along a common axis with uniform speeds ε with respect to S and ε' with respect to S', respectively, then the two Lorentz transformations from S to S' and from S' to S" can be replaced by a single Lorentz transformation from S to S" with the relative speed of S" with respect to S of $\varepsilon'' = \frac{\varepsilon - \varepsilon\prime}{1 - \varepsilon\varepsilon'}$", Miller 1986, 84.

$$x' = \gamma l(x + vt), \qquad y' = ly, \qquad z' = lz, \qquad t' = \gamma l(t + vx),$$

and it is assumed that the dependence of the factor l on the velocity v is not at all affected by replacing v with $-v$.

Then he considered the case of the inverse transformations:

$$x' = \frac{\gamma}{l}(x + vt), \qquad y' = \frac{y}{l}, \qquad z' = \frac{z}{l}, \qquad t' = \frac{\gamma}{l}(t + vx),$$

and noticed that the only way in which they can be a part of the group is by establishing that l=1. It is clear that if l has this value, the inverse transformations are identical to the transformations considered in the y-axe rotation case. Once again, the velocity that is relevant for the transformations and their inverses is simply the *relative* velocity (v or $-v$) between the frames, not the 'absolute' one with respect to the ether.

Two remarks are relevant with respect to this issue. First, notice the different way in which Lorentz and Poincare determined that the value of l has to be 1. Lorentz obtained it via dynamical considerations, for it was necessary for the expressions of the v-dependence of inertial mass obtained from the correspondence states theorem and the general contraction hypothesis to be equivalent to the one obtained from his model of the electron. Poincare, on the other hand, determined it by means of simple *mathematical* features of the transformations.

Second, Poincare *explicitly* showed that the correct interpretation of the transformations is symmetrical, i.e., that the relevant velocities involved are the relative ones. This feature was not originally noticed by Lorentz, he interpreted them as asymmetric –his view of the transformations was such that the transformation to *go back* to the ether-rest frame delivers *un*contracted lengths, for example, whereas Poincare showed that *both* systems S and S' determine that the lengths of bodies in (relative) motion get contracted. Curiously, as Janssen points out, Lorentz only saw this via Einstein, not via Poincare; even though he was certainly aware of his work –Lorentz openly accepted the introduction of the *Poincare-pressure*, for instance[74]. However, this amendment introduced by Poincare led to a feature that is crucial when an evaluative comparison between Lorentz's theory and SR is to be made. Even though the ether is an essential part of the ontology and of the conceptual content of the theory, the symmetry of the transformations imply that it becomes completely undetectable; and from a theoretical point of view, the symmetry also entails that the transformations do not even refer to the ether.

Before finishing this section and turning to Einstein's theory, I will make a brief and general remark about a fascinating and important issue: did Poincare independently discover SR? Some authors state that he did –Giedymin, Zahar and Whittaker[75]. The main arguments for this conception are Poincare's analysis of the measurement of time and the

[74] Janssen 1995, section 3.5.

[75] Zahar 2001, Ch. 4; Whittaker 1953, Ch. 2; Giedymin 1982, Ch. 5. I thank Professor Roberto Torretti for showing me how subtle this subject is and for his generous suggestions and corrections on my views about it.

determination of distant simultaneity, his amendments on the meaning of the Lorentz transformations, and the fact that he derived some *mathematical* results from the transformations that clearly prefigure Minkowski's space-time –he noticed that $ct^2 - x^2 - y^2 - z^2 = ct'^2 - x'^2 - y'^2 - z'^2$, so that the transformations can be understood not only as rotations around the y and z-axes, but also around the x-axis and a fourth axis it –and he also noticed that many physical quantities can be expressed as determined by four components, and that so expressed they remain invariant under Lorentz transformations.

In spite of the many results that Poincare obtained, and in spite of the many epistemological considerations which resemble some of the ones that Einstein also did; I think that Poincare did not discover SR. Even though he made a step towards it with his right foot, his left foot and his whole body stayed in the core of classical mechanics. His *relativistic glimpses* are only spots in a non-relativistic backdrop. To prove that, allow me to quote at length a passage that shows his commitment to an ether that stands still with respect to the motion of bodies in a Euclidian-space-through-absolute-time:

> [in order to account for the negative result of the ether-wind experiments] The most ingenious [hypothesis] was that of local time. Imagine two observers who wish to adjust their timepieces by optical signals; they exchange signals, but as they know that the transmission of light is not instantaneous, they are careful to cross them. When station B perceives the signal from station A its clock should not mark the same hour as that of station A at the moment of sending the signal, but this hour augmented by a constant representing the duration of the transmission. Suppose, for example, that station A sends its signal when its clock marks the hour *0*, and that station B perceives it when its clock marks the hour t. The clocks are adjusted if the slowness equal to t represents the duration of the transmission, and to verify it, station B sends in its turn a signal when its clock marks *0*; then station should perceive it when its clock marks t. The timepieces are then adjusted.
>
> And in fact they mark the same hour at the same physical instant, but on the one condition, that the two stations are fixed. Otherwise the duration of the transmission will not be the same in the two senses, since the station A, for example, moves forward to meet the optical perturbation emanating from B, whereas the station B flees before the perturbation emanating from A. The watches adjusted in that way will not mark, therefore, the true time; they will mark what may be called the *local time*, so that one of them will gain on the other. It matters little, since we have no means to perceive it. All the phenomena which happen at A, for example, will be late, but all will be equally so, and the observer will not perceive it, since his watch is slow; so, as the principle of relativity would have it, he will have no means of knowing whether he is at rest or in absolute motion.
>
> Unhappily, that does not suffice, and complementary hypotheses are necessary; it is necessary to admit that bodies in motion undergo a uniform contraction in the sense of motion. One of the diameters of the earth, for example, is shrunk by one two-hundred-millionth in consequence of our planet's motion, while the other diameter retains its normal length.[76]

This passage clearly shows that, according to Poincare, the exchange of light signals between the systems can be used to synchronize clocks, with the travel time *expressed as L/c* for each trip, only if the clocks are at rest with respect to the ether. If the procedure is used between clocks in relative rest, but which move with velocity v with respect to the ether, it synchronizes the clocks with respect to 'local time', but not with respect to the real time. If the moving-with-respect-to-the-ether observers in A and B want to achieve *real* synchronization, the procedure must be such that if the station A sends the signal at time *0* in

[76] From *L'État Actuel et l'Avenir de la Physique Mathématique* (1904); in Poincare 1958, 99-100.

its clock, the clock in station B must set its clock to the time $\frac{L\left(\sqrt{1-v^2/c^2}\right)}{c+v}$, and when the signal returns to A its clock should read $\frac{L\left(\sqrt{1-v^2/c^2}\right)}{c-v}$. However, this procedure cannot be applied, for the length contraction effect on the measurement of distance and the effect of local time on the readings of the clocks –which taken together determine that the velocity they measure for the light signals is always c–, entail that the results obtained will be the same as if they were at rest in the ether. It is clear that this line of reasoning is not relativistic at all, for it presupposes a Euclidian-space-through-absolute-time, and an immobile ether which determines certain dynamical effects that 'deceive' observers in motion with respect to it.

III. Special Relativity

Now I turn to Einstein's SR. In this section I offer an exposition of the theory as presented by Einstein in his two famous papers of 1905.

a) Motivation and the two principles

As I mentioned at the beginning of this chapter, there was a time when an essential connection was stated between the Michelson-Morley experiment and SR. That is, the former was thought to have been a direct motivation for Einstein to develop his theory. After the work of historians like Hirosige, Holton, Stachel, Miller and others, this view has finally been shown to be wrong. Einstein's motivation was not to solve the problem of the ether. What really took him to create a radically new theory were some foundational issues that he, unlike the rest of the scientific community –of which he was not a part ca. 1905–, understood as problematic.

More specifically, the main problematic feature he found out was that

> It is known that Maxwell's electrodynamics –as usually understood at the present time– when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena. Take, for example, the reciprocal electrodynamic action of a magnet and a conductor. The observable phenomenon here depends only on the relative motion of the conductor and the magnet, whereas the customary view draws a sharp distinction between these two cases in which either the one or the other of these bodies is in motion.[77]

If a conductor is considered as at rest in the ether, and a magnet moves with respect to it, then an electric current is induced; and if the magnet is considered as at rest in the ether and the conductor is in motion, exactly the same result obtains –an electric current is observed. However, electromagnetic theory provides a different explanation in each case. In the first one, the motion of the magnet creates an electric field which in turn causes the current; whereas in the second case, the current is the result of an electromotive force in the conductor produced by its motion with respect to the magnet, but no electric field is

[77] From *On the Electrodynamics of Moving Bodies*; in Lorentz *et al.* 1952, 37.

involved. The problematic aspect, for Einstein, was that in his view the only relevant feature for the phenomenon produced –the electric current– was simply the relative motion of the bodies involved. Such motion is totally symmetrical, of course; nevertheless, the theoretical explanation that electrodynamics provided was asymmetric.

After this example of the theoretical asymmetries that do not correspond to the observed phenomena, Einstein briefly refers to "the unsuccessful attempts to discover any motion of the earth relatively to the light medium", and states that these failed attempts "suggest that the phenomena of electrodynamics as well as of mechanics possess no properties corresponding to the idea of absolute rest"[78]. These two observations are the only reasons that Einstein mentions in his 1905 paper as leading him to the formulation of his *relativity principle*. That the only specific experiment he mentions is the one of electromagnetic induction and his generic reference to the failed attempts to detect ether-wind effects clearly point out that the Michelson-Morley experiment had no special relevance for the formulation of the theory. After all, this experiment is only one more among the unsuccessful attempts to discover any motion of the earth relatively to the light medium. This does not mean that the experiment was totally irrelevant for Einstein, of course; it only means that the problematic issue that really motivated him to create a new theory was a much more general one, a foundational flaw –of which the Michelson-Morley experiment was yet another instance–, rather than a specific empirical problem. Actually, the explanation that Lorentz's theory provided for its null result can easily be conceived in analogous terms with respect to the magnet-conductor case: in the ether rest frame, the pattern of interference obtained depends only on the velocity of light and the length of the arms of the interferometer; whereas in the moving frame, the very same interference pattern depends also on the length contraction and on local time, but the only observable difference is the relative motion of the interferometers[79].

Einstein, as a conclusion of his analysis of this foundational problematic issue, introduces the first of the two basic principles of the theory, the *relativity principle*: "the same laws of electrodynamics and optics will be valid for all frames of reference for which the equations of mechanics hold good"[80]. That is, all the laws of physics have the same form in all inertial frames. In classical mechanics this principle had been assumed and tested without any problems, but in electrodynamics, the fact that Maxwell's equations were interpreted as valid for the ether-rest frame implied that they should change their form in a frame in motion through the ether. In the previous section I showed how Lorentz's theory was able to

[78] *Ibid*, 37. The expression 'absolute rest' in this passage must be understood as 'rest with respect to the ether'. As the great scientists in electrodynamics had explicitly stated, Lorentz and Hertz for example, there was no reason to think that the ether was at rest with respect to absolute space, for there was no observable feature to distinguish this case from a case in which the ether is in uniform motion with respect to absolute space.

[79] In spite of some comments that Einstein made many years later, claiming that he did not know about the Michelson-Morley experiment by 1905, there are good reasons to think that he actually read Lorentz's 1899 work. In that case, it is clear that he knew about it. Moreover, there is a recently discovered document –notes taken from a talk that Einstein gave in 1921 at the Parker school in Chicago– which strongly indicates that, some time around 1899, he knew the experiment (see van Dongen 2009). However, it is also clear that he did not assign any special importance to it. For a historical survey of this issue, see Holton 1988 [1973], Stachel 1982, and Stachel 2002 and van Dongen 2009. I thank professor Torretti for drawing my attention to Stachel's beautiful 2002 paper.

[80] Lorentz *et al.* 1952, 37-8.

cope with the complete failure of detecting any effects of the Earth's motion through the ether, and how Poincare related this achievement to *his* principle of relativity The difference between Einstein's and Poincare's relativity principles consists in that, for the latter, it was the result of empirical tests and their theoretical explanation; that is, it was the outcome of dynamical features like the Lorentz-Fitzgerald contraction and local time. In the case of Einstein, the principle, even though empirically suggested, takes the form of a constrictive axiom of the theory, i.e., it is not an assumption to be explained. I will return to this difference below.

After these considerations Einstein –rather abruptly and without any further justification– introduces the second principle of the theory, the *light principle*: "light is always propagated in empty space with a definite velocity *c* which is independent of the state of motion of the emitting body"[81]. The lack of further comments by Einstein on his motivations to introduce this principle might suggest that the failed experiments on ether-wind effects also counted as the drive behind it, or that the first principle itself led him to the second. However, Stachel has compellingly shown that this is not the case[82]. The real motivation for the formulation of the principle was grounded in Einstein's thoughts on the electromagnetic nature of light. His famous *light-rider* thought experiment suggested him that a light ray looks the same to any observer, regardless of his velocity with respect to the emitting source. In a 1912 letter to Paul Ehrenfest, Einstein wrote:

> I well knew that the principle of the constancy of the velocity of light was something quite independent of the relativity principle; and I weighed which was more probable: the principle of the constancy of *c*, as required by Maxwell's equations, or the constancy of *c* exclusively for an observer at rest with respect to the source of light. I decided for the former, because I was convinced that any light is completely defined by frequency and intensity, quite independent of whether it comes from a moving light source or one at rest. It did not even enter my mind to imagine a deflected radiation propagated through a point could behave differently than radiation newly emitted at the point in question. Such complications seem to me much more unreasonable than those which the new concept of time involves.[83]

The second principle, just like the first one, is also a constraining axiom. Even though empirically justified –Einstein did not so consider it, but the negative results of ether-wind experiments in optics could be interpreted as empirical support for it– it is not an assumption that requires an explanation. Once again, this differentiates Einstein's approach from Lorentz's and Poincare's. In their theory, the *measured* velocity of light is *c* in all inertial frames; but in the ones which move with respect to the ether this measurment is a 'deception', for it is the outcome of compensating-conspiring dynamical effects. The velocity of light is *really c* only in the ether-rest frame.

[81] *Ibid*, 38.
[82] Stachel 1981 and 2002.
[83] Quoted in Stachel 1982, 51 (footnote 29).

b) Relative simultaneity

Taken in isolation, both principles are rather natural even from the point of view of classical mechanics and electrodynamics. The first one is part of the very core of classical mechanics; and in electrodynamics, as Lorentz and Poincare showed, it obtained as the outcome of compensating dynamical effects. The second principle is totally consistent with electromagnetic theory, but only for one special inertial frame, the ether-rest frame; only in it light is emitted with constant velocity *c* regardless of the state of motion of the source. In any other inertial frame, the speed of light must respect the Galilean law of addition of velocities: it is the sum of *c* as defined in the ether-rest frame and the velocity of the moving frame with respect to the ether. But if the two principles are taken together it follows that the speed of light is *c* in *any* inertial frame –and therefore, there is no privileged inertial frame any more– regardless of their relative velocity and regardless of the state of motion of the source. This is a blatant violation of the classical expression for the addition of velocities that I just mentioned.

However, and this is one of the (many) genius insights that Einstein had, he noticed that the contradiction is only apparent; there is no real contradiction at all. The law for the addition of velocities, Einstein found out, rests upon a certain definition of distant simultaneity that is reflected in the methods to determine it, and that presupposes physical assumptions:

> If we wish to describe the motion of a material point, we give the values of its co-ordinates as functions of the time. Now we must bear carefully in mind that a mathematical description of this kind has no physical meaning unless we are quite clear as to what we understand by "time". We have to take into account that all our judgments in which time plays a part are always judgment of *simultaneous events*. If, for instance, I say, "That train arrives here at 7 o'clock", I mean something like this: "The pointing of the small hand of my watch to 7 and the arrival of the train are simultaneous".
> […] In fact such a definition is satisfactory when we are concerned with defining a time exclusively for the place where the watch is located; but is no longer satisfactory when we have to connect in time series of events occurring at different places, or –what comes to the same thing– to evaluate the times of events occurring at remote places from the watch.[84]

This passage shows that Einstein realized that the determination of the simultaneity between two *distant* events is an *inference*, rather than an *a priori* relation defined in terms of absolute time[85]. In order to provide a sound method according to which simultaneity between distant events can be determined –and that does not require any knowledge of the

[84] Lorentz *et al.* 1952, 39.

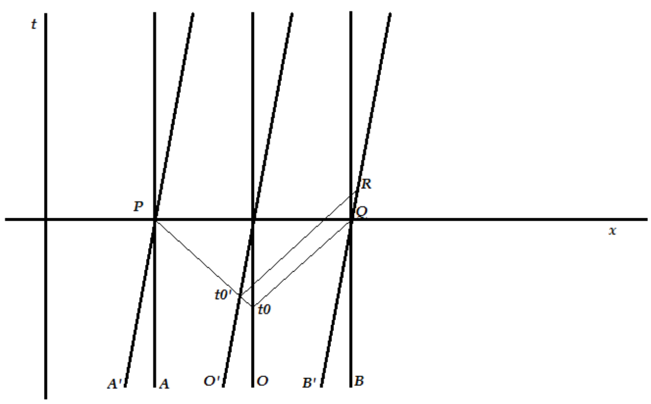
[85] With respect to Einstein's motivations and philosophical background for this insight, Stachel comments: "Here, I believe, Einstein was really helped by his philosophical readings. He undoubtedly got some help from his readings of Mach and Poincare, but we know that he was engaged in a careful reading of Hume at about this time; and his later reminiscences attribute great significance to his reading of Hume's *Treatise on Human Nature*. What could he have gotten from Hume? I think it was a relational –as opposed to an absolute– concept of time and space. This is the view that time and space are not to be regarded as self-subsistent entities; rather one should speak of the temporal and spatial aspects of physical processes; "The doctrine," as Hume puts it, "that time is nothing but the manner in which some real object exists". I believe the adoption of such a relational concept of time was a crucial step in freeing Einstein's outlook, enabling him to consider critically the tacit assumptions about time going into the usual arguments for the 'obvious' velocity addition law". Stachel 2002,

distance between the events or between the clocks– Einstein proposes his famous method to synchronize distant clocks:

> If at the point A of space there is a clock, an observer at A can determine the time values of events in the immediate proximity of A by finding the positions of the hands which are simultaneous with these events. If there is at the point B of space another clock in all respects resembling the one at A, it is possible for an observer at B to determine the time values of events in the immediate neighborhood of B. But it is not possible to compare, in respect of time, an event at A with an event at B. We have so far defined only an "A time" and a "B time". We have not defined a common "time" for A and B, for the latter cannot be defined at all unless we establish *by definition* that the "time" required by light to travel from A to B equals the "time" it requires to travel from B to A [this *definition* is exactly what the second postulate allows]. Let a ray of light start at the "A time" t_a from A towards B, let it at the "B time" t_b be reflected at B in the direction of A, and arrive again at A at the "A time" t'_a.
> In accordance with the definition the two clocks synchronize if $t_b - t_a = t'_a - t_b$. We assume that this definition of synchronism is free from contradictions, and possible for any number of points.[86]

This method for the synchronization of distant clocks, and to determine distant simultaneity, is the key to dissolve the 'contradiction' between the postulates. The method is *by definition* consistent with the two principles, for the value of the velocity of light is assumed to be c in *all* inertial frames[87]. Moreover, as we will see, the kinematics this definition of simultaneity entails –along with the two principles– contain a law for the *composition* of velocities which is not the Galilean one –though when v is much smaller than c the relativistic law reduces to the Galilean formula– and that is totally consistent with the principles of the theory.

A very important consequence that follows from the two postulates and the synchronization method just presented is the *relativity of simultaneity*. That is, a statement like *two events are simultaneous* becomes meaningful only *with respect to a specific inertial frame*. If we consider two inertial frames in relative motion, two events that are simultaneous in one of them will not be so in the other one. This result can be better visualized in the following figure:



[86] Lorentz *et al.* 1952, 39-40.

[87] "In agreement with experience [and with the second postulate] we further assume the quantity $\frac{2AB}{t'_a - t_a} = c$, to be a universal constant –the velocity of light in empty space" *Ibid*, 40.

A, *O* and *B* are three observers at rest with respect to each other, the distance from *O* to *A* is the same as the distance from *O* to *B*; and the three of them carry synchronized clocks according to the method just described. *A′*, *O′* and *B′* are three other observers at rest with respect to each other, and *A′* and *B′* are equidistant with respect to *O′*. As the diagram shows, these three observers are in inertial motion with respect to *A*, *O* and *B*; and they are also equipped with clocks synchronized by the same method. At t_0, *O* emits a light ray towards *A* and another towards *B*. At t_0' the light ray from *O* to *A* reaches *O′*, and at that same instant *O′* emits a light ray towards *A′* and another towards *B′*. Since *A* and *B* are equidistant with respect to *O*, the events which coincide with the arrival of the light rays sent from *O* at t_0 are simultaneous, i.e., when the light ray reaches *A*, its clock marks the same time as the time that the clock in *B* marks when it receives the other light ray. On the other hand, since *A′* and *B′* are equidistant with respect to *O′*, the events which coincide with the arrival of the light rays sent from *O′* at t_0' are simultaneous; when the light ray reaches *A′* its clock marks the same time as the time that the clock in *B′* marks when it receives its corresponding light ray. The events at which the light rays towards *A* and *A′* arrive are the same, namely, *P*. However, the events *Q* in *B* and *R* in *B′* that coincide with the arrival of their light rays are not the same. The situation is then that according to the clocks in *A*, in *O* and in *B*, *P* and *Q* are simultaneous events; but according to the clocks in *A′*, in *O′* and in *B′*, the events *P* and *R* are simultaneous. The explanation for this discrepancy is that simultaneity is relative. Simultaneous events for observers *A*, *O* and *B* are not so for the observers *A′*, *O′* and *B′* –and the other way around– because they are in relative inertial motion[88].

The relativity of simultaneity has consequences also for how the lengths of bodies are conceived in SR. The length of an object, say, a rod, is defined as the spatial distance between two simultaneous events: the ones that coincide with its end points. But we just saw that inertial motion between two frames determines that the events which are simultaneous in one of them will not be so in the other one. Therefore, the events which define the length of an object in one of the frames are not the same events which define the length of the same object in the other frame. This means that the relativity of simultaneity implies the relativity of length.

c) Lorentz′ transformations derived

These considerations are mainly qualitative. Einstein's next step was to obtain the specific quantitative transformations that relate coordinates of events in one inertial frame with their corresponding coordinates in a different frame. The first postulate can be stated as saying that any experiments whatsoever will have the same results in any inertial frame. This implies that result of an experiment will be the same even if its initial conditions differ only in terms of a translation and/or rotation in some inertial frame. Moreover, identical

[88] Or more precisely, since the inertial rest frame of *A*, *O* and *B* is in a state of inertial motion with respect to the inertial rest frame of *A'*, *B'* and *O'*; simultaneous events in one of them are not so in the other. This statement also precludes any misunderstanding about the subjectivity of the observers as being involved in this issue. Observers could be totally dispensed of and the result would be the same.

experiments carried out in inertial frames at different times will also yield the same outcomes. These two remarks mean that the first postulate implies the homogeneity and isotropy of space and time:

> It will also turn out, as a direct consequence of the relativity principle, that all inertial frames are spatially homogeneous and isotropic, not only in their assumed Euclidean geometry but for the performance of all physical experiments. By this we mean that the outcome of an experiment is the same whenever its initial conditions differ only by a translation (homogeneity) and rotation (isotropy) in some inertial frame.
>
> Again, as a consequence of the relativity principle, it will presently turn out that inertial frames are temporally homogeneous, i. e., that identical experiments (relative to a given inertial frame) performed at different time yield identical results.[89]

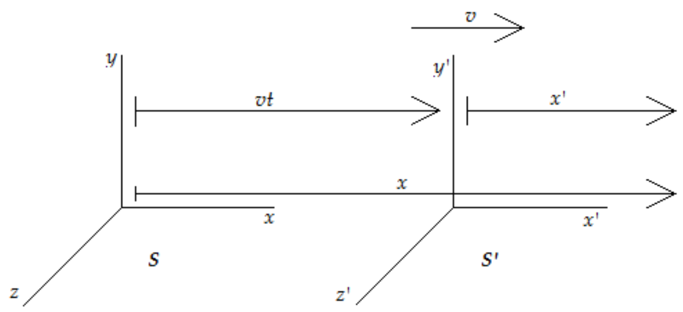
The spatiotemporal homogeneity and isotropy implied by the relativity postulate leads to a constraint in the nature of the coordinate transformations connecting events in different inertial frames: the transformations must be linear[90]. If they were not so, then, for example, the coordinates of a freely moving particle in one inertial frame, when transformed to a different one, would not yield an inertial description of its state of motion in the second frame; and this would be a blatant violation of the first principle of the theory[91].

To obtain the transformations in a simple way a harmless assumption can be made: the frames to be related by the transformations can be considered as in *standard configuration*, that is, they satisfy the following conditions: *i)* the motion of the frame S' occurs only along the x-axis of the frame S –and that the planes defined by the y=0 and z=0 coordinates of the frames coincide with the planes defined by y'= 0 and by z'= 0; *ii)* when the origins of the two frames coincide the time coordinate of that event in both frames is 0; *iii)* that the coordinate plane in the moving frame S' defined by x'= 0 coincides with the plane in S defined by $x = vt$ –where v is the velocity of S' with respect to S–; and *iv)* the transformations are invariant under a reversal of the x and z-axes in both frames and an interchange of primed and unprimed coordinates –the same holds for reversal around x and y-axes. In simple words, the frames considered for the derivation of the transformations are two identical frames S and S' –with their origin and their three axes coinciding–, but such that at the instant 0 of both frames the frame S' is set in uniform motion with velocity v with respect to S along its x-axis:

[89] Rindler 1991 [1982], 6-7.

[90] Einstein's statement that the linearity follows from the homogeneity and isotropy of space and the homogeneity and isotropy of time, taken separately, has been challenged. According to Torretti (1996 [1983], § 3.6), for instance, the correct view is that the homogeneity and isotropy of *space-time* is what really entails the linearity of the transformations. For example, it can be shown that the coordinate transformations between a resting and a uniformly rotating system are *not* spatiotemporally homogeneous, even though they are spatially homogeneous and temporally homogeneous. Spatiotemporal homogeneity requires that a space-time location-independent variation applied to the coordinates of one system leads to a space-time location-independent variation in the coordinates of another. This is stronger requirement than the conjunction of spatial homogeneity (that a spatial location-independent variation in the coordinates of one system leads to a spatial-location-independent variation in another) and temporal homogeneity (that a time-independent variation in the coordinates of one system leads to a time-independent variation in another).

[91] For a formal proof that the transformations have to be linear, see Rindler 1991 [1982], 11.



If we begin for the transformation for the y to y' coordinates, the linearity condition states that $y' = Ax + By + Cz + Dt + E$, where the coefficients are v-dependent constants. The first assumption above implies that if $y = 0$, then $y' = 0$; therefore $y' = By$. Applying the x-z invariant reversal we have that $y = By'$, and then $B = \pm 1$; but if the velocity v tends to 0 the transformation must lead to an identity transformation, so that B can only be 1. The resulting transformation is then $y' = y$; and completely analogous reasoning leads to $z' = z$.

Turning now to the x and x' coordinates, we have that because of linearity, $x' = Ax + By + Cz + Dt + E$. Assumption *iii)* above implies that if $x = vt$, then $x' = 0$; and therefore the transformation reduces to $x' = A(x - vt)$ and an x-y reversal yields $x = A(x' + vt')$. In a classical mechanics scenario, $t' = t$, and that would lead to $A = 1$, which in turn yields the Galilean transformations. However, Einstein's second postulate entails that the coordinates of the same light ray in S and S' are given by $x = ct$ and $x' = ct'$ respectively. These expressions can be plugged in the transformations so that one obtains $ct' = At(c - v)$ and $ct = At'(c + v)$. Then, the following equation can be set, $ct \cdot ct' = At'(c + v) \cdot At(c - v)$, and solving for A it follows that $A = \frac{1}{\sqrt{1 - v^2/c^2}}$. It is quite clear that the coefficient A is mathematically identical to the factor γ in the Lorentz transformations. Finally, t' can be determined by replacing ct', ct, and x/c, for x', x, and t respectively in $x' = A(x - vt)$, so that $ct' = A(ct - v\frac{x}{c})$; and solving for t' it follows that $t' = A(t - \frac{vx}{c^2})$.

Summarizing, and adapting the notation, from Einstein's two postulates –plus the assumptions about the configuration of the frames– a set of coordinate transformations follow which are mathematically equivalent to the Lorentz transformations:

$$x' = \gamma(x - vt), \qquad y' = y, \qquad z' = z, \qquad t' = \gamma(t - vx/c^2);$$

and assumption *iv)* above entails, just as Poincare showed, that the transformations are symmetric, i. e.:

$$x = \gamma(x' + vt'), \qquad y = y', \qquad z = z', \qquad t = \gamma(t' + vx'/c^2).$$[92]

[92] The derivation of the transformations I just presented is not the one that Einstein performed in § 3 of his 1905 paper, but a simpler one in Rindler 1991 [1982], 11-6. For a detailed analysis and commentary of Einstein's own derivation, see Miller 1998 [1981], 195-205; and Torretti 1996 [1983], § 3.4.

In spite of their mathematical equivalence, the Lorentz transformations as conceived by Lorentz-Poincare and as conceived by Einstein are rather different in their physical meaning. The main difference is that for the Dutch and the French they are the *outcome* of a set of dynamical effects, the Lorentz-Fitzgerald contraction and local time; whereas Einstein did not introduce any dynamical grounds for his *derivation* of the transformations. In a word, the Lorentz transformations, for Lorentz and Poincare, are *dynamically* grounded; whereas for Einstein they are *kinematically* grounded. For the Dutch and the French, the interaction between matter and the ether underlies them. Einstein does not make any assumption about the ultimate nature of matter, the two postulates are enough; and therefore, the ether becomes superfluous[93]. This point of view allows Einstein, unlike Lorentz and Poincare, to conceive that the only physically relevant velocity involved in the transformations is the *relative one* between the frames. As we saw above, even though Poincare showed that the transformations are mathematically symmetric, the ether was still underlying their physical meaning. On this respect Miller comments:

> In the (1895) or (1904), Lorentz's plausibility argument for the Lorentz contraction hypothesis involved cross-multiplying the quantities $\sqrt{1-v^2/c^2}$ in the spatial portions of the S'' and S'_r transformations, respectively. But since in special relativity K and k were equivalent, then Einstein could move between k and K by changing v to $-v$, and interchanging Greek and Roman letters. For Lorentz in 1904 this interchange had no physical meaning because the system K was fixed in the ether. In 1905 Poincare attributed only a mathematical interpretation of the reciprocity property of the Lorentz transformations –that is, reciprocity corresponded to a rotation of K and k by 180 about their common y-axes.[94]

With the coordinate transformations already presented, the meaning of the relativity of simultaneity and the relativity of length becomes much more precise and quantitative. Consider first the relativity of length. Frames S and S' are configured in the same way I assumed above. A rod of length $\Delta x'$ lies at rest in the x'-axis of S'. We want to find out what is its length Δx in the S frame. This length is given by the spatial distance between two simultaneous events in S which coincide with the end-points of the rod in S. The formula we need is the Δ-form of the transformation from x' to x, that is, $\Delta x' = \gamma(\Delta x - v\Delta t)$. The requirement of the simultaneity in S of the events which determine the length of the rod implies that $\Delta t = 0$. Therefore, the length of the rod Δx in S is equal to the length of the rod $\Delta x'$ in S' divided by γ. In other words, $L_s = \sqrt{1-v^2/c^2}\, L_{s'}$. This means that observers in S, for whom the rod is in motion, will measure it as contracted by the factor $\sqrt{1-v^2/c^2}$ with respect to its rest-length or *proper length* in S' –the frame in which the velocity of the body is 0 measures the largest possible length, and this largest possible length is called the *proper length* of a body. It is clear that the grounding of this contraction is the relativity of

---

[93] In Einstein's own words: "These two postulates suffice for the attainment of a simple and consistent theory of the electrodynamics of moving bodies based on Maxwell's theory of stationary bodies. The introduction of a 'luminiferous ether' will prove to be superfluous inasmuch as the view here to be developed will not require an 'absolute stationary space' provided with special properties"; Lorentz *et al.*, 38. With respect to the *absolute stationary space* Einstein refers to, see footnote 78 above.

[94] Miller 1998 [1982], 204. The frames S'' and S_r correspond to the frames S and S' in Lorentz's three-step method involving S_0, S and S', according to the notation I used. Einstein's K and k frames correspond to S and S' in the derivation I just presented, and the interchange between Greek and Roman letters corresponds to the interchange between primed and unprimed coordinates I mentioned.

simultaneity, not dynamical effects such as the Lorentz-Fitzgerald contraction which results of the motion of objects across the ether. This can be noticed by considering that if the rod were at rest in S, in that case its length in S' would be contracted with respect to its proper length in S.

Now we consider the case of time in two different inertial frames in relative motion. Once again let us assume that S and S' are configured as before. Suppose that a clock w is fixed in S and that two events at that clock –their spatial coordinates are the same– are separated by Δt according to that clock. We want to find out what is the $\Delta t'$ between those two events as marked by a clock w' stationary in S'. We use the Δ-form of the transformation for the time coordinate $\Delta t' = \gamma(\Delta t - v\Delta x/c^2)$, and since $\Delta x = 0$ then $\Delta t' = \gamma\Delta t$. This means that for the observer in S for which the clock w' is in motion, such clock measures a time interval between the two events which is dilated by a factor $\sqrt{1 - v^2/c^2}$ with respect to the *proper time* that her stationary clock w measures between the same two events –the frame in which the spatial distance between the events is 0 measures the shortest possible time-interval between them, i.e., their *proper time*. The formal expression for this *time dilation* effect is thus $T_{S'} = \frac{T_S}{\sqrt{1-v^2/c^2}}$. More generally, clocks that move in an inertial frame go slower than clocks at rest in that same frame. Just like length-contraction, this is a *kinematically* grounded effect. There are no *dynamical* processes affecting the rates of the clocks. This can be easily seen by shifting the roles between the frames. In S the moving clock w' goes slower with respect to the stationary clock w, but in S' the moving clock w goes slower with respect to the stationary clock w'.

I will now mention some other consequences of the coordinate transformations that Einstein obtained –in order to make a case for the predictive equivalence of his theory with respect to Lorentz's. First, the expression for the composition of velocities: assume frames S and S' to be configured in the standard way, so that the x-velocity of a particle is given in S by $u_x = \frac{dx}{dt}$ and in S' by $u'_x = \frac{dx'}{dt'}$. From the coordinate transformations we have that $dx' = \gamma(dx - vdt)$, and that $dt' = \gamma(dt - vdx/c^2)$. Plugging the right hand of these equations in the expression for u'_x, then $u'_x = \frac{dx - vdt}{dt - v\frac{dx}{c^2}}$ obtains. Comparing this result with the expression for u_x, it follows that $u'_x = \frac{u_x - v}{1 - u_x v/c^2}$. If we remind that in his derivation Poincare set c to 1, it is obvious that the velocity composition law he obtained is identical to Einstein's –which is rather natural since both were carried out from identical coordinate transformations.

This last result was used in 1907 by Laue to derive the Fresnel coefficient in a very simple way. Suppose a container filled with a transparent medium with a refractive index $n = c/u'$, where $u' = c/n$ is the velocity of light in the medium when it is considered at rest in the ether. Recall that Fresnel found out that if the medium is set in motion –with respect to the ether– with a velocity v, the velocity u of the light ray in the refractive medium is given by $u = u' + v(1 - 1/n^2)$, where the term in brackets is Fresnel's drag coefficient. What Laue did was to show that this last formula, in the context of SR, is nothing but a consequence of the velocity composition law. The derivation is quite simple and very meaningful. First, the ether-rest frame does not play any role, v is simply the relative velocity between frames S and S'. From this perspective, u is the velocity of light across the refractive medium as measured in S –for which the container moves with velocity v–, and u' is the velocity of light in the refractive medium as measured in S' –for which the container is at rest. Therefore, $u = \frac{u'+v}{1+u'v/c^2}$[95]. This last formula can be approximated, neglecting terms of second order of v/c, to $u = (u' + v)(1 - \frac{u'v}{c^2})$ and then to $u = (u' + v)(1 - \frac{u'^2}{c^2})$. Comparing the second factor with the formula for n, it turns out that $u = u' + v(1 - 1/n^2)$, in agreement with Fresnel's coefficient. The fact that this derivation was carried out only from the law of composition of velocities indicates that, in the context of SR, Fresnel's drag is a kinematically grounded effect that does not need any dynamical explanation in terms of the interaction between light and the medium on the one hand, and the ether on the other[96].

Now I will dedicate a few words for the velocity dependence of inertial mass in SR. Einstein's first step in his derivation was to consider a frame S in which an electron is at rest at a time t_0, but in motion at the *next instant of time* t_1 –that is, both times differing by an infinitesimal amount. The motion of this electron in the frame S is described by Newton's formula $\boldsymbol{F} = m\boldsymbol{a}$, and the net force $\boldsymbol{F}$ is given by the influence of an external electric field $\boldsymbol{E}$ on the electron charge e. Therefore, the equations of motion for the electron in the frame S are

$$m_0 d^2x/dt^2 = eE_x \qquad m_0 d^2y/dt^2 = eE_y \qquad m_0 d^2z/dt^2 = eE_z;$$

with the proviso that the motion of the electron is slow –the reason for this condition is to neglect any change that the relativity of simultaneity could produce in the formulation of the Newtonian law– and where m_0 is the inertial mass m in Newton's law.

The next step was to suppose that at the instant t_0 the electron is moving with velocity v with respect to the frame S, and that the frame S' is such that at that same instant the electron is at rest. According to the relativity postulate, at the *next instant of time* in S', the equations of motion for the electron will have the same form as the equations above –expressed in terms of x', y', z', and t'; of course. By applying the coordinate transformations to the equations of motion in S' it follows that in the frame S,

$$\frac{d^2x}{dt^2} = \frac{e}{m_0\gamma^3}E_x \qquad \frac{d^2y}{dt^2} = \frac{e}{m_0\gamma}(E_y - \frac{v}{c}B_z) \qquad \frac{d^2z}{dt^2} = \frac{e}{m_0\gamma}(E_z - \frac{v}{c}B_y).$$

In order to obtain the expressions for the transverse and longitudinal mass of the electron, Einstein formulated the last equations in the following way:

---

[95] This is, of course, the inverse velocity transformation, in which -v replaces v, and in which the primed and unprimed quantities have shifted roles.

[96] Remember that though Lorentz's 1895 derivation of Fresnel's coefficient was rather general and it did not presuppose any electromagnetic features, in 1886 he had provided a derivation in that the drag of light was dynamically explained in terms of the interaction of the constituting molecules of the transparent medium and the ether. His 1895 derivation cannot be a rejection of this explanation, for it was crucial for his assumption of an immobile ether –the electromagnetic explanation he provided for Fresnel's coefficient was not based on an ether-partial-drag.

$$m_0\gamma^3 \frac{d^2x}{dt^2} = eE_x = eE_{x'}$$

$$m_0\gamma^2 \frac{d^2y}{dt^2} = e\gamma(E_y - \frac{v}{c}B_z) = eE_{y'}$$

$$m_0\gamma^2 \frac{d^2z}{dt^2} = e\gamma(E_z - \frac{v}{c}B_y) = eE_{z'}$$

The terms in the left hand of these expressions are the product of mass and acceleration, so that the formulas for the longitudinal and transverse mass become $m_L = m_0\gamma^3$ and $m_T = m_0\gamma^2$. If one compares the value of the *transverse* mass that Lorentz obtained with the one that Einstein derived it turns out that they are not the same. But the difference is only apparent. The middle and right-hand terms in the equations above represent the net force in the frame S and in the frame S' respectively, so that $\boldsymbol{F} = \boldsymbol{F'}$. As we saw above, in the case of Lorentz's theory, the equivalence between the forces in the frames only holds in the case of the x and x'-axes –and that is the reason why the *longitudinal* masses are equivalent in both theories. However, in Lorentz's theory the force equality does not hold in the two other cases. Therefore, the discrepancy between Lorentz's and Einstein's transverse masses is only a matter of a different definition of force. This is obvious when one considers that the equations above, for the y and z-axes, can be written as

$$m_0\gamma \frac{d^2y}{dt^2} = e(E_y - \frac{v}{c}B_z) \qquad m_0\gamma \frac{d^2y}{dt^2} = e(E_y - \frac{v}{c}B_z).$$

In this case, the value of the transverse mass, just as in Lorentz theory, becomes $m_0\gamma$; and the equality between F_y and $F'_{y'}$ –and also the corresponding equality correspondent to the z and z'-axes– is no longer the case. But this is not a problem, for the forces in the two frames remain connected via the coordinate transformations. Therefore, the velocity dependence of the inertial mass is the same both in Lorentz's and Einstein's theories. However, notice that, unlike Lorentz, Einstein did not speculate about the ultimate nature of the electron mass, its rest-mass m_0 is just taken as a given feature.

d) Mass and energy

Finally, I will turn to the relation between mass and energy contained in SR. Einstein published this result in September 1905 in a paper entitled *Does the Inertia of a Body Depend upon its Energy Content?* –three months after the publication of *On the Electrodynamics of Moving Bodies*. The derivation of the relation at issue was grounded on a result he had already obtained in the latter article. Suppose that a plane wave of light possesses the energy l in a frame S; and that the direction of the ray makes an angle φ with respect to the x-axis of S. Consider now a frame S' moving with velocity v with respect to S –and they are configured in the standard way. The value of the energy of the system of plane waves in S' is given by $l' = \frac{l(1-\frac{v}{c}\cos\varphi)}{\sqrt{1-v^2/c^2}} = \gamma l(1-\frac{v}{c}\cos\varphi)$.

Now Einstein proposes the following thought experiment: a body at rest in S simultaneously emits a light wave of energy $L/2$ at an angle φ with respect to the x-axis, and another light wave with the same energy in the opposite direction, so that after the emission the body remains at rest. E_0 is the energy of the body before the emission and E_1 is its energy after it. Considering energy conservation, then $E_0 = E_1 + L$.

Turning to the description of this same situation in the frame S', and applying the formula for the transformation of energy of the light wave he had already obtained, it follows that in the S' frame $E'_0 = E'_1 + \frac{1}{2}L\gamma\left(1-\frac{v}{c}\cos\varphi\right) + \frac{1}{2}L\gamma\left(1+\frac{v}{c}\cos\varphi\right) = E'_1 + \gamma L$, where E'_0 and E'_1 are the energies of the body before and after the emission, respectively.

The difference between E'_0 and E_0, that is, $E'_0 - E_0 = E'_1 - E_1 + (\gamma-1)L$, can be related to the kinetic energy of the body before and after the emission in S'. According to Einstein, the difference $E' - E$ is equal to the kinetic energy of the body in S', for the latter is defined as the difference between the body's energy when in motion and when at rest, which in this case can be represented as E' and E, respectively[97]; so that $K_0 = E'_0 - E_0$, and $K_1 = E'_1 - E_1$ –where K_0 and K_1 are the kinetic energies of the body before and after the emission, respectively. Therefore, $K_0 - K_1 = L(\gamma - 1)$; which means that the kinetic energy of the body in S' diminishes as the outcome of the emission of the light wave.

The connection of this result with the mass of the body follows from the fact that, neglecting terms of fourth and higher orders of v/c, $K_0 - K_1 = \frac{1}{2}\frac{v^2}{c^2}L$; and from this equation it follows that "*if a body gives off the energy L in the form of radiation, its mass diminishes by L/c*[2]. The fact that the energy withdrawn from the body becomes energy of radiation evidently makes no difference, so that we are led to the more general conclusion that the mass of a body is a measure of its energy content"[98].

IV. Comparing the theories

Now that both theories have been presented, I turn to an evaluative comparison between them. First I will deal with some subtleties that need to be considered if they are going to be understood as predictively equivalent. Secondly, I will outline why the two theories are *different* and *rival*s.

---

[97] The rationale for the possibility of considering E –which is the rest-energy of the body *in the frame S*– as the rest energy of the body *in the frame S'*, relies on the relativity postulate: the energy of a body at rest must be the same in *any* frame.

[98] Lorentz *et al.*, 1952, 71. Einstein is assuming that, according to the binomial theorem, $\gamma = (1-v^2/c^2)^{-1/2} = 1 + \frac{1}{2}(v/c)^2 + \frac{3}{8}(v/c)^4 + \cdots$; and that –with $v \ll c$– $K = \frac{1}{2}mv^2$; so that $K_0 - K_1 = \frac{1}{2}\frac{v^2}{c^2}L$ reduces to $m_0 - m_1 = L/c^2$ –where m_0 and m_1 are the masses of the body before and after the emission, respectively. It is important to underscore that the low-velocity assumption does not imply that mass-energy relation is approximate, for "all the quantities in equation (7.9) [$m_0 - m_1 = L/c^2$] are measured in the body's rest frame. The relation between them cannot depend on the velocity of the auxiliary frame S' in which the kinetic energy is expressed. We are free to assign to S' any velocity we please. Hence the result is rigorously true"; Sartori 1996, 205.

a) Empirically equivalent

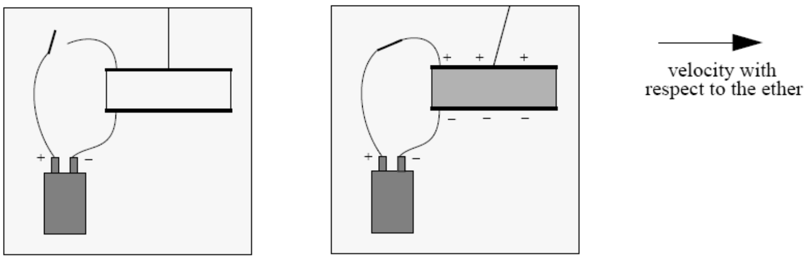
At first sight it looks pretty clear that Lorentz's ether theory and Einstein's SR are predictively identical. As I showed above, both predict an equal velocity dependence of inertial mass, both allow to derive Fresnel's drag coefficient, and they both entail a longitudinal length contraction and effects on the readings of clocks as the outcome of the state of motion of bodies. Even though these features have a different physical meaning in each theory, their mathematical forms and values are equal.

The root of this issue lies of course in the identical coordinate transformations that the theories include. Since the empirical predictions of both theories are entailed by these transformations it is quite natural that if the transformations are identical then the predictions will also be. Moreover, if we consider that the very core of SR is given by its two postulates, we can find a sort of *Lorentzian* version of them in the other theory. As I showed in the sections dedicated to Poincare's contributions, the French scientist explicitly formulated a *principle of relativity* that he regarded as a requirement that Lorentz's theory should accomplish in order to be fully satisfactory. That is, Poincare thought that Lorentz's theory had to be formulated in a way such that the outcome of *all* experiments set out to measure any kind of ether-wind effect should be negative: Poincare demanded a full Lorentz-invariance for the laws of physics. We saw that the main difference between Poincare's and Einstein's relativity principle was that for the former it was the result of a set of dynamical-compensating effects that precluded the possibility to detect any sort of alteration in the laws of physics; whereas for Einstein it was simply an axiomatic constraint of the theory which did not need any kind of dynamical explanation.

Something similar holds for the light postulate. Einstein simply took the principle that light has a constant velocity c in any direction independently of the state of motion of the source as a point of departure. Together with the first postulate, it follows that the velocity of light is c in all directions in *any* inertial frame, not only in the ether-rest frame. I also showed that the commitment to these two principles requires a radical reconsideration of the nature of simultaneity and time. In the case of Lorentz's theory, the analogous feature with respect to the light postulate is that the same compensating dynamical effects that resulted in the Poincarean version of the relativity principle, also determined that in any frame moving inertially with respect to the ether the *measured* value of the velocity of light is c. Its real velocity is $c+v$, but the dynamical effects inexorably *deceive* the observer, he has no way to find out its real value. One of the compensating effects that participate in this conspiracy is local time. Analogously to what happens in SR, in Lorentz's theory the events that observers in a frame that moves with respect to the ether describe as simultaneous are not so for observers in a different frame with a different velocity with respect to the ether. However, and unlike Einstein's theory, Lorentz's theory describes this effect as a *deception*, for the *real* time measurements can be carried out only in the ether-rest frame.

So far, so good, but what about $E = mc^2$? The relationship between mass and energy was not formulated as an explicit result in Lorentz's theory, and therefore this result must be carefully considered if the empirical equivalence of the theories is to be argued. Actually, if one considers only Lorentz's view of his own theory, the absence of the famous formula seems to imply that, after all, there is a crucial experiment to ground an empirically based decision –namely, the Trouton experiment.

In 1900 Frederick Trouton, based upon an original idea by George F. Fitzgerald, designed an ingenious experiment in order to look for an ether-wind effect. The interesting feature of this particular experiment was that, unlike many of the attempts to detect an effect of the motion of the Earth across the ether, it was not based on optics: according to Fitzgerald, if a capacitor in motion through the ether is charged or discharged it should suffer an impulse. Trouton's experiment was designed to measure this effect. In a terrestrial laboratory –which, of course, moves with respect to the ether– a hanging capacitor at rest is connected to a battery. When the battery is switched on an electromagnetic field is produced between the plates of the capacitor. If it were at rest in the ether, only an electric field would be induced, but its motion through the ether adds the generation of a magnetic field. Fitzgerald reasoned that the extra energy needed to produce the magnetic field had to come from the kinetic energy of the capacitor, and the loss of kinetic energy must result in a jolt in the direction opposite to the motion across the ether. The actual display of the experiment was such that the capacitor was a part of a torsion pendulum, and the charges and discharges were done by a clock-work at time intervals corresponding to the free period of swing of the pendulum, so that the jolt effect would accumulate with its natural oscillation and so become more easily observable. A simple sketch of the experiment is depicted in the following figure[99]:



Just as in every ether-wind experiment performed, the result of Trouton's was negative, no impulse in the capacitor was observed. In any possible outcome, though, this experiment posed a deep theoretical challenge. Let us recall that Newton's third law is closely associated with the law of conservation of momentum and with the center of mass theorem. If the outcome of the Trouton experiment is positive, the jolt of the capacitor constitutes a clear violation of the center of mass theorem[100]; but if the result is negative, then the law of conservation of momentum is violated. The dilemma consists in that it seems that

[99] Taken from Janssen 2003, 31.
[100] This theorem states that no process in an isolated system can change the state of motion of the system's center of mass: for any system with no external forces, the center of mass moves with constant velocity.

a theoretical explanation of the experiment necessarily implies the abandonment of one of these two core tenets of classical physics.

As I showed above, Poincare had already found a similar problem. In 1900 he introduced a fictitious fluid in the ether carrying energy and momentum that allowed him to save the momentum conservation law and also the center of mass theorem –and with this maneuver he got very close to a formulation of the energy-mass relation. However, Lorentz rejected Poincare's solution. The fact that a fictitious ether-fluid carries momentum means that the ether is in motion under certain conditions, and this would be at odds with the central assumption of an immobile ether in his theory:

> Lorentz's opinion of Poincare's valiant attempt at saving the principle of action and reaction was, in his words: "I must claim to you that it is impossible for me to modify the theory in such a way that the difficulty that you cited disappears." Lorentz went on to emphasize several times that his ether acted on bodies but that there was no reaction on the ether. He explained that the "phenomena of aberration," that is, fist order effects, had "forced him" to assume a motionless ether [...]. Lorentz continued in his letter: "I deny therefore the principle of reaction in these elementary actions." In mechanics, Lorentz continued, action and reaction were instantaneous because disturbances were not mediated by an ether; however, in electromagnetic theory the reaction of an emitter of radiation was not compensated simultaneously by the action on the absorber. Poincare had avoided this problem by attempting to satisfy the principle of reaction separately by emitter and absorber. Consistent with his desire to maintain an absolutely immobile ether, Lorentz protested Poincare naming the quantity in Eq. (5.10) [$G = \frac{1}{4\pi c}\int E \times B \, dV$; the momentum of the *fictitious fluid*] which Lorentz compared to Poynting's vector, to be an electromagnetic momentum. To Lorentz the term momentum, of course, connoted motion. Lorentz was willing to concede only that Poincare's electromagnetic momentum was formally "'equivalent' to a momentum". Thus, in the 1901 letter, Lorentz informed Poincare of his own sensitivity toward adding further hypothesis to an already overburdened theory, especially theses invented solely to save a principle whose violation permitted the theory's formulation in the first place.[101]

Accordingly, Lorentz's attitude towards the Trouton experiment was simply to hold one of the horns of the dilemma. His account of the experiment did consider an electromagnetic momentum, but interpreted as Abraham did in his 1903 *Principles of the Dynamics of the Electrons*. Abraham's electromagnetic momentum was carried by the electromagnetic field, and not by an ether-fluid, so that Lorentz's immobile ether was not threatened. Lorentz's explanation of the Trouton experiment was that the capacitor's loss of kinetic energy and momentum caused by the production of the magnetic field was compensated by the electromagnetic momentum carried by the magnetic field; so that the total momentum of the system as a whole remains constant. This explanation implies that the impulse in the capacitor does happen, and consequently, that the center of mass theorem does not hold. However, since Lorentz did not hesitate in abandoning Newton's third law in the name of all the achievements of his theory, this was not a problem for him:

> I take this opportunity for mentioning an experiment that has been made by Trouton at the suggestion of Fitzgerald, and in which it was tried to observe the existence of a sudden impulse acting on a condenser at the moment of charging or discharging; for this purpose the condenser was suspended by a torsion balance, with its plates parallel to the earth's motion. For forming

[101] Miller 1986, 5-7. Miller's article contains a full reproduction of the letter of Lorentz to Poincare, dated January 20th, 1901.

an estimate of the effect that may be experienced, it will suffice to consider a condenser with ether as dielectricum. Now if the apparatus is charged there will be an electromagnetic momentum $\mathfrak{C} = \frac{2U}{c^2} w$ [where U is the energy of the charged condenser at rest and w is its velocity with respect to the ether] (terms of the third and higher orders are here neglected). This momentum being produced at the moment of charging and disappearing at that of discharging, the condenser must experience in the first case an impulse $-\mathfrak{C}$ and at the second an impulse $+\mathfrak{C}$. However Trouton has not been able to observe these jerks.

I believe it may be shown (though his calculations have led him to a different conclusion) that the sensibility of the apparatus was far from sufficient for the object Trouton had in view.[102]

Besides Lorentz's abandonment of the action and reaction law –and of the theorem of the center of mass[103]– it must also be noticed that his account of the Trouton experiment would imply a violation of the principle of relativity as conceived by Poincare, and hence, a breakdown of the empirical equivalence with respect to SR. By 1904, Lorentz did not conceive his theory as an expression of the *full* invariance of the laws of physics under his coordinate transformations. In the same article in which he introduced his explanation of the Trouton experiment, he wrote that his goal was to show that: "by means of certain fundamental assumptions, and without neglecting terms of one order of magnitude or another, that *many* electromagnetic actions are entirely independent of the motion of the system"[104]. In other words, Lorentz did not think that any way to detect ether-wind effects would fail from the outset. Therefore, in order to make a case for the predictive equivalence of the theories at issue an account of the Trouton experiment that respects the principle of relativity must be offered in the context of Lorentz's theory.

This explanation can be carried out only by considering $E = mc^2$. If energy has mass, a transfer of energy from the battery to the capacitor is also a transfer of mass from the former to the latter; and in a frame in which they are both moving, a transfer of momentum is involved as well. Therefore, by charging the capacitor, it gains an amount energy, mass and momentum, while the battery loses the same amount of these quantities. Total momentum is then conserved. However, if energy has mass, the momentum circulation within the system does not imply a change in the velocity of its parts, for the increase of the capacitor momentum is given by a *mass*-gaining, not by a change in its *velocity*. On the other hand, and from the point of view of Fitzgerald's interpretation of the experiment –that the extra energy needed to produce the magnetic field came from the kinetic energy of the capacitor, not from the battery–, the kinetic energy lost by the capacitor and which is taken away by the magnetic field is, indeed, accompanied by a loss of momentum of the capacitor which is compensated by the electromagnetic momentum of the field –just as Lorentz said. Nevertheless, the loss of momentum of the capacitor means that it loses mass, not velocity –and the center of mass theorem still holds.

[102] From Lorenzt's "Weiterbildung der Maxwellschen Theorie Elektronentheorie", quoted in Janssen 2003, 37.

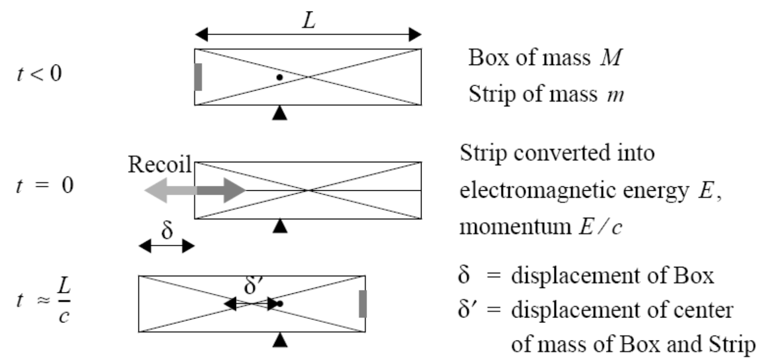
[103] The action and reaction law is violated in Lorentz's theory because it includes an ether that affects ponderable matter but which is never affected by it; see footnote 36 above. The center-of-mass theorem is violated because the jolt produced by the emission implies that the center of mass of the system gets accelerated, even though no external forces have been applied on the system; see footnotes 37 and 38 above.

[104] Quoted in Janssen 2003, 38. As I mentioned above, later on Lorentz –under the influence of Einstein, not of Poincare– ended up interpreting his theory as in full predictive equivalence with respect to SR. See Janssen 1995, section 3.5.

The connection between the Trouton experiment and the energy-mass relation becomes even clearer when one considers Einstein's second derivation of $E = mc^2$. In his 1906 *The Principle of the Conservation of Motion of the Center of Gravity and the Inertia of Energy*, Einstein established, by means of a thought experiment, that if the center of mass theorem is to hold in systems in which electromagnetic processes take place along with mechanical ones, it is necessary to consider the mass-energy relation. He considered a box of mass *M* and length *L*. In the left wall of the box an amount of energy E is stored. At time *t*=0 the energy is converted into electromagnetic radiation and travels to the other side of the box, where it is absorbed and converted to its original form and stored in the other wall. The emission of the electromagnetic radiation produces a recoil in the box in the opposite direction of the motion of the wave; and when it is received in the opposite wall another recoil of the same magnitude occurs, so that the box is brought back to rest. Electromagnetic theory says that the momentum of the radiation is E/c. Conservation of momentum requires that the recoil experienced by the box has that same amount momentum. The thought experiment shows that a closed system can move itself. If the energy involved has no mass –as classical physics says–, then we are facing a violation of the center of mass theorem.

The way out of the problem is simply $E = mc^2$. If the energy E stored in the left wall has a mass $m \ll M$, then, before its emission, the center of mass of the system formed by the box plus E lies slightly to the left of the middle of the box. When the emission of E occurs, a mass m travels to the opposite wall, and when it is received and stored, the center of mass of the system has moved to somewhere slightly to the right of the center of the box. If δ' is the displacement of the center of mass, its value can be calculated from the condition which determines where a wedge supporting the system should have to be placed in order to keep it perfectly balanced, namely, $M\left(\frac{\delta'}{2}\right) = m\left(\frac{L-\delta\prime}{2}\right)$. The value of the center of mass displacement is then $\delta' \approx \left(\frac{m}{M}\right)L$. The center of mass theorem is respected only if δ' is exactly compensated by δ, the displacement of the box in the opposite direction between its two recoils. The time during which the box displaces is given by L/c, and the velocity of the displacement is given by its momentum divided by its mass, that is, $\frac{E/c}{M}$. Hence, $\delta \approx \left(\frac{E/c^2}{M}\right)L$. Then, $\delta = \delta'$ if and only if $(E/c^2) = m$. In other words, the center of mass theorem holds only if $E = mc^2$. The Trouton experiment is clearly an empirical instance of Einstein's thought experiment. Actually, in the former, a tiny recoil occurs when the capacitor is charged, but it is compensated by a displacement of the center of mass, and the theorem is respected. The following figure is a good depiction of these remarks:[105]

[105] The figure is taken from Janssen 2003, 40. A full and precise explanation of the Trouton experiment in the context of Lorentz's theory that includes the energy-mass relation must consider some further subtleties. The definition of electromagnetic momentum involved is not Lorentz invariant, insofar as it includes a special reference to the ether-rest frame. The *hyperplane of simultaneity* that is the base for the integration over space that defines the electromagnetic momentum is the one that corresponds to the ether-rest frame. Accordingly, the explanation must also consider *Laue's effect*, namely, that stresses in the capacitor considered at rest give rise to momentum in a frame in which it moves. This momentum must be added to the electromagnetic one in order to explain Trouton's experiment. See Janssen 2003, 44-9.



These statements clearly show that if Lorentz's theory is to be proved predictively equivalent to SR, then the energy-mass relation has to be considered. Accordingly, it must also be shown that this relation can be derived from Lorentz's theory. Otherwise, the implications of $E = mc^2$ could be used as a crucial experiment. Herbert Ives has shown that the famous equation is indeed contained in Lorentz's theory, and that Poincare was very close to formulate it. Actually, in his 1906 paper, Einstein himself acknowledges that his derivation is quite similar to Poincare's result in his 1900 criticism of Lorentz violation of Newton's third law. Ives offers an analysis of Poincare's introduction of his fictitious fluid that underscores the mentioned similarity. Departing from the fact that the French scientist derived the expression $M = S/c^2$ for the fictitious fluid, where M is the momentum of the radiation, and S is the *flux* of radiation; Ives points out:

> Consider how Poincare got his numerical result. He was using his formula, derived in his article, for the momentum radiation, $M = S/c^2$, and he was putting down the expression for the conservation of momentum in the recoil process. Putting μ for the mass of the recoiling body, and v for its velocity, his working equation is then $\mu v = S/c^2$. For S, the energy flux, he put the energy E times c. He then has $\mu v = S/c^2 = Ec/c^2 = E/c^2 \cdot c$.
> The significant thing for our present study is that Poincare in his calculation used E/c^2 for the coefficient of c in stating the momentum of the radiation, that is E/c^2 plays the role of mass. The relation $E = m_R c^2$ was thus contained in his relation $M = S/c^2$.[106]

What Ives means by m_R is the mass equivalent for free radiation. Since in his 1900 paper Poincare, as we saw above, established that the fictitious fluid was not indestructible in the sense that it could not be entirely transferred in the emission or absorption of energy, and hence it always had to appear as energy in other guises; then Poincare precluded the possibility of interpreting m_R also as m_M. By m_M Ives designates the mass of matter, in opposition to the mass equivalent of free radiation. In other words, Poincare was not in a position to interpret the expression $E = m_R c^2$ underlying his formula for the fictitious fluid momentum –even if he had actually seen that it was there– as an expression referred to the gain or loss of mass that matter experiences when it emits radiation.

However, Ives also points out that, had Poincare tackled the issue from the point of view of *his* relativity principle, and even within the context of an ether-theory ontology, he had all the tools needed to obtain the mass-energy relation as referring both to m_R and to m_M:

106 Ives 1952, 540.

Consider a body suspended loosely, as by a nonconducting chord, in the interior of an enclosure, the whole system being stationary with respect to the radiation transmitting medium [the ether]. Let the body emit symmetrically in the 'fore' and 'aft' directions the amount of energy $\frac{1}{2}E$. The momenta of the two oppositely directed pulses cancel each other, the body does not move, and no information can be obtained as to its change of state.
Now let the whole system of enclosure and suspended particle be set in uniform motion with respect to the radiation transmitting medium with the velocity v. the body now possesses the momentum $mv[1-(v^2/c^2)]^{1/2}$ and the problem is to determine the effect on this momentum of the two emitted wave trains. Now the energy contents of the two wave trains emitted for the same (measured) period of emission, taking into account the change of frequency of the source and the lengths of the trains, are

$$\frac{E}{2}\frac{[1+(v/c)]}{[1-(v^2/c^2)]^{1/2}} \quad \text{and} \quad \frac{E}{2}\frac{[1-(v/c)]}{[1-(v^2/c^2)]^{1/2}}.$$

The accompanying momenta, from Poincare's formula, are

$$\frac{E}{2c^2}\frac{[1+(v/c)]}{[1-(v^2/c^2)]^{1/2}}c \quad \text{and} \quad \frac{E}{2c^2}\frac{[1+(v/c)]}{[1-(v^2/c^2)]^{1/2}}c.$$

These being oppositely directed, the net imparted momentum is $Ev/c^2[1-(v^2/c^2)]^{1/2}$. Forming the equation for the conservation of momentum we have

$$\frac{mv}{[1-(v^2/c^2)]^{1/2}} = \frac{m\prime v\prime}{[1-(v^2/c^2)]^{1/2}} + \frac{Ev}{c^2[1-(v^2/c^2)]^{1/2}},$$

where v' is the velocity of the body after the emission of the radiation.
Now according to Poincare's principle of relativity, the body must behave in the moving system just as in the stationary system first considered, that is, it does not change its position or velocity with respect to the enclosure, hence $v' = v$, and we get

$$\frac{(m-m')v}{[1-(v^2/c^2)]^{1/2}} = \frac{Ev}{c^2[1-(v^2/c^2)]^{1/2}},$$

giving exactly $(m-m') = E/c^2$, a relation independent of v, and so holding for the stationary system. The radiating body losses mass E/c^2 when radiating mass E. This is the relation $E = m_M c^2$.[107]

In simple words, what Ives' analysis shows is that –just as Einstein himself acknowledged– in the context of Lorentz's theory it is possible to carry out a derivation of $E = mc^2$ that is analogous to the one that Einstein performed in 1906 –with the required provisos, namely, that electromagnetic momentum and the ether must be considered. Moreover, a much more simple derivation is available once Poincare's stress in considered and plugged into Lorentz's expression for electromagnetic momentum, as I showed above.

I think that these remarks are enough in order to see that the mass-energy relation derived by Einstein in the context of SR is mathematically and physically contained in Lorentz's theory. It is true, though, that from a historical point of view the famous equation was not directly discovered by Lorentz or Poincare, but it had to be *borrowed* from Einstein's results. However, in the context of this research, the fact that $E = mc^2$ was *conceptually* contained in Lorentz's theory is enough to make a complete case for the predictive equivalence of the theories at issue. On the other hand, the fact that many of the

---

107 Ives 1952, 541. One could object that this derivation rests upon the momentum $M = S/c^2$ carried by the fictitious fluid; and we saw the reasons that Lorentz gave for his rejection of this concept. However, the generality of Ives' analysis allows that the electromagnetic momentum *as understood by Abraham* could be used as well, i.e., as the momentum carried by the *electromagnetic field* –which added to the momentum of the recoil of the emitter gives the total momentum that is conserved.

amendments, extensions and interpretations that Poincare introduced are crucial to argue for the empirical equivalence, it seems more correct to state that the theories that are really predictively identical are SR and something like a *Lorentz-Poincare theory*, rather than Lorentz's. It must be remarked that what we can dub the Lorentz-Poincare theory (LPT) is a *conceptual* reconstruction which is only possible with the benefit of hindsight. This tag I propose is not meant to refer to a theory which existed from a *historical* standpoint. No textbook about it was ever written.

Before turning to a treatment of the differences between the theories which ground their rivalry, I will briefly mention one argument that has been put forward in order to deny their empirical equivalence. Arthur Miller, for example, claims that Lorentz's theory cannot yield the relativistic Doppler effect[108]. However, as Janssen points out, if Poincare's contributions are considered, that is, if the factor l is set to unity and if the transformations are understood as symmetric; then the relativistic expression for the Doppler effect does follow from the ether theory[109]. This remark is yet another reason to consider the *LPT* as empirically equivalent with respect to SR.

b) Different and rivals

The first difference between Einstein's SR and the ether theory of Lorentz and Poincare that I will address was already mentioned above. In this section I will simply state it in more precise terms. In a newspaper article he wrote in 1919, Einstein introduced a distinction between two kinds of scientific theories, namely, between *constructive* theories and theories *of principle*:

> We can distinguish between various kinds of theories in physics. Most of them are constructive. They attempt to build up a picture of the more complex phenomena out of the materials of a relatively simple formal scheme from which they start out. Thus the kinetic theory of gases seeks to reduce mechanical, thermal, and diffusional processes to movements of molecules, i. e., to build them up out of the hypothesis of molecular motion. When we say that they have succeeded in understanding a group of natural processes, we invariably mean that a constructive theory has been found which covers the processes in question.
>
> Along with this most important class of theories there exists a second, which I will call 'principle-theories'. These employ the analytic, not the synthetic, method. The elements which form their basis and starting point are not hypothetically constructed but empirically discovered ones, general characteristics of natural processes, principles that give rise to mathematically formulated criteria which the separate processes or their theoretical representations of them have to satisfy. Thus the science of thermodynamics seeks by analytical means to deduce necessary conditions, which separate events have to satisfy, from the universally experienced fact that perpetual motion is impossible.
>
> The advantages of the constructive theories are completeness, adaptability, and clearness, those of the principle theory are logical perfection and security of the foundations.
>
> The theory of relativity belongs to the latter class. In order to grasp its nature, one needs first of all to become acquainted with the principles on which it is based.[110]

[108] Miller 1986, 232.
[109] See Janssen 1995, section 3.3.5. Janssen also mentions that Miller acknowledged this point.
[110] From Einstein's *My Theory*, quoted in Dieks 2009, 2.

From what has been said so far it is quite clear that the *principles* of SR are the relativity postulate and the constancy of light postulate. Just as Einstein's definition of a theory of principle states, these principles do not presuppose any conceptions about the ultimate nature of the processes they refer to; rather, they are general features –empirically suggested– that work as constraints the physical processes must satisfy. On the other hand, it is also clear that the LPT is a constructive theory, for some particular electrodynamical descriptions of matter and physical processes are the features that determine the picture of the world the theory provides.

Even though this is a very important difference between the theories, it is not sufficient to establish their rivalry. It is just a formal or schematic dissimilarity which does not determine that the LPT and SR are contending opponents. For instance –and using Einstein's own example–, though thermodynamics and the kinetic theory of gases are different in the sense at issue, and even though they offer an account of the same phenomena, they are not rivals, but complementary.

Actually, the early reception of SR and of Lorentz's theory was such that the principle-constructive difference between them was somewhat noticed; but most of the scientific community understood that the generalization that Einstein had introduced with respect to the achievements of Lorentz did not imply that they were contenders. This fact is quite apparent when one looks at the way in which the Kaufmann experiments were interpreted. In 1901 Walter Kaufmann started a series of experiments with β-radiation –electrons produced in radioactive decay and emitted with a velocity close to c– set out to test the v-dependence of the inertial mass of the β-particles. These experiments were considered as a way to empirically decide between three alternative theories: Abraham's theory, Bucherer and Langevin's, and the *Lorentz-Einstein* theory. That is, scientists, until around 1911, considered that the difference between the theories at issue was quite similar to the difference between thermodynamics and the kinetic theory of gases. This curious feature becomes understandable if one considers that since the reception of the theories occurred in the context of the Kauffmann's experiments, the scientific community paid more attention to the electrodynamical part of Einstein's paper than to its kinematical part[111].

The radically new approach to physics that Einstein's paper contained with respect to the notions of space and time were crucially developed and clarified by the work of Hermann Minkowski. In his famous paper of 1908 *Space and Time* he elucidated that Einstein's theory implied a revolutionary reformation of the meaning of these concepts by the introduction of a four-dimensional manifold that we now call *space-time*.

Minkowski's paper departs from a critical consideration of the way in which the geometrical and kinematical transformations of coordinates valid for Newtonian mechanics were usually regarded:

> The equations of Newton's mechanics exhibit a two-fold invariance. Their form remains unaltered, firstly, if we subject the underlying system of spatial-coordinates to any arbitrary *change of position*; secondly, if we change its state of motion, namely, by imparting to it any

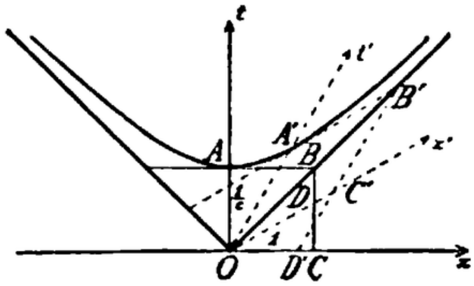
[111] See Miller 1998 [1981], section 7.4.

uniform translator motion. […] Each of them by itself signifies, for the differential equations of mechanics, a certain group of transformations. The existence of the first group is looked upon as a fundamental characteristic of space. The second group is preferably treated with disdain, so that. […] Thus the two groups, side by side, lead their entirely apart. Their utterly heterogeneous character may have discouraged any attempt to compound them. But it is precisely when they are compounded that the complete group, as a whole, gives us to think.[112]

Minkowski's contribution was precisely to compound the transformation groups from the point of view of Einstein's theory. In simple words, he undertook a geometrical approach to kinematics. To achieve his goal, Minkowski proposed us to consider a point in space and in time –a *world-point*– in terms of a set of *Cartesian-like* coordinates x, y, z, t; so that "the multiplicity of all thinkable x, y, z, t systems of values we will christen the *world*"[113]. Minkowski's world is then a *four-dimensional space-time*. Then he pays attention to the path of one particular world-point. The differential of its coordinates that define that path determine his curve in the world, its *world-line*.

After these preliminary considerations, he introduces his crucial idea: "The whole universe is seen to resolve itself into similar world-lines, and I would fain anticipate myself by saying that in my opinion physical laws might find their most perfect expression as reciprocal relations between these world-lines"[114]. In order to materialize the idea he anticipates, he tells us to consider the positive parameter c and the graphical representation of the hyperbola $c^2t^2 - x^2 - y^2 - z^2 = 1$ –or more precisely, he considers the sheet of the hyperbola in the region $t > 0$. He then considers "those homogeneous linear transformations of x, y, z, t into four new variables x', y', z', t', for which the expression for this sheet in the new variables is of the same form". The rotational and translational transformations, just as in Newtonian mechanics, are the ones that entail that $x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2$ –provided that *t=0* determines the same instant in both systems. In the kinematic case, the transformations looked for –presupposing a standard configuration of the systems– are the ones that entail that $c^2t^2 - x^2 = c^2t'^2 - x'^2$. Minkowski concludes that the group of the rotational and translational transformations plus the kinematical transformations –the group G_c– for which the expression of the hyperbola retains its form depends on the parameter c. In order to illustrate his line of thought, Minkowski depicted this famous diagram:[115]

[112] Lorentz *et al.* 1952, 75-6.
[113] *Ibid.*
[114] *Ibid.*
[115] "We consider the sheet in the region $t > 0$, and now take those homogeneous linear transformations of x, y, z, t into four new variables x', y', z', t', for which the expression for this sheet takes the same form. It is evident that the rotations of space about the origin pertain to these transformations. Thus we gain full comprehension of the rest of the transformations simply by taking into consideration one among them, such that y and z remain unchanged. We draw the section of this sheet by the plane of the axes of x and t –the upper branch of the hyperbola $c^2t^2 - x^2 = 1$, with its asymptotes. From the origin O we draw any radius vector OA' of this branch of the hyperbola; draw the tangent of the hyperbola at A' to cut the asymptote on the right at B'; complete the parallelogram $OA'B'C'$ […]. Now if we take OC' and OA' as axes of oblique co-ordinates x', t', with the measures OC'=1, OA'$=1/c$, then that branch of the hyperbola again acquires the expression $c^2t'^2 - x'^2 = 1$, $t' > 0$, and the transition from x, y, z, t to x', y', z', t' is one of the transformations in question. With these transformations we now associate the arbitrary displacements of the zero point of space and time, and thereby constitute a group of transformations, which is also, evidently, dependent on the parameter c. this group I denote by G_c". Lorentz *et al.* 1952, 77-8.



These kinematical-geometrical considerations acquire their *Einsteinean* physical meaning simply by identifying c with the velocity of the propagation of light. In that case, the kinematical transformations of G_c are, of course, the Lorentz transformations. Minkowski's own evaluation of the deep physical significance implied by his geometrical contribution was the following:

> The existence of the invariance of natural laws for the relevant group G_c would have to be taken, then, in this way:
> From the totality of natural phenomena it is possible, by successively enhanced approximations, to derive more and more exactly a system of reference x, y, z, t, by means of which these phenomena then represent themselves in agreement with definite laws. But when this is done, this system of reference is by no means unequivocally determined by the phenomena. *It is still possible to make any change in the system of reference that is in conformity with the transformations of the group G_c, and leave the expression of the laws of nature unaltered.*[116]

The revolutionary conception of space and time implied by SR that Minkowski's contribution clarifies becomes quite clear by comparing it with the Newtonian framework. In the latter, the *place* in which two successive events occur is relative to a specific inertial framework, but the *instants* in which those events occur is an absolute feature. In Einstein's theory the Lorentz transformations entail that the time at which successive events occur is also frame-relative. Minkowski's space-time contributes to make clear that this relativity is inherently associated to the geometric-kinematic properties of a four-dimensional space-time whose invariants are determined by G_c.

A second important difference lies on the *invariant* interval between events of Minkowski's space-time. Whereas in the Newtonian world the geometrical and kinematic transformations leave intact the distance $\Delta x^2 + \Delta y^2 + \Delta z^2$ and leave intact the time interval Δt between to events; Minkowski shows that in a four-dimensional space-time governed by G_c the invariant interval between events is given by $c\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$, which is normally denoted by $\Delta \boldsymbol{s}^2$. Moreover, the Newtonian distance between E_1 and E_2 is greater than 0, or equal to 0 only if E_1=E_2; but the Minkowskian 'distance' can be greater, equal or less than 0. If the $\Delta \boldsymbol{s}^2$ between two events is greater than 0, the 'distance' between them is called *time-like*, if it is equal to 0 it is called *null*; and if it is less than 0 the interval is *space-like*. Therefore, and with respect to a given event *E,* Minkowskian space-time gets divided in three regions: the set of events whose interval is time-like; the set of events whose interval is null, and the set of events whose intervals is space-like. The set of events whose 'distance'

[116] Lorentz *et al.* 1952, 79.

from E is null form E's light cone, the time-like events with respect to E are situated within its cone of light, and the space-like ones are outside the cone. This arrangement of events in turn reflects the causal structure of space-time. Only the events in the surface or inside the cone of light of E can be causally connected with it; in the former case the only possible connection is given by a light ray between them, in the later there is always the possibility of an observer passing between two time-like events. Events outside the cone of E cannot be causally connected with it.

This classification can also be applied to the 'lengths' of curves connecting events and to define the Minkowskian geodesics or straight lines. If a curve C that connects two events in space-time is such that its length $\int_C d\boldsymbol{s}$ is larger than the length of any other curve connecting the same two events, then C is a geodesic –if it is a time-like, null or space-like geodesic depends on the value of the corresponding $\Delta\boldsymbol{s}^2$, of course. In turn, geodesics in space-time allow a geometric account of inertia. Consider the world-line of a particle p. If p is a freely moving particle, then its inertial motion is represented by a time-like geodesic; and if the word-line of p is *not* a time-like geodesic, then its motion is accelerated.

After this brief revision of the main tenets of Minkowski's space-time, it is possible to assert that, in Einstein's theory, the Lorentz transformations express the structure and the metric properties of a Minkowskian space-time. This is a simple way to understand the main difference which underlies its rivalry with respect to the LPT. It is true that the LPT can be expressed in terms of a 'Minkowskian description' –recall that Poincare discovered the invariant interval and other four-dimensional features–, but that description only express some *mathematical niceties* of the theory. The LPT does *not* describe the world as characterized by a Minkowskian structure and metric. The metric and geometry of the 'Lorentzian space-time', one might say, is still Newtonian[117]. The Lorentz transformations do not express geometric and kinematical features of that world, but the *dynamical* features which govern the behavior of objects inhabiting a Newtonian space-time. SR and the LPT are not rivals because the former offers a top-down explanation and the latter a bottom-up one. The rivalry is grounded in the fact that the constructive approach of the LPT lies upon a conception of the nature of space-time that is inconsistent with the corresponding conception contained in the principles of SR.

This way to characterize the difference and rivalry between the theories at issue allows a straightforward rejection of an interesting argument proposed by Lászlo Szabó that aims to show that SR and the LPT are the very same theory. He claims that

> According to [the] widespread view, special relativity was, first of all a new theory about space and time. A theory about space and time describes a *certain group of objective features* of physical reality, which we call (the structure of) space-time. Consider claims like these:

[117] Minkowski's four-dimensional approach allows to construct a four-dimensional *Newtonian* space-time: "If we now allow c increase to infinity, and $1/c$ therefore to converge towards zero, we see form the figure that the branch of the hyperbola bends more and more towards the axis of x, the angle of the asymptotes becomes more and more obtuse, and that in the limit this special transformation changes into one in which the axis of t' may have any upward direction whatever, while x' approaches more and more exactly to x. In view of this is clear that group G_c when $c = \infty$, that is the group G_∞, becomes no other than the group which is appropriate to Newtonian mechanics". Lorentz *et al.*, 78-9.

- According to classical physics, the geometry of space-time is $\mathbb{E}^3 \times \mathbb{E}^1$, where $\mathbb{E}^3$ is a three-dimensional Euclidean space for space and $\mathbb{E}^1$ is a one-dimensional Euclidean space for time, with two independent invariant metrics corresponding to the space and time intervals.
- In contrast, SR claims that the geometry of space-time is different: it is a Minkowski geometry $\mathbb{M}^4$.

> The two statements are usually understood as telling *different things about the same* objective features of physical reality. One can express this change by the following logical schema: Earlier we believed in $G_1(\widehat{M})$, where $\widehat{M}$ stands for (the objective features of physical reality called) space-time and G_1 denotes some predicate (like "of type $\mathbb{E}^3 \times \mathbb{E}^1$"). Then we discovered that $\neg G_1(\widehat{M})$ but $G_2(\widetilde{M})$, where G_2 denotes a predicate different from G_1 (something like "of type $\mathbb{M}^4$").
> *This is however not the case*. Our analysis will show that the correct logical schema is this: Earlier we believed in $G_1(\widehat{M})$. Then we discovered for some *other* features of physical reality $\widehat{M} \neq \widetilde{M}$ that $\neg G_1(\widetilde{M})$ but $G_2(\widetilde{M})$. Consequently, it still may (and it actually does) hold that $G_1(\widehat{M})$. In other words, in comparison with the pre-relativistic Galileo-invariant conceptions, special relativity tells us nothing new about the geometry of space-time. It simply calls something else "space-time", and this something else has different properties. We will also show that all statements of special relativity about those features of reality that correspond to the original meaning of the terms "space" and "time" are identical with the corresponding traditional pre-relativistic statements. Thus the only new factor in the special relativistic account of space-time is the terminological decision to designate something else "space-time".
> So the real novelty in special relativity is some $G_2(\widetilde{M})$. It will be also argued, however, that $G_2(\widetilde{M})$ does not contradict to what Lorentz claims. Both, the Lorentz theory and special relativity claim that $G_1(\widehat{M})\&G_2(\widetilde{M})$. In other words: *SR and the Lorentz theory are identical theories about space and time in all sense of the words.*[118]

It might be true that $G_1(\widehat{M})$ holds both in Einstein's and in Lorentz's theories. But according to what I have said above, I do not agree with the second part of Szabó's argument. It is not true that both theories assert $G_2(\widetilde{M})$. If by this statement we understand something like 'the physical world is characterized by the properties of the Minkowskian space-time', I just explained that the LPT does not claim that. When Szabó claims that "then we discovered for some *other* features of physical reality $\widehat{M} \neq \widetilde{M}$ that $\neg G_1(\widetilde{M})$ but $G_2(\widetilde{M})$", I think that the right interpretation of this view is that the LPT claimed that $G_1(\widetilde{M})$ –understanding this last statement as something like 'the physical world is a *Newtonian* space-time in which some dynamical features expressed by the Lorentz transformations govern the behavior of objects[119].

---

[118] Szabó 2011, 1-2.

[119] A possible reply could be that if $\widetilde{M}$ is defined in terms of *measurements results*, then it could be said that both the theories assert that $G_2(\widetilde{M})$. However, at least in SR $\widetilde{M}$ is not defined in that way. Moreover, this maneuver would be quite close to a logical-positivistic verificationist criterion of meaning, and my criticism to Szabó's argument can be taken as yet another example to show that, contrary to the claim of logical positivists, the meaning of a scientific theory does not reduce to its empirical consequences; these are only *a part* of that meaning. Actually, a logical positivistic-like semantic criterion seems to underlie Szabó's argument: the title of his paper is *Lorenztian theories vs. Einsteinian SR –a logico-empiricist reconstruction*. Moreover, when he refers to his methodological principles, he states that "it is to be noted that our analysis is based on the following very weak operationalist/verificationist premise: physical terms, assigned to *measurable physical quantities*, have different meaning if they have different empirical definitions"; 20.

V. On the reasons to choose

Now we are in position to deal with the epistemological problem involved. We have two physical theories which have the same empirical consequences. However, they are inconsistent. Then, and under a thorough *hypothetical-deductive* conception of the confirmation of theories, the choice to be done between them is deeply underdetermined regarding empirical evidence. In this section I will propose an analysis of the possible reasons to make a choice, along with an evaluation concerning which of them are good, and which of them are bad.

1. Bad reasons

a) The Lorentz-Fitzgerald contraction and ad-hocness

The competition between SR and Lorentz's (or the LPT) theory has been the object of philosophical debate for many years –and in many cases the debate goes further than empirical equivalence issues. One of the very first subjects in the philosophy of science which was associated to it was the concept of an *ad-hoc* hypothesis. For example, Popper took the Lorentz-Fitzgerald contraction as a paradigmatic case of an *ad-hoc* hypothesis, and he stated that its presence within Lorentz's theory was enough to prefer Einstein's, for the latter did not contain any such hypothesis. Popper's view was very influential in the treatment of the Einstein *vs.* Lorentz-Poincare case, so I will first consider and evaluate the accusation of *ad-hoc*ness features as a possible reason to make a decision in the competition.

The concept of an *ad-hoc* hypothesis is a whole subject on its own in the philosophy of science, and its very definition is not at all clear and definitive –for example, Hempel has argued that a purely logical criterion for *ad-hoc*ness simply does not exist, only pragmatic ones are viable[120]. This complication gets reflected in the treatment of the specific case of the Lorentz-Fitzgerald hypothesis: all of the authors who refer to it introduce their own definitions of *ad-hoc*ness[121]. However, it is possible to extract two common factors from the diverse accusations and defenses of the contraction hypothesis as being *ad-hoc*: issues concerning the testability of the hypothesis, and issues concerning its plausibility.

It is quite clear that, both from a conceptual and historical point of view, Lorentz's –and also Fitzgerald's– motivation for the introduction of the contraction hypothesis was the null outcome of the Michelson-Morley experiment. In that sense, it was *cooked up* with the goal of explaining one single empirical problem in their theories. The problem with this maneuver is that it is sometimes affirmed that once the contraction hypothesis was introduced, its *only* contribution was the explanation of the mentioned experiment –so that it cannot receive any further empirical confirmation. This also means that the hypothesis is somewhat *conceptually isolated* within the theory, for it is not relevant to derive any further

[120] Hempel 1966, 30.
[121] Some relevant papers on this issue are Grünbaum 1959, Zahar 1973, Schaffner 1974, Miller 1974, Leplin 1975, Grünbaum 1976, Janssen 2002.

empirical consequences other than the negative outcome of the Michelson-Morley experiment.

These problematic features of the hypothesis were indeed the case in what I called *the first stage* of Lorentz's theory. The first version of the theorem of corresponding states allowed an explanation of the negative result of only *first-order* experiments. The contraction hypothesis was explicitly and independently introduced to explain the famous *second-order* experiment. In this explanation, the first version of the theorem did not play any role, and, in turn, the contraction hypothesis did not play any part in the explanation of the first-order experiments. Moreover, Popper's accusation also states that the hypothesis entails a negative result for all the optical Michelson-Morley-like experiments, and since it did not entail any further consequences other than these, the hypothesis is un-falsifiable and *pseudo*-scientific. This last claim is not totally true. Adolf Grünbaum showed that the hypothesis, when applied to the Kennedy-Thorndike experiment –a variation of Michelson-Morley's in which one of the arms of the interferometer is shorter than the other–, predicts a positive outcome, and therefore, a test that could falsify it. Popper accepted Grünbaum's rebuttal, but given the modest source of falsification provided by the Kennedy-Thorndike experiment he concluded that Lorentz's theory was *more ad-hoc* than Einstein's[122]. From the point of view of the first stage of the theory, Popper is right, and the different degrees of *ad-hoc*ness could then be invoked to decide between the theories at issue.

However, and even though the original motivation for the introduction of the hypothesis gets fixed once and for all, in the *second stage* of the development of Lorentz's theory the accusation of *ad-hoc*ness does not hold –whether in terms of problems about falsifiability or in terms of lack of independent confirmation for the hypothesis. I showed above that in the final version of the theorem of corresponding states –interpreted in connection with the generalized contraction hypothesis– the Lorentz-Fitzgerald contraction was a part of the core of the theory, for it was a specific consequence of the general hypothesis and the theorem. Moreover, I also showed how the velocity dependence of inertial mass follows in Lorentz's works of 1899 and 1904; and in both cases it is a consequence of the theorem and the generalized hypothesis. Therefore, the issue can be put this way: after 1904, Lorentz's theory provided an explanation of the null result of optic experiments to any order of v/c in which the theorem of corresponding states entailed the Lorentz-Fitzgerald contraction and the local time effect, which taken together imply the negative result. Moreover, the Lorentz-Fitzgerald contraction –as a consequence of the theorem and the generalized contraction hypothesis– was also involved in the most spectacular prediction of the theory: the velocity dependence of mass. These two last remarks show that in a more developed version of the theory, an acussation of *ad-hoc*ness does not hold. And since it is the LPT the one which is equivalent to SR, not Lorentz's theory of 1895, this specific argument of *ad-hoc*ness is irrelevant for the decision to be made[123].

[122] See Grünbaum 1959, 48-50.
[123] In this point I follow Janssen 2002, 431-6.

The second feature on which most of the accusations of *ad-hoc*ness are grounded is an alleged problem of plausibility. In simple words, the Lorentz-Fitzgerald contraction is sometimes understood as a sort of *rabbit in the hat* maneuver. The lack of theoretical rationale to make it plausible thus determines its being *ad-hoc*. This accusation does not work either. As is showed above, since its very first introduction, both Lorentz and Fitzgerald proposed similar plausibility arguments for the contraction. It is true that such arguments were not *proofs* of the reality and of the specific nature of the molecular forces causing the contraction, but they were enough as to consider the contraction a *physically grounded* hypothesis. Moreover, in the definitive version of the theory, the theorem of corresponding states and the generalized contraction hypothesis are the basis for Lorentz's assumption that *all* forces whatsoever transform like electromagnetic ones. In turn, this assumption is quite natural within the context of the electromagnetic worldview. If mechanical features are rooted in electromagnetic grounds, it is rather expectable that all forces behave like electromagnetic forces do; for, in the end, they are *all* electromagnetic. Thus, one could say that –in its 1904 version– the molecular forces that cause the contraction are derived from the core of the theory.

The accusation of *ad-hoc* features in Lorentz's becomes more serious if the burden of plausibility is transferred to the hypothesis that all forces transform like the electromagnetic ones do. I just said that in the context of a thorough electromagnetic worldview it is a rather natural assumption. However, as I showed above, the introduction of the Poincare-stress –which was crucial for the consistency of Lorentz's theory– involves a problem for that assumption. We have to recall that the model of the electron proposed by the Dutch physicist involved a contraction not only of the distance between the molecules of a body –which can be plausibly understood under the assumption about the behavior of forces–, but also of the *electron itself*. Poincare showed that the electron can be made stable only by means of the introduction of a *non-electromagnetic* force which ended up providing the rationale of the electron's contraction: "as the region where $P_{Poincare}(x)$ is non-vanishing always coincides with the ellipsoid-shaped region occupied by the moving electron, *these forces make the electron contract by a factor γ in the direction of motion*"[124]. As Kenneth Schaffner points out, the fact that the Poincare-pressure is definitively non-electromagnetic makes it hard to give an account of why it is connected to a contraction of the electron which has that specific value:

> There were however serious reservations about the satisfactory applicability of *M.F.H.* [molecular forces hypothesis] to electrons. First, it was shown by Poincare (1906) that the contractile electron could be considered a stable entity only if a definitively non-electromagnetic counter-pressure were invoked. To extend the *M.F.H.* to cover *this* type of force would violate the reduction thesis (of the M.F.H to electromagnetic forces) which provided the plausibility (and independent support) for the original M.F.H.[125]

I think this is a much more serious accusation of *ad-hoc*ness. The plausibility argument provided by Lorentz does not include the Poincare-pressure in its scope. However,

[124] Janssen & Mecklenburg 2007, 31; emphasis added.
[125] Schaffner 1974, 52.

Poincare did offer a sort of plausibility argument for the non-electromagnetic stress he introduced –an argument that Schaffner does not refer to in his paper. After the mathematical formulation of his amendment to Lorentz's theory, Poincare states that

> *the pressure due to our supplementary potential is proportional to the fourth power of the experimental mass of the electron.*
> Since the Newtonian attraction is proportional to the experimental mass, one is tempted to infer that there exists a general relation between the causes giving rise to gravitation and those which give rise to the supplementary potential.[126]

I think it is true that this plausibility argument is not as good as the one offered by Lorentz regarding the molecular forces. However, it still *is* a plausibility argument. Moreover, in the context of Poincare's goal –the construction of a theory within the spirit of the electromagnetic worldview–, it is an argument that fits into that program. The last sections of Poincare's work of 1906 were dedicated to an attempt of reducing gravitation theory to electromagnetism. So, if the stress introduced could be shown to be gravitationally determined, it was not madness to think that it was therefore an electromagnetic feature in the end –and that would be shown by a successful electromagnetic account of gravitation. More generally, I think that the quoted passage shows that Poincare was clear about the need of a justification for the force he introduced. He offered a tentative argument which was more a *promise* of an explanation than an actual explanation. Such a promise, nevertheless, was one which could perhaps be kept –at least before the problems that quantum-related phenomena put in the face of electromagnetic theory began to occupy a central place in physics. Schaffner argues that when the incapability of electromagnetic theory to account for quantum-related phenomena became more and more clear, the plausibility arguments for the molecular force and the Poincare stress fell down; and consequently, they became *ad-hoc* hypotheses[127]. But I think that if this is the case, the problems with the LPT were much more general and deeper than the inclusion of *ad-hoc* hypotheses, so it would be trivial to invoke *ad-hoc*ness as a relevant reason to reject it. I will return to this issue below.

From a more general point of view, one could ask what is wrong with *ad-hoc* hypotheses *in themselves*. Even if the hypotheses considered here could be characterized as *ad-hoc* in some of the many senses of the term, one could still argue that all the theoretical and empirical achievements they provide are enough for not considering them as a harmful feature of the theory. I agree with Grünbaum and Laudan when they claim that it is dubious that if a certain hypothesis can in some sense be considered as *ad-hoc* then that hypothesis necessarily carries unscientific consequences[128]. I think that since *i)* the accused hypotheses

---

[126] From Poincare's *On the Dynamics of the Electron*; quoted in Miller 1986, 120.
[127] Schaffner 1974, 73-6.
[128] "Assuming that adhocness is understood in this way [a hypothesis which is *cooked up* with the specific and only goal of solving an empirical anomaly], we are entitled to ask: *what is objectionable about it?* if some theory T_2 has solved more empirical problems than its predecessor –*even just one more*– then T_2 is clearly preferable to T_1, and, *ceteris paribus*, represents cognitive progress with respect to T_1. [...]
In urging that adhocness (so defined) is a cognitive virtue rather than a vice, I am clearly not implying that ad hoc theories are invariably better than non-ad hoc ones. My claim, rather, is that an ad hoc theory is preferable to its non-ad hoc predecessor (which was confronted with known anomalies).
But it might be argued that I have missed the point of the critics of *adhocness*. They might say 'yes, of course, T_2 is better than its *refuted* predecessor T_1; but the relevant comparison is between T_2 and some other theory T_n which is

were conceptually connected to the rest of the theory in its definitive form, and *ii)* there were plausibility grounds to regard those hypotheses as scientifically reasonable ones; then any sense of *ad-hoc*ness that could still be assigned to them is not an immediate argument to state that the theory they are a part of is flawed one[129].

b) Mathematic-aesthetic features

A second reason that can be invoked in order to make a decision in the case we are dealing with is based upon *aesthetic* considerations. Actually, from a historical point of view, the fact that the members of the scientific community judged that SR was a mathematically simpler and more elegant theory than Lorentz's played an important role in turning the balance in Einstein's favor[130]. Consider, for instance, the following passage by Max von Laue in his textbook on SR (the first ever published):

> Though a true experimental decision between the theory of Lorentz and the theory of relativity is indeed not to be gained, and that the former, in spite of this, has receded into the background, is chiefly due to the fact that, close as it comes to the theory of relativity, it still lacks the great simple universal principle, the possession of which lends the theory of relativity an imposing appearance.[131]

The comparative simplicity of Einstein's theory can be easily noted in that it needed only two basic postulates, along with a deep reform of the concept of simultaneity, in order to achieve all what Lorentz's theory did only after a long and winding road –paraphrasing The Beatles– and by means of introducing further and further amendments and hypotheses.

In addition to the mathematical and structural simplicity of Einstein's theory, its formulation in terms of the four-dimensional language introduced by Minkowski was followed by an evaluation of SR as being a very elegant theory. For example, in his 1911 paper *Relativitätsprinzip und Äther*, Emil Wiechert made such claim:

> Wiechert wrote that SR theory was "brought by Mikowski to a highly mathematically-finished form." He continued:
>
> > It was also Minkowski who, with bold courage, drew the extreme consequences of the theory for a new spacetime intuition and contributed so very much to the theory's renown.

---

not ad hoc but still solves as many problems as T_2'. Einstein's special theory of relativity might exemplify T_n while the Lorentz'modified aether theory was T_2. The obvious reply to such criticism is to ask why the admittedly *ad hoc* character of the Lorentz contraction constitutes a decisive handicap against it comparing it with SR. If the empirical problem solving capacities of the two theories are, so far as we can tell, equivalent, then they are (empirically) on a par; defenders of the view that the adhocness of T_2 makes it distinctly inferior to T_n must spell out why, in such cases, the comparable problem-solving abilities and equivalent degree of empirical support can be thrown to the winds simply by stipulating that ad hoc theories are intrinsically otiose. […]
To the extent that these same detractors set an epistemic premium on theories which work for the first time around, without any juggling or ad hoc adjustments, we are entitled to ask for the rationale for such a preference". From Laudan's *Progress and its Problems*; quoted in Grünbaum 1976, 358-9. One does not need, of course, to commit oneself to the Laudan's 'problem-solving model' to state that his claim is an appealing one.

[129] Moreover, also in the case of SR any concrete model of the electron must refer to something like a Poincare-stress to insure its stability. I thank Dennis Dieks for this remark.

[130] For a historical survey of the reasons on which the scientific community accepted the theory of relativity, see Brush 1999.

[131] From Laue's *Das Relativatätsprinzip*; quoted in Schaffner 1974, 74.

> It was precisely Minkowski's spacetime intuition, or his identification of the extreme consequences of this intuition, that had made the theory of relativity famous in Wiechert's view.[132]

With respect to simplicity, what I said in the first chapter concerning its relevance for the choice of a theory holds here too. It is true that this feature is a welcome one and that it is related to many pragmatic virtues. Were Einstein's and Lorentz's only two different formulations of the same theory, to favor the simpler one would be justified in many situations. Actually, it is rather probable that at the times when the scientific community still talked about the Lorentz-Einstein theory simplicity considerations may have somewhat turned the balance in favor of the 'Einstein-version'. However, if we take a sort of *van Fraasean* standpoint and claim that between two empirically equivalent *rival* theories the choice can only be done in terms of pragmatic virtues, we have that simplicity, for instance, cannot be related to the empirical success of the chosen one. This stance is quite close to quit to the basic requirement that theories are good or bad in terms of what they say about the world. Let me explain my point with some counterfactual history –counterfactual history is always tricky, but I think that my example is harmless. Imagine that the comparison in terms of simplicity in the case of Lorentz *vs.* Einstein was the other way around, so that Lorentz's theory was simpler than Einstein's, so we choose the former. As I will show below, it turned out that the electromagnetic theory on which Lorentz's was built was inconsistent with the explanations of the quantum-related phenomena discovered at the dawn of the 20th century. On the other hand, Einstein's theory did not contain such inconsistency. This difference is actually one of the good reasons to make a choice between the theories –as I will argue below. Therefore, in our fictitious example, simplicity would have led us to the *wrong* decision from the point of view of empirical success and adequacy[133]. Hence, simplicity issues cannot be trusted as a ground for theory choice –if we want that the theories we accept are accepted in terms of what they say about the world.

In the case of the mathematical-aesthetic features like *elegance*, something similar holds: they cannot be invoked if we want that our decision between rival theories may pick the empirically superior theory. In the previous example we can replace 'mathematical elegance' for 'simplicity' and the result would be the same. Moreover, in the case of SR *vs.* the LPT, we have that all the elegant features of the former theory which are connected to its formulation in terms of Minkowski's space-time can also be assigned to the LPT. Peter Galison's analysis of what were precisely the mathematical-aesthetic virtues introduced by Minkowski's work in connection with SR help to clarify this issue.

The first aesthetic virtue that Galison considers is *symmetry*. He is quite careful, though, in detaching the *geometric* Minkowskian notion of symmetry from the *physical* symmetry of Einstein –recall his dissatisfaction with 'the theoretical asymmetries which do not reflect in the phenomena'– and the *formal* symmetries of Poincare –the symmetry of the

---

[132] Walter 2010, 16.

[133] By *wrong* I mean that 'given a certain state of scientific knowledge' the decision would have been a mistake. I do not mean that by choosing a theory we grasp an absolute and everlasting truth, of course. Theories are empirically successful or unsuccessful "as far as we know".

coordinate transformations grounded on their group properties. Minkowski's view on symmetry, according to Galison, was rooted in his four-dimensional geometric-kinematic approach to physics:

> Minkowski, like Einstein, objected to the prevailing theory on what could be called aesthetic grounds. He objected to a lack of symmetry in the old physics, but a lack of geometric, rather than physical symmetry. Minkowski's new, geometrical symmetry is grounded in Poincare's *x*, *y*, *z*, *ict* formalism. In "The Principle of Relativity" Minkowski begins with Poincare's four-space and goes on to show that the Lorentz transformation is an orthogonal transformation for all vectors which transform like x, y, z, t. finally he reasons that physical laws composed of these four-vectors will be covariant. [...] He claims that covariance follows from the Lorentz transformation alone, that is, without any discussion of the status of the relativity principle. As he puts it, covariance follows "as a pure triviality, that is without the introduction of any new, previously unincluded law…". Only in the next section, on matter, does he introduce the "new law" of relativity.
>
> Different observers assign different coordinates to a given event. Minkowski reasons that since $t^2 - x^2 - y^2 - z^2$ is Lorentz-invariant, the four-dimensional hyperboloid $t^2 - x^2 - y^2 - z^2 = constant$ represents the set of all possible space-time coordinates of one event. The principle of relativity tells us that "absolute rest corresponds to no properties of the phenomena". Since in four dimensions there is a non-zero vector lying on the hyperboloid and corresponding to zero velocity, any point (x, y, z, t) on the hyperboloid can be transformed to lie on the t-axis. *Such a Lorentz transformation will take the hyperboloid back into itself. This is the geometric symmetry which Minkowski introduces into relativity. Its physical consequence is that no particular measurement of the coordinates of an event can indicate absolute rest.* […]
>
> The four-dimensional representation places rest and motion on equal graphical footing. *Since any four-vector can be transformed to the "rest-vector", leaving the hyperboloid of the appropriate invariance unchanged, the principle of relativity, i.e., that no phenomena are attached to absolute rest, stands fully exposed*. Such a symmetry is clearly distinct from the physical symmetry of Einstein and the formal group or group symmetries of Poincare.[134]

In spite of their differences, the symmetries considered by Einstein and Minkowski are connected, of course. The important point is that, according to Galison, Minkowski understands his *geometrical* symmetry as *grounding* Einstein's *physical* symmetry. This remark is quite coherent with the fact that the mathematician's work helped to develop and clarify the meaning of Einstein's theory.

Galison's second aesthetic feature consists in the generality that Minkowski's approach introduced. By this he simply means that the four-dimensional geometry and language that the German mathematician created permits to express different groups of invariant transformations, transformations which in turn determine different space-times with different metrics. Minkowski's four-dimensional stance is able to precisely describe the geometry and kinematics of both Newton's and Einstein's theories[135].

The third aesthetic factor that Galison refers to is given by the invariant interval that defines the metric of Minkowskian space-time:

> The existence of invariants for the relativistic transformations forms the third aesthetic criterion Minkowski considers in his four-dimensional relativistic theory. "The innermost harmony of these [electrodyamic] equations", he writes, "is their invariance under the transformations of the expression $dx^2 + dy^2 + dz^2 - dt^2$ into itself". In Newtonian space-time the free t-axis prevents us from constructing such an invariant expression. Like symmetry and generality,

[134] Galison 1979, 104-5.
[135] See the quote in footnote 117 above.

> invariance is an aesthetic geometric-criterion which supports the new conception of space-time.[136]

An evaluation of these three aesthetic factors in connection with Lorentz's theory, or with the LPT, shows that when Galison affirms that they *support the new conception of space-time*, this new conception cannot be Einstein's theory, but the approach to physical theories from the point of view of Minkowski's four-dimensional language. First, the invariance factor can be applied to both the theories, but with a different meaning of course. Whereas in the case of Einstein's it defines the metric of the space-time it depicts, in the case of the LPT it is just a mathematical nicety grounded by the dynamical features which govern bodies in motion through the ether. It suffices to remember that Poincare explicitly noticed this invariant. Second, something similar holds for the *symmetry* factor. It is true that the *geometric* symmetry of Minkowski is different from the formal one that Poincare made explicit. However, there is a weaker sense in which a sort of geometric symmetry can be assigned to the *Poincarean* principle of relativity. Minkowski's work shows that in the case of the *Einsteinean* relativity postulate, it is grounded in the geometry *of* the depicted space-time. In the case of Poincare's version, the four-dimensional approach allows to *represent* the principle in a geometric way –though we have to remind that it would not be an expression of the particular metric of the space-time depicted by the theory, but the outcome of dynamical features; but then again the mathematical properties of the quantitative expressions of those effects are such that they allow a geometric *representation* of the principle. Finally, the *generality* factor mentioned by Galison does not need any analysis in order to show that it holds for both the theories.

Therefore, these aesthetic factors have to be considered as reasons for the adoption of four-dimensional physics as a language and method, rather than as reasons for the adoption of *Einstein's* four-dimensional physics over *Lorentz's* four-dimensional physics. Galison's analysis is not posed within the context of the competition between these theories, but his conclusion is coherent with what I claim:

> If one grants that Minkowski can pass from good mathematics to productive physics, it remained for him to ground the new physics on mathematics alone. He accomplishes this by comparing Newtonian and relativistic theories on the basis of three criteria of geometrical elegance that emerge from his visual thinking: symmetry, generality and invariance. Together they seem to form the motivation and the justification for Minkowski's adoption of the new physics.[137]

> From his belief in the "pre-established harmony" [between mathematics and physics] and his discovery of these geometrically satisfying properties, Minkowski concludes that the four-dimensional theory is superior to Newtonian three-dimensional physics.[138]

In these two passages *the new physics* and the *four-dimensional* physics must be understood as referring to both Einstein and Lorentz's theories, whereas the Newtonian rival

[136] Galison 1979, 111-2.
[137] Galison 1979, 103
[138] *Ibid.*, 109.

is simply classical mechanics. Actually, as Galison clearly shows[139], Minkowski was a supporter of the electromagnetic worldview, and even when he wrote his famous papers on four-dimensional space-time he was thinking in terms of the Lorentz-Einstein theory –but he did not have the time to see the clarification that his work involved, for he died shortly after he published his seminal papers.

c) Minkowski space-time and explanatory power

Now I will consider one last reason to decide between SR and the LPT that I think does not work either. One could pick Einstein's theory in terms of its explanatory power or explanatory virtues, virtues and power which are absent in Lorentz's. The deepest and most precise argument along this line that I found in the relevant literature is Janssen's[140]. We already saw that the main difference between the theories at issue is that "according to the ether theory, the effects of length contraction and time dilation are due to peculiarities of all laws governing physical systems, causing them to deviate from the normal spatio-temporal behavior in the Newtonian space-time posited by the theory. In special relativity, these phenomena are simply part of the normal spatio-temporal behavior of systems in Minkowski space-time"[141]. Janssen points out that the program of the electromagnetic worldview expected to reduce all of these peculiarities to the basic properties of electromagnetic equations. However, the introduction in 1906 of the Poincare-stress resulted in the acknowledgment that the achievement of a purely electromagnetic description of the physical world was not possible. This failure, according to Janssen, results in that the ether-theorist has to commit to a rather striking coincidence: non-electromagnetic features of the world are such that the same peculiarities of electromagnetic laws hold for them too. The explanation of the Trouton experiment, as we saw, and also the explanations of the Trouton-Noble and the Kaufmann experiments, require to consider how energy and momentum transform; and in this context,

> for the ether theorist […] it is an unexplained coincidence that the stress-energy-momentum of both the electromagnetic and the non-electromagnetic parts of the Lorentz-Poincare electron and the Trouton-Noble condenser can be described by a quantity that transforms as a second rank tensor under Lorentz transformation, which, in fact, is just the property of these systems that accounts for the results of Kaufmann and Trouton and Noble.[142]

In Einstein's theory –in its Minkowskian formulation– there is no such striking coincidence. The Lorentz-invariance of all physical features is a direct consequence of the metric of space-time. According to Janssen, the explanation of Lorentz-invariance of physical laws that SR theory is able to provide is a reason to pick it in detriment of Lorentz's –the same holds in the case of the LPT, of course. Janssen's own explanation of his *common-cause* argument is the following:

---

[139] *Ibid.*, 90-5.
[140] Janssen 2002; and Janssen 1995, chapter 4.
[141] Janssen 1995, section 4.2.1.
[142] Janssen 1995, section 4.2.1.

> Lorentz invariance manifests itself in many different phenomena. Ultimately, these phenomena form the input of the common-cause argument. The most obvious examples are length contraction and time dilation […]. In the Newtonian space-time of Lorentz's theory this is a consequence of the laws. In the Minkowski space-time of SR, it is a consequence of the way in which particular space-time slices are used to define the length of a system or the duration of a process.[143]

I think that the situation of the electromagnetic worldview supporter was not that bad. Janssen simply assumes that the program went *bankrupt* with Poincare's work of 1906. However, I showed above that the French scientist had a hunch and a plausibility argument about the ultimate nature of his non-electromagnetic force that kept the electron stable, and this plausibility argument was such that it contained a promise for the possibility of including the Poincare-pressure in the electromagnetic picture of the world. Moreover, as late as 1909, there were big names in science that still believed that, after all, the electromagnetic program could be successful. Galison shows, as I mentioned above, that Minkowski was one of them; and Miller reports that Max Born, even after Minkowski's death, tried to complete the program[144]. Moreover, in the first decade of the 20th century, the very striking coincidence that Janssen mentions –understood as an anomaly to be solved– could be interpreted as an indication that the electromagnetic worldview was the right one after all. Born and others might have still found reasonable to expect that all of physics was electromagnetic in the end, in spite of the difficulties carried by Poincare's work. Actually, the acknowledgment that the program was unrealizable was the outcome of the inconsistence of its foundations with quantum physics; not of the introduction of the Poincare-stress.

In spite of this criticism, I think that Janssen's argument can be run anyway by means of an even deeper coincidence underlying the ether theory, one which could not be explained even if the electromagnetic program would have been successful. Assume that the program was successful, so that the ultimate nature of all the physical peculiarities that Janssen mentions were of electromagnetic origin. In that case, one could still ask why all those peculiarities –even if they had a *common* dynamical cause and foundation– are such that the Poincarean version of the principle of relativity holds. For example, why the Lorentz factor γ has the exact required value in order to make the ether-wind totally undetectable? This question can be made with respect to the whole electromagnetic core of the theory. The coincidence is analogous to the case of inertial and gravitational mass in Newtonian mechanics: it is a core part of the theory, but there is no theoretical explanation for it. In SR, of course, there is no such coincidence, for it is all a consequence of the particular metric of the space-time the theory depicts.

In think, though, that even in this more general formulation, the argument does not work. The first flaw in it that I will mention is that it rests on certain philosophical assumptions about the 'entities' that can act as a *cause* that are far from being totally justified. Janssen claims that the metric of space-time is the *cause* of the Lorentz-invariance of the laws

[143] Janssen 2002, 439.
[144] Miller 1998 [1981], 230-1.

of physics. One could immediately ask in what sense the metric of space-time, or space-time itself, can work as a cause of physical features. Actually, Janssen himself refers to this issue in his 1995 presentation of the argument:

> The reason for calling this a 'common cause'-type argument rather than a common cause argument, is that Minkowski space-time does not seem to be a common cause in quite the same sense that a shrimp cocktail contaminated with the salmonella bacteria is the common cause of the sudden death of half the population of a cheap Dutch old folks home.
>
> Although the status of the 'common cause' obviously needs further philosophical clarification, it is safe to say, I think, that this is a very strong argument for preferring special relativity over an empirically equivalent classical ether theory.[145]

Unlike Janssen, I think that the need for philosophical clarification of the status of the common cause invoked makes it *unsafe* to use it as an argument to decide between the theories. At first sight, it seems that a *substantivalist* position is taken with respect to space-time. As most of the philosophical debates, the one about the ultimate ontology of space-time is open, and to offer a criterion for theory choice which is based on a specific position in the context of an open philosophical debate is, I think, quite risky. Actually, and in a *relationist* spirit, one could say that there is no way in which space-time can be a cause, and that it is the Lorentz-invariance of physical laws what explains the metric of space-time rather than the other way around.

This position has important and relevant support. In a seminal paper published in 1976, John Bell introduced his *Lorentzian-pedagogy*, a didactic approach to SR sympathetic with a view in which the dynamical foundations of the world determine its kinematics and geometric features. More recently, and in a more philosophical stance, Harvey Brown has published several papers and a very influential book in which he explicitly argues that the *geometrical-kinematics* described by the Minkowskian space-time require a *dynamical* foundation. Allow me to quote at length a passage in which he clearly explains the basis of his interesting position:

> In his masterful review of relativity theory of 1921, Wolfgang Pauli was struck by the difference between Einstein's derivation and interpretation of the Lorentz transformation in his 1905 paper and that of Lorentz in his theory of the electron. Einstein's discussion, noted Pauli, was in particular "free of any special assumptions about the constitution of matter", in strong contrast with Lorentz's treatment. He went to ask
>
>> Should one, then, completely abandon any attempt to explain the Lorentz contraction atomistically?
>
> It may surprise some readers to learn that Pauli's answer was negative. Be that as it may, it is a question that deserves careful attention, and one that, if not haunted him, then certainly gave Einstein unease in the years that followed the full development of his theory of relativity.
> Einstein eventually came to realize that the first, 'kinematic' section of his 1905 paper was problematic; that it effectively rested on a false dichotomy. What is kinematics? In the present context it is the universal behavior of rods and clocks in motion, as determined by the inertial coordinate transformations. And what are rods and clocks, if not, in Einstein's own words "moving atomic configurations"? They are macroscopic objects made of micro-constituents – atoms and molecules– held together largely by electromagnetic forces. But it was the second,

[145] Janssen 1995, section 4.2.1

dynamical section of the 1905 paper that dealt with the covariant treatment of Maxwellian electrodynamics. Einstein came to see that the first section was not wholly independent of the second, and that the treatment of rods and clocks in the first section as primitive, or "self-sustained" entities was a "sin". The issue was essentially the same one that Pauli had stressed in 1921:

> The contraction of a measuring rod is not an elementary but a very complicated process. It would not take place except for the covariance with respect to the Lorentz group of the basic equations of electron theory, as well as those laws, as yet unknown to us, which determine the cohesion of the electron itself.

Pauli is here putting his finger on two important points: that the distinction between kinematics and dynamics is not fundamental, and that to give a full treatment of the dynamics of length contraction was still beyond the resources available in 1921, let alone 1905. And this latter point was precisely the basis of the excuse Einstein later gave for his 'principle theory' approach –modeled on thermodynamics– in 1905 in establishing the Lorentz transformations. […]
The main lesson that emerges, as I see it, is that the special theory of relativity is incomplete without the assumption that the quantum theory of *each* of the fundamental non-gravitational interactions –and not just electrodynamics– is Lorentz-covariant. This lesson was anticipated as early as 1912 by W. Swann, and established in a number of papers culminating in 1941. […] It is consistent with the didactic approach to special relativity advocated by J. S. Bell in 1976. […].
Swann's unsung achievement was in effect to spell out in detail the meaning of Pauli's 1921 warning above. His incisive point was that the Lorentz covariance of Maxwellian electrodynamics, for example, has no clear connection with the claim that the theory satisfies the relativity principle, unless it can be established that the Lorentz transformations are more than just a formal change of variables and actually codify the behavior of moving rods and clocks in motion. But this last assumption depends for its validity on our best theory of the micro-constitution of stable macroscopic objects. Or rather, it depends on a fragment of that quantum theory (for it could be no other than a quantum theory): that at the most fundamental level all the interactions involved in the composition of matter, whatever their nature, are Lorentz covariant[146].

In simple words, Brown's claim consists in that theories of principle do not offer a *complete* explanation of their *explananda*. A full understanding of them requires that the constructive foundations are clear. I think this is a rather appealing position. As he shows, Einstein himself was aware that he was quitting to a 'bottom-up' explanation in his formulation of SR. It is true that the *sin* of grounding the theory on rods and clocks without offering a theory describing the behavior of their micro-components was solved by Minkowski –for he grounded the theory on geometric-kinematic terms instead, as we saw above. However, this solution did not imply that a constructive explanation was unnecessary precluded from the scope of SR; there is certainly available room in it for such approach, and Einstein himself was totally clear about it[147].

[146] Brown 2005, 4-5.
[147] "In *Geometry and Experience* Einstein states: "The idea of the measuring rod and the idea of the clock in the theory of relativity do not find their exact correspondence in the real world. It is also clear that the solid body do not in the conceptual edifice of physics play the part of irreducible elements, but that of composite structures, which must not play any independent part in theoretical physics". This is very relevant to our topic: as we have seen, Einstein's original derivations of the contraction and dilations effects proceeded through the identification of distances and periods with the indications given by rods and clocks, without mentioning anything about the causal processes going on in the interior of these devices. Nothing at all was said about their atomic or molecular constitution and about the forces that keep them together, and as we have seen this could easily create the impression that no ordinary causal processes are involved at all in the contractions and dilations. But in the quoted passage Einstein emphasizes that rods and clocks are ordinary bodies with a microscopic structure and therefore determined in their macroscopic features by what occurs at the microscopic level.
In the same 1921 lecture, Einstein continues: "It is my conviction that in the present stage of development of theoretical physics these concepts [rods and clocks] must still be employed as independent concepts; for we are still far from possessing such certain knowledge of the theoretical principles of atomic structure as to be able to construct solid bodies and clocks theoretically from elementary concepts". [...] It is pretty clear that Einstein never

However, if one step further is taken, in the sense that the dynamical-constructive approach is stated as superior or more fundamental than the principle-kinematic one, then problems arise again. To pose the subject of constructive-dynamical explanations *vs*. principle-kinematic ones in terms of a question like which one is the cart and which one the horse is, I think, mistaken. There is no absolute and ultimate answer for the question of what explains what, 'does the metric of space-time explain the Lorentz-invariance, or the other way around?' The situation relating the two kinds of explanations, in the context of SR, is better posed in terms of *the hen and the egg*. If one assumes that kinematic top-down explanations are all what is needed, substantivalist presuppositions about space-time seem to be required. On the other hand, a view according to which *the* explanations have to be constructive is quite close to a relationist position. Therefore, I think that the *safest* position is simply to conceive the metric of space-time and the dynamics of objects in that space-time as two sides of the same coin.

This suggestion can be further supported by the description of the nature of scientific explanations that I offered in the first chapter. Explanations are context-dependent; as van Fraasen puts it, they are an answer to a determinate *why* question. Consequently, what is the form and content of a satisfactory explanation depends on the form and content of the particular *why* question that we are trying to respond. Dieks also argues for a similar view:

> Explanations and ways of achieving understanding are contextual in physics, no less than in other disciplines. As a consequence there exists no uniquely best explanatory scheme. Instead, there is a plurality of possible physical explanations and ways of understanding physical processes, and it depends on the type of question that is asked and on the aim and interests of the scientist that poses the question which one is the most appropriate. In other words, what is the best explanation and the best strategy for achieving understanding depends on contextual, pragmatic factors.[148]

In the specific case of SR, the contextual dependence of which is the most satisfactory explanation is rather clear. What Einstein's theory predicts in the case of the famous *Ehrenfest paradox* of the rotating disk, or John Bell's thought experiment of the space rockets connected by a rope, demands for a *dynamic* bottom-up consideration for its proper understanding. On the other hand, if what is at stake is the comparison between different inertial frames and what happens to bodies at rest in each of them, then it is a top-down *kinematic* explanation what provides an adequate understanding[149]. However –and this is just a suggestion– I think that the question concerning whether a *Lorentzian-like* approach is sufficient and adequate to provide a dynamical explanation of, for example, length-contraction in the context of SR is a difficult one. The Lorentzian physical foundation for the contraction is the electromagnetic interaction of the body with the ether; so from this point of view, without an ether the contraction becomes groundless, for there is no wind-effect to cause it. Furthermore, and more importantly, Lorentz's view of the contraction is a contraction *in a Newtonian space-time*

thought that general principles about rods and clocks should *replace* considerations about atomic constitution and causal processes." Dieks 2009, 6-7.

[148] Dieks 2009, 1.

[149] See Dieks 2009, 11-2.

which is not connected to its metric, and this view is inconsistent with what SR asserts –for Einstein's theory states that space-time is Minkowskian, not Newtonian. So, if a dynamical bottom-up account of the Lorentz-invariance of physical laws governing the microstructure of bodies is to be provided, it must be one such that the mentioned inconsistency and foundational conflict are avoided.

Coming back to Janssen's argument, I think that the way in which Lorentz's theory has been presented in this work makes it quite clear that it provides an electromagnetic, dynamical, bottom-up explanation of the 'relativistic' behavior of bodies in motion. On the other hand, Einstein's theory provides a top-down explanation –and they are incompatible in the sense that the LPT asserts that we live in a Newtonian space-time, whereas SR claims that we inhabit a Minkowskian space-time. It is true that Lorentz's view contains the striking coincidence I mentioned above, and that it cries for an explanation –that the theory does not provide. However, this shortcoming is compensated insofar as it *does* provide reasonable *dynamical* foundations, a feature which is absent in SR[150]. Simplifying, Janssen's argument is grounded on the assumption that SR is a better theory *because* it provides an explanation *of principle*. But if a position in which the adequacy of an explanation is context-dependent is taken, then the *explanatory* difference between the theories cannot be a reason to prefer one of them. It is quite obvious that these remarks constitute an instance of the general reason why explanatory power cannot count as a solution to the EE and UD problem that I stated in the first chapter: explanations are context-dependent.

2. Good reasons

a) The phantasmagoric ether

The first reason that I think is good in order to make an evaluative comparison between the theories refers, just like the ones I considered in the previous subsection, to a

[150] Actually, something like this might have been what Lorentz had in mind as a motive to prefer his own theory over Einstein's: "the main reasons why Lorentz never accepted the special theory of relativity were methodological. From his first published statement of his reservations in 1909 to his last in 1927, the theme remains the same: the special theory assumes what should be proven"; Nersessian 1986, 230.
"Thus, although he came to see that the "local time" and "universal time" variables must be treated mathematically as equal, he nevertheless always believed that there must be a difference in interpretation. Several times Lorentz discussed the counterintuitive and "paradoxical" nature of Einstein's interpretation. He argued that these problems disappear if "we decorate or should I say disfigure" one system with an ether with respect to which we can in principle leave open the possibility of determining the real time of events and the real dimension of objects. If we eliminate the aether, Einstein's interpretation follows, but then:

> One risks creating the impression that the question here is about 'apparent' things rather than about a real physical phenomenon... As against this one may remark that when we observe an 'alteration'... according to the customary linguistic usage (and why should we not cling to that?), this 'alteration' represents a physical phenomenon. The shortening of a fixed rod, that is made to move with respect to K, is just as real as the expansion by raising the temperature.

Lorentz's objection here –since of course the 'alteration' of length predicted in relativity theory is "a real physical phenomenon"– is best interpreted as a concern about causality. That is, when a measuring rod expands due to increase temperature, we can provide a causal explanation for the phenomenon. The same is true of the contraction of a measuring rod according to the theory of electrons: the phenomenon is given a causal explanation in terms of "molecular forces" and the ether. But, the special theory provides no causal explanation for this phenomenon –nor does it need to"; *Ibid*, 233. After the provisos introduced by Brown and Dieks, for example, one should be careful when interpreting what does it mean that Einstein's theory does not need to provide a causal explanation.

non-empirical feature. I propose that the status of the ether in the LPT is such that it undermines the theory concerning what one expects a scientific theory must be. It is often remarked in the literature on Lorentz' theory that the ether it postulates is undetectable. That is quite true. The interaction of bodies in motion with the ether produces compensating effects that lead to the Poincarean principle of relativity: the (deceptive) measured quantities in any inertial frame are the same as the (real) quantities in the ether-rest frame; therefore, there is no possible experiment capable to show any trace of motion relative to it.

This feature is sometimes invoked to prefer Einstein's theory, for in it the ether is totally superfluous, and by a simple application of Ockham's razor, one can take it as stating that the ether simply does not exist. However, I think that this is not a feature that one can use to make the choice. First of all, in spite of its un-observability, there might be an ether anyway. The fact that a certain entity is described as non-observable in the context of a theory is not a reason to immediately discard that theory. Actually, many successful scientific theories contain entities that can be considered as such. Janssen poses this issue –in the context of the discussion of the *ad-hoc*ness of Lorentz's theory– with the following words:

> As part of his analysis of the doubly-amended theory [Lorentz's theory including the generalized contraction hypothesis and the dilation of local time as an observable feature] Grünbaum offers a more accurate diagnosis of the trouble with Lorentz's theory. What makes the theory unsatisfactory are not the elements that are added, but some of the original elements, notably the ether and Newtonian space-time, that are rendered more and more invisible with every amendment. This suggests that Ockham's razor is all that is needed to settle the case of Einstein *versus* Lorentz. The problem with this type of argument is that it derives its force from a blanket rejection of unobservables in scientific theories, whereas it is widely accepted that such elements should not be banned automatically. Rather than condemning unobservables in general, I think it is wiser to demand that arguments put forward on a case-by-case basis to show why a particular unobservable is otiose.[151]

Moreover, the ether that the LPT postulates is the basis for the constructive-dynamical explanation it provides; and as we saw above, this feature can be understood more as a virtue than as a flaw.

Nevertheless, I think that Janssen's demand for a case-by-case argument to show that the problematic entity is otiose can indeed be put forward. To do that we must recall one of the crucial amendments that Poincare introduced. In his 1906 work he showed that the Lorentz transformations form a group only if the factor l is set to unity, and that by doing so the transformations become totally symmetric: the inverse transformation of a Lorentz transformation is simply one in that v is replaced by $-v$ and in which primed and unprimed quantities shift roles. This amendment, we also saw, is completely necessary if one is to make a case for the empirical equivalence between the theories at issue. If we analyze what is the consequence of this adjustment in terms of the physical meaning of the Lorentz

[151] Janssen 2002, 438. One must be careful here, though. The kind of unobservables which are widely accepted, according to Janssen, as a non-problematic part of scientific theories must be terms and entities which are not *directly* observable but that at least leave a *trace* in certain observable phenomena. I think that the ether, in the original meaning of the Lorentz transformations, can be understood in that way. The length of a rod at rest in a frame is L, but when set in motion in that same frame rest its length becomes $L\sqrt{1-v^2/c^2}$. This effect is not enough as to detect the ether-rest frame or the velocity of motion with respect to the ether, but it could be understood as a *trace* of the ether. However, once the symmetry of the transformations is considered, this view becomes problematic.

transformations, we have that the velocities involved become simply the *relative* velocities of the frames. Accordingly, the velocity of frames *with respect to the ether* becomes irrelevant for the coordinate transformations. The situation is then that the ether is the basis for the rationale of the contraction included in the Lorentz transformations. Interpreted from the point of view of the generalized contraction hypothesis, their meaning is that bodies moving with a velocity v with respect to the ether contract because they interact with it; however, in their corrected form and interpretation, the v included in the transformations is no longer relative to the ether, but only to the frames involved.

I think this is a very unsatisfactory feature of the theory. It is not only true that the entity at issue is undetectable and unobservable, it also becomes otiose from a *theoretical* point of view. Before Poincare's amendment, even though undetectable as a consequence of the compensating effects, Lorentz's theory asserted that there *is* a frame with respect to which the inverse transformation is an asymmetrical one. But after Poincare's contribution the full symmetry of the transformations implies that the special frame becomes almost a *theoretical fiction* which is not referred to by the coordinate transformations –I say *almost* because it still explains why the contraction happens, though one could say that this explanatory role makes the entity even more problematic if one considers the theoretical *evaporation* of the ether. For example, in the original meaning of the asymmetric Lorentz-transformations, the length contraction phenomenon could be considered as a *trace* of an undetectable ether, for the effect was related to the interaction between a moving rod and the ether expressed in the velocity v. Once the symmetry of the transformations is considered, the physical background for the contraction is still the ether –for it is the cause of the moving electron to change its shape and the cause for the molecular forces to transform in the way they do–. However, since the relevant v in the transformations is simply the relative velocity between the frames involved, then the length-contraction effect cannot be so easily conceived as an observable trace of the interaction with the ether, for it is related only to relative motion[152].

I think that this non-empirical feature can be used to claim that Einstein's is a better theory than LPT. Unlike considerations of simplicity and considerations of explanatory matters, this line of reasoning does not include subjective or context-dependent elements. The status of the ether within the LPT is an *objective* feature; and on the other hand, the fact that the phantasmagoric nature of the ether is a flaw, I think, does not depend on pragmatic-contextual issues. The reason why it becomes a problematic entity is because its theoretical *evaporation* gives it a rather metaphysical flavor. This is more than merely stating that it is a

[152] Assume that S is the ether-rest frame, and that the length of a rod at rest in it is L. When set in motion with velocity v with respect to S, the rod's length in S becomes $L\sqrt{1-v^2/c^2}$, of course. Consider now a frame S' which moves with the same velocity v with respect to S and the ether, so that the moving rod in S is at rest in S'. Within an asymmetric conception, the corresponding transformation for the length of the rod from S' to S should be the *inverse-function*, that is, a transformation that gets the rod *un*-contracted –or from the point of view of S', expanded. However, when the full symmetry of the transformations is considered this interpretation is no longer possible: the rod at rest in S' is contracted from the point of view of S, and the rod at rest in S is also contracted from the point of view of S'. The contraction is associated only to the relative motion of the frames. However, in the LPT, the ether still provides the explanation of why the contraction happens.

(directly) unobservable entity, and I think it is a widely accepted and justified desideratum for scientific theories that they do not include this kind of quasi-metaphysical items.

I claim that this is a good reason to state that –from the point of view of what we expect a scientific theory may be– SR is a better theory than its rival. That is, from the standpoint of the general and basic goals of science, we can say that Einstein's accomplishes *one* of those goals in a more satisfactory way than the LPT. However, I also claim that if we consider the most important goal of science, namely, empirical success, this non-empirical criterion could also lead to the wrong decision. If we apply the counterfactual history view that I used above, the result is that later developments might have perfectly shown that the LPT was more empirically successful than Einstein's after all. In that case, I think that the LPT is a theory which in spite of the problematic ether it includes is scientific enough as to be accepted –had it defeated its rivals on the empirical battlefield. My stance is that, if the very basic conditions of theoreticity –such as the ones I described in the first chapter– are accomplished, then the ultimate criteria of theory choice will always be empirical. It is true that some non-empirical considerations can be qualified as objectively virtuous or vicious, and therefore as welcome or unwanted; but they remain subsidiaries with respect the empirical supreme-court when it comes to theory choice. I will return to a more general evaluation of this issue.

b) The LPT, classic electrodynamics and quantum physics

Now I will consider two *empirically and evidentially grounded* reasons that do work as a motive for the choice. The first one is rooted in the inter-theoretical relation between the LPT and quantum physics. The second relies on the logical and conceptual connection between SR and the general theory of relativity. I will explain them in turn.

Lorentz's theory is built on Maxwell's equations, which constitute the core of classic electrodynamics and, *a fortiori*, the core of the electromagnetic world view. This theoretical core, interpreted in the classic way, started to face deeper and deeper problems to provide a satisfactory account of new phenomena at the dawn of the 20th century. The first of them was the problem of black-body radiation. The solution for this problem that Planck introduced turned out to be inconsistent with the foundations of classic electrodynamics. This particular problem was not a problem *directly* for Lorentz's theory –neither for the LPT. However, since it showed that the deep puzzle was rooted in the very foundations of classic electrodynamics, this empirical riddle *flowed* to Lorentz's theory –and to the LPT too. On the other hand, SR, as we saw above, did not contain any specific assumptions about the ultimate nature of matter –meaning mass *and* energy– and it was not essentially grounded on classic electrodynamics. Therefore, the empirical problem I mentioned did not affect it.

I showed above that by the last years of the 19th century the electromagnetic worldview arose as a revolutionary program to replace mechanics as the basic and universal framework under which physical science was to be understood. Lorentz and others, up to a certain extent, succeeded in reducing Newtonian mechanics to electromagnetic laws. In 1900

Lorentz also made an attempt to include gravitation into the scope of the electromagnetic view[153], attempt which was further developed by Wilhelm Wien. Even though it was not a successful endeavor some of its results were received as promising. It is true that the introduction of non-mechanical forces was a drawback for the program, but the real problems started with Planck's work.

Between 1900 and 1903 Lorentz published a series of papers devoted to thermodynamics from the point of view of electron theory[154]. The main goal of these works was to found Boltzmann's and Wien's radiation laws on properties of the electrons, based on a study of the thermodynamic characteristics of metals. His results allowed him to derive a formula for the black-body radiation which was consistent with Planck's in the domain of long wavelengths and low frequencies. Lorentz evaluated this outcome very positively, for he inferred from his results on the issue that the goal of reducing thermodynamics to electron theory was achievable. With his characteristic scientific honesty, Lorentz also pointed out that the main difficulty was the disagreement of his formula with Planck's –and with the relevant observed phenomena– in the case of short wavelengths and high frequencies. However, later developments showed this difficulty to be unsolvable for the electromagnetic view. Moreover, the satisfactory solution available, namely, Planck's, was inconsistent with the foundations of classic electrodynamics:

> In 1908 Lorentz came out in support of Planck's theory; it was then that he emphasized the profound antithesis between the quantum hypothesis and the electron theory. At a mathematical congress in Rome that year Lorentz spoke on Planck's and James Jeans' theories of black-body radiation. His object was to prove that the union of electron theory with Hamilton's equations for motion and J. W. Gibbs' statistics leads inescapably to Jean's radiation law, which, like his own of 1903, agrees with experience only in the case of long wavelengths. He said that the alternative, Planck's theory, demands far-reaching changes in electron theory. He pointed out that this is easily seen, since an accelerating electron should emit rays of all wavelengths, a result incompatible with the hypothesis of energy elements whose magnitude depends on wavelength. At the time of his lecture he had not decided between the two theories. Wien, however, called his attention to experiments showing that for short wavelengths a body emits much less light in proportion to its absorbing power than that predicted by Jean's theory. This proves, Lorentz said in a note appended to the published version of this talk, that any theory that bases itself on the electron theory and the equipartition theorem has to be profoundly revised. […]
>
> Lorentz thus accepted the quantum theory as the only capable of explaining the complete spectrum of black-body radiation, while at the same time regarding it as very incompletely understood in its connection with the other branches of physics and in particular with electron theory.[155]

The electromagnetic framework proved incapable to provide an empirically successful account of black-body radiation. In order to obtain one, Planck introduced the hypothesis of the discrete quantum of energy, but this hypothesis was inconsistent with the classic-electromagnetic framework underlying Lorentz's electron model and its attempted connection with thermodynamics –and this framework was also the foundation of the

---

[153] See McCormmach 1970b, 476-7.

[154] *"The Theory of radiation and the Second Law of Thermodynamics"* (1900), *"Boltzmann's and Wien's law of radiation"* (1901), and *"On the Emission and Absorption by Metals of Rays of Heat of Great Wavelengths"* (1903). For the full references see McCormmach 1970b, 486.

[155] McCormmach 1970b, 487.

electromagnetic worldview. Lorentz's ether theory –and also the LPT–, insofar as it was grounded on the very same framework, was affected by this problem. This was an *indirect* empirical complication, for Lorentz's ether theory does not make any specific predictions on the realm of thermodynamic phenomena –radiation and heat are physical features that are outside its scope. Therefore, the failure of the theoretical accounts of the black-body spectrum of radiation cannot be considered as a direct falsification of it.

The historical course of events indicates that this problem was very important when it comes to the competition between Einstein's and Lorentz's theory. McCormmach points out that in the first Solvay Congress in 1911 there was general agreement –with exceptions, of course– about the electron theory based on classic electrodynamics being irreconcilable with the quantum hypothesis. Bohr's 1913 first contributions to the model of the atom deepened that feeling, for the quantum hypothesis was also a crucial assumption in it. I think it is not coincidental that around these years the expression "Lorentz-Einstein theory" disappeared from scientific vocabulary.

On the other hand, it is important to underscore that the abandonment of a program of physics based on classic electrodynamics did not have the form of a straightforward falsification. The process was more complex and deeper than that. The empirical anomalies that classic electrodynamics faced were understood as a manifestation of the inadequacy –both empirical and *theoretical*– of its foundations, not only as a particular cases of empirical failure. As Helge Kragh reports on the reasons of the fall of the electromagnetic worldview,

> More important [than the problematic outcomes of the Kaufmann experiments] was the competition from other theories that were either opposed to the electromagnetic view or threatened to make it superfluous. Although the theory of relativity was sometimes confused with Lorentz's electron theory or claimed to be compatible with the electromagnetic worldview, about 1912 it was evident that Einstein's theory was of a very different kind. It merely had nothing to say about the structure of electrons and with the increasing recognition of the relativistic point of view, this question –a few years earlier considered to be essential– greatly changed in status. To many physicists, it became a pseudo-question. As the rise of relativity theory made life difficult for electromagnetic enthusiasts, so did the rise of quantum theory. Around 1908, Planck reached the conclusion that there was a fundamental conflict between quantum theory and the electron theory, and he was cautiously supported by Lorentz and other experts. It seemed that there was no way to derive the blackbody spectrum on a purely electromagnetic basis. As quantum theory became more and more important, electron theory became less and less important. The worst thing that can happen to a proclaimed revolution is that it is not needed.
>
> In general, electron theory had to compete with other developments in physics that did not depend on this theory, and after 1910 new developments in physics attracted interest away from the electron theory. So many new and interesting events occurred, so why bother with the complicated and overambitious attempt to found all physics on electromagnetic fields? Rutherford's nuclear atom, isotopes, Bohr's atomic theory, the diffraction of x-rays by crystals, Stark's discovery of the electric splitting of spectral lines, Mosley's x-ray-based understanding of the periodic system, Einstein's extension of relativity to gravitation, and other innovations absorbed the physicists' intellectual energy and left the electromagnetic worldview behind. It was a beautiful dream indeed, but was it physics? Much progress took place in atomic physics and as the structure of the atom became better understood, it became increasingly difficult to uphold the electromagnetic view.[156]

---

[156] Kragh 1999, 115.

SR, on the other hand, did not face any of these complications, for it was free of any presuppositions of the ultimate nature of mass and energy. This feature seems to have been explicitly considered by Einstein when he formulated it. If we recall that the first published paper of his *annus mirabilis* was the one on the hypothesis of the light-quantum, it is rather natural that he may have had the conflict of classic electrodynamics with the quantum hypothesis in the back of his mind when he invented SR. As Einstein himself stated in his *Autobiographical Notes*,

> Reflections of this type [the clash between the quantum hypothesis and classic electrodynamics] made it clear to me as long ago as shortly after 1900, i.e., shortly after Planck's trailblazing work, that neither mechanics nor thermodynamics could (except in limiting cases) claim exact validity. By and by I despaired of the possibility of discovering the true laws by means of constructive efforts based on known facts. The longer and the more despairingly I tried, the more I came to the conviction that only the discovery of a universal formal principle could lead to assured results.[157]

Moreover, Einstein's maneuver of stating the second postulate as the constancy of the velocity of light independently of the motion of the source rather than as a statement claiming that Maxwell's equations –which govern the behavior of light– hold in any inertial frame, might have been grounded in the goal of avoiding any special commitment to electrodynamics. If he had taken the second alternative, the constancy of the velocity of light would be simply a specific instance of the invariance of Maxwell's equations. However, this move would have carried along the conflict that got him in despair in the first place.[158]

c) Special and general relativity

In order to explain the third reason that can be used to decide the case between the theories at issue it is necessary to take a glimpse on Einstein's road to the formulation of the general theory of relativity (GR)[159]. Soon after he introduced SR, he thought that the principle of relativity should be extended and generalized to encompass also non-inertial motion. He found that the reference to inertial forces included in his 1905 theory –which in turn refer to absolute space-time– was a problematic feature, for example. By applying the relativity postulate to accelerated motion, he believed, the reference to inertial forces would be no longer needed. On the other hand, Einstein also wanted to develop a theory of gravitation grounded on his results of 1905.

The crucial insight that allowed him to tackle both issues came from what he called *the happiest idea of his life*. This idea, which is the seed of his crucial *principle of equivalence*, was

[157] Schilpp (ed.) 1949, 21.
[158] See Zahar 1973, 233-5; and Schaffner 1974, 56-9. Schaffner offers an analysis in which Einstein's theory was favored over Lorentz's by means of inter-theoretical considerations: the conflict between the quantum hypothesis and classic electrodynamics. However, he claims (pages 75-6 of his article) that the way in which the conflict decided the competition was grounded on the fact that the rise of quantum theory made the generalized contraction hypothesis *ad-hoc*. It might be right that the plausibility arguments for the contraction that Lorentz offered became untenable after quantum theory appeared, but in the best case this is simply a specific instance of the deepest foundational problem at hand. Actually, as Kragh's quoted passage indicates, the whole electromagnetic view was put aside by quantum theory, not only Lorentz's ether theory. The *ad-hoc*ness was simply *collateral damage*, not the fatal wound.
[159] On this subject I closely follow Torretti 2003, § 1.4.

simply that a freely falling observer in a gravitational field does not feel his own weight, so that if the observer were to perform experiments in physics their outcome would be the same as if the he was at rest or in inertial motion. The general conclusion of this thought experiment is that a frame which freely falls in a homogenous gravitational field is equivalent to an inertial frame –and this statement constitutes the 'first half' of the principle of equivalence. On the other hand, it is also the case that an observer in an elevator at rest in a homogeneous gravitational field –which is therefore resisting the pull of the massive object which generates the field– is indistinguishable from a case in which an observer is inside an elevator in outer space, free from any gravitational attraction, but in accelerated motion –with the acceleration having the same value of the gravitational pull but opposite direction. More generally, a frame at rest in a homogenous gravitational field is physically equivalent to a frame accelerating with the same value but in the opposite direction of the attraction of the field –and this statement constitutes the second part of the principle of equivalence[160].

The connection just stated between acceleration and gravitation was thus the milestone for Einstein's path to GR. It allowed a solution for a long-lasting puzzle in the context of Newton's theory: the coincidental equality of the values of inertial and gravitational mass for any object. Inertial mass –the *stubbornness* of bodies to remain in their state of motion–, and gravitational mass –the measure of a body to gravitate– have exactly the same value within Newton's theory, even though they represent different physical properties –and the theory does not provide any explanation for this coincidence. The link that the principle of equivalence establishes between acceleration and gravitation allowed a way out of the puzzle: inertial and gravitational mass are the same property[161]. And if acceleration and gravitation are two sides of the same coin, Einstein's goal of extending his theory to accelerated motion and to gravitational phenomena could be achieved by the same theoretical stroke.

The full significance of the principle of equivalence was put in front of Einstein's eyes by means of a group of problematic features he found. Maybe the most important one was given by the analysis of what happens in a rigid rotating disk. Imagine a flatlander equipped with a measuring rod living on the surface of the disk. If he were to make a measurement of the circumference of the disk he would find it to be greater than $2\pi r$, for his rod would get contracted when put along circumference but would remain the same when put along the radius –the contraction occurs only in the direction of motion, and the radius is always transversal to it. The surprising result is then that the geometry of the rotating disk is not Euclidean, and given the principle of equivalence, the geometry in a frame at rest in a gravitational field is not Euclidean either.

---

[160] The restriction to homogeneous gravitational fields is required to neglect the effects of the different values of the gravitational potential in different points of the field in the real cases, and the effects produced by the fact that the vectors of the field are radials, not parallel. Such effects would, of course, break the equivalence.

[161] More precisely, consider the Newtonian equations $F = m_i a$, where m_i is the inertial mass; and $W = m_g g$, where W is the weight, m_g the gravitational mass and g the acceleration of gravity. Applying these equations for the acceleration in free falling –substituting W for F– one obtains $a = (m_g/m_i)g$. As experience shows, the gravitational acceleration exerted on *any* freely falling body is the same. This can only be the case if the ratio between the two different masses is the same for any object. This is the unexplained coincidence.

The theoretical framework that Einstein created to provide an explanation of this issue was the result of a sort of analogical reasoning. During the 19th century Gauss had created mathematical methods to deal with the intrinsic geometric properties of curved surfaces. If x^1 and x^2 are Cartesian coordinates on a surface, the length of a line C in that surface is given by the expression $\int_C ds = \int_C \sqrt{(dx^1)^2 + (dx^2)^2}$. The *line element ds* can also be written by means of the Kronecker symbol, such that $\int_C ds = \int_C \sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} \delta_{ij} dx^i dx^j}$. If the Cartesian coordinates are replaced by *curved* coordinates u^1 and u^2, then the lines defined by $u^1 = constant$ and $u^2 = constant$ are curves which form variable angles in their intersection points –whereas the corresponding lines defined by $x^1 = constant$ and $x^2 = constant$ in the Cartesian case *always* cut each other orthogonally. The length of a line in terms of the curved coordinates can be expressed in a similar way, but the constant factors δ_{ij} must be replaced by factors that vary with position. Designating those factors as g_{ij}, then the length can be stated as $\int_C ds = \int_C \sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} g_{ij} du^i du^j}$; and a geodesic line in the curved surface is given by $\delta \int_C \sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} g_{ij} du^i du^j} = 0$. The line elements $\sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} \delta_{ij} dx^i dx^j}$ and $\sqrt{\sum_{i=1}^{2} \sum_{j=1}^{2} g_{ij} du^i du^j}$ are, of course, different[162]. The former is defined for Cartesian coordinates, whereas the latter is defined for curved ones. However, they coincide when infinitesimally or locally considered. Locally or infinitesimally speaking, the expression for length in curved coordinates reduces to the expression for Cartesian ones. Finally, the approach introduced by Gauss also allowed to quantitatively evaluate the degree of curvature of a surface. For example, flat surfaces have a 0 curvature, the surface of a sphere is constant and positive, and the surface of an egg is positive but varies with position.

The adoption of the Minkowskian point of view allowed Einstein to notice the analogy between the geometrical work of Gauss and the results he obtained from his principle of equivalence. In an article of 1912, entitled *On the Theory of the Static Gravitational Field*, Einstein wrote the equation for the world-line of a material point freely falling in a static gravitational field as $\delta\left\{\sqrt{c^2 dt^2 - dx^2 - dy^2 - dz^2}\right\} = 0$. As I showed above, this is the equation for a time-like geodesic in a Minkowskian space-time, the equation for an *inertially moving* particle. However, one of the results that his principle of equivalence entailed was that light gravitates, and therefore in inhomogeneous gravitational fields the speed of light depends on the specific position-dependent value of the gravitational potential. Consequently, a more accurate and general way to write the equation was $\delta\left\{\sqrt{[c(x,y,z)]^2 dt^2 - dx^2 - dy^2 - dz^2}\right\} = 0$. If we replace x, y, z, t with x^0, x^1, x^2, x^3, and if we define $g_{00} = c(x^1, x^2, x^3)$, $g_{kk} = -1$ if k is greater than 0, and $g_{kh} = 0$ if $k \neq h$; then the equation can be written as $\delta \int_C \sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} g_{ij} dx^i dx^j} = 0$. This equation is different from a

[162] More precisely, only in a Euclidean plane it is possible to define Cartesian coordinates x^1 and x^2 such that the g_{ij} quantities –expressed as a function of the coordinates– satisfy the relation $g_{ij} = \delta_{ij}$.

Gaussian geodesic only by the amount of dimensions at stake, and the genius analogy that Einstein noted was that it represents the equation for a geodesic in space-time. It had to be a *curved* space-time because he had already discovered that the spatial metric of an accelerated frame –and of its equivalent frame at rest in a gravitational field– are not Euclidean, and thus the spatio-temporal metric in those frames could not be Minkowskian; i.e., the g_{ij} factors are position-dependent. The curvature expressed by the specific g_{ij} was associated to acceleration and gravitation determining the space-time considered: the non-Euclidean geometry of the rotating disk was the outcome of its acceleration, and by the principle of equivalence one has that gravitation results in the same feature. This line of reasoning naturally suggested Einstein that gravitation is not a force acting at a distance between distant massive bodies, but simply the curvature of space-time that the bodies produce. Since $\delta \int_C \sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} g_{ij} dx^i dx^j} = 0$ defines a time-like geodesic in a curved space-time, it defines the *inertial* motion of a particle in a region of space-time determined by a gravitational field: free-fall is simply inertial motion in a curved region of space-time, and the curvature is produced by the presence of mass-energy; *this* is gravitation, not a distant-instant force. Yet another important feature of the analogy with the Gaussian approach was that the 'Minkowskian line element' relates to the line-element defined for the curved space-time case just as the Cartesian line element relates to the one defined for curved coordinates; i.e., as a local-infinitesimal region of the global surface.

This was, in a nutshell, the line of reasoning that led Einstein from SR to GR. The *Gaussian analogy* he discovered got reinforced by Einstein's collaboration with Marcel Grossman, who introduced him to the work of Riemann on the generalization of the Gaussian approach for *n*-dimensional manifolds. The road to the specific quantitative relation between the mass-energy distribution and the specific metric of space-time was a difficult and intricate one, but the goal was finally achieved in 1916 with the introduction of the field equations that form the core of the new theory.

Coming back to our subject, we have that the relevance that GR has with respect to the Einstein *vs.* Lorentz case is that it reduces the special theory to a limiting case. The Minkowskian space-time that SR defines is, from the point of view of the general theory, a specific solution of the field equations in which there is no mass-energy to produce any curvature-gravitation-field, or simply a local-infinitesimal piece of a global space-time defined by the field equations –just as the line element defined for Cartesian coordinates can be considered as determining an infinitesimal piece of a curved surface, according to the Gaussian approach. More simply, the Minkowskian world describes either an empty universe, or a tiny piece of a universe which does contain mass-energy. The crucial point is that GR entails empirical consequences that SR does not. The most spectacular and historically relevant ones were the bending light effect and the predicted advance of Mercury's perihelion. The observation of both phenomena determined a bombastic acceptance of the theory; and to accept GR *logically entails* to accept SR, in the sense that one has to accept that the space-time the latter describes holds for the specific cases mentioned –the empty space-time or the infinitesimal region of a curved space-time.

The essential mathematical and *physical* connection between the special and general theories does not exist between the LPT and the GR. If we remind ourselves that in the context of the Lorentz-Poincare view the invariant interval does *not* refer to the metric of a four-dimensional space-time –it is nothing but a *mathematical nicety*–, we can easily see that we cannot say that the world the LPT describes is a specific solution of the field equations or a local-infinitesimal piece of a global curved space. It is true that the LPT can be presented as a *mathematical* consequence of Einstein's field-equations under certain constraints –its mathematical structure is identical to SR, after all. However, from a physical-semantic point of view the connection does not hold. GR states that an empty universe would be determined by a Minkowskian *metric*, and that locally speaking the *metric* of a curved space-time can be described by the 'Minkowskian line element'. But the invariant interval, within the context of the LPT, has nothing to do with the metric of space-time, it is simply a mathematical nicety produced by the specific value of compensating dynamical effects such as the Lorentz-Fitzgerald contraction and local time. The real metric of the space-time that the theory defines is Newtonian. More generally, it is the *physical meaning* of GR what precludes that the LPT could be understood as the expression of a limiting case. In order to do that, the meaning of GR should have to be severely altered, in a way such that the theories could become coherent. Therefore, the acceptance of GR does *not* entail the acceptance of the LPT. Moreover, since the former claims that an empty universe would have a *Minkowskian metric* and that a *Minkowskian metric* describes an infinitesimal portion of a curved space-time, its acceptance entails a *rejection* of the LPT, for this theory claims that the metric of space-time would be *Newtonian* in both cases. More generally, even though the mathematical structure of the LPT can be derived from GR, the different meanings of the two theories make them incompatible.

Before moving to the general conclusions of this work, I would like to clarify the nature of the views I just offered. I proposed three reasons I consider effective and justified in order to accept Einstein's theory over the LPT. These reasons are proposed from a *conceptual* point of view, and they do not intend to describe *historical facts* concerning the reasons that the scientific community of the times endorsed to make the choice. Actually, the third reason I dealt with was not really important from a historical perspective. The final formulation of GR occurred in 1916, and by then SR had already won the competition over the Lorentz-Poincare approach. The other two reasons were indeed important from a historical standpoint, and also were the *aesthetic* features which I described as *bad* reasons for the choice. However, I do not intend to claim that the acceptance of the Einstein's theory ca. 1911 was an irrational choice insofar as it considered aesthetic factors. My intuition on this subject is that the aesthetic features of SR, the problematic ether and the incoherence between the LPT and quantum theory were reasons that worked all together and at once, and that the scientific community did not care about making a thorough analysis of the issue. However, this is just a hint, for I am not writing a thesis on the history of physics.

CONCLUSIONS

As I mentioned in the Introduction of this work, my main goal is to evaluate the general philosophical solutions that have been provided for the problem of EE and UD by means of an analysis of a *real* example of the problem: the Lorentz-Poincare *vs*. Einstein case. Consequently, I will now present such an evaluation. I will proceed by considering each of the three good reasons I proposed in turn.

The first of the reasons I offered was based on a non-empirical feature. I claimed that within the LPT there is an entity, the ether, which is deeply problematic. It forms an essential part of the theory, for it plays the role of the rationale grounding the Lorentz-Fitzgerald contraction. However, it becomes a problem when one realizes that it is not only an unobservable entity, but also an entity expressed by a concept that *theoretically evaporates* once the full symmetry of the coordinate transformations is considered. The concept of the ether in the LPT is a quasi-metaphysical one.

The status of this theoretical evaporation can be clarified by considering some of the remarks that John Norton made on the problem of EE. As I showed in the first chapter, he states that if two theories are EE:

> The two sets of theoretical structures may be interconvertible without loss; or they may not be. In the latter case, there would be additional structures present in one theory but not in the other. However, any such additional structure will be unnecessary for the recovery of the observational consequences. That follows since the additional structure has no correlate in the other theory, yet the other theory has identical observational consequences. Thus any additional structures will be strong candidates for being superfluous, unphysical structures[1].

I think it is rather clear that this is a good description of the problem of the ether in the LPT. Once the symmetry of the Lorentz-transformations is considered, it becomes apparent that the ether is not needed for the derivation of empirical predictions. Even though in the first formulation of Lorentz's theory the ether was already directly undetectable, the fact that the velocity included in the Lorentz transformations referred to it made it plausible to interpret the length contraction effect as an observable *trace* of the ether. However, after Poincare's contribution, the length contraction effect gets related only to the relative velocities between the frames –and the ether no longer participates in the derivation of this effect. The resulting situation is that the ether is directly unobservable and there are no observable *traces* of its existence. Therefore, it gets accurately described in terms of Norton's concept of *superfluous and unphysical structure*[2]. Within Einstein's theory, on the

[1] Norton 2008, 35.

[2] Recall that Norton takes his general observation as an indication that EE theories are simply two different formulations of the same theory –in one of them there is simply extra and superfluous structure. Recall also that he explicitly states that his remarks count as a *suggestion* that we are dealing with identical theories. In the case of Newtonian theory and the van Fraassean alternative formulations that include a value for the absolute velocity of the solar system, the suggestion looks to be totally true. However, in the Einstein *vs*. Lorentz case things are not so simple. The difference between the theories is not simply the presence or absence of the ether. If we get rid of it in the LPT we do not obtain SR right away. The theories also differ in terms of the metric they assign to space-time. In the LPT is still possible to say that, even without the ether, local-time and length-contraction are dynamical effects which are not related to the metrical properties of space-time. This would be a very implausible stance, of course, but it shows that the difference between the theories is not simply the presence of the ether; SR and the LPT are

other hand, there is no such problem. Therefore we could say that SR is a better theory than the LPT because it does not include any kind of superfluous-unphysical structure.

To evaluate how this reason to choose fits within the views I proposed in the first chapter, we have that it relates to the concept of *theoreticity* I analyzed. This concept refers to the requirements that a hypothesis or a theory must fulfill in order to be considered as genuinely and satisfactorily scientific –and those requirements are formal and pragmatic, not empirical. I suggested that some basic requirements of this sort are that a hypothesis must be non-parasitic and non-superfluous, and that a hypothesis must be plausible. If a theory contains statements that do not meet these constraints, it cannot play the game of science. This suggests that the non-empirical flaw I referred to with respect to the LPT is not as basic as the ones just mentioned, for the LPT is a theory that is clearly scientific enough to play the game of science. From a historical point of view, it is totally obvious that Lorentz's theory was considered as genuinely scientific by the community of the time, and the contributions introduced by Poincare did not change this opinion. From a conceptual point of view, even though it is true that the ether becomes a quasi-metaphysical term, it nevertheless played the role of giving the plausibility argument for the Lorentz-Fitzgerald contraction and it was also involved in the explanation of the difference between local and real time. That is, in spite of being problematic in the sense described, the ether was the ground for the constructive-dynamical explanations that Lorentz's theory and the LPT were able to provide. This role of the ether somewhat 'softens' its unphysical and superfluous status in Norton's sense.

However, the phantom-ether *is* a flaw insofar as it does carry problems regarding the epistemological foundations that one expects from a firmly constituted theory. We prefer theories which do not include quasi-metaphysical concepts or entities. Thus, one conclusion that can be extracted from all this is that an accurate account of what are the *theoreticity* conditions should include *levels*: some of them are basic in the sense that they are mandatory requirements for a theory to be scientifically considered, whereas some others play a role *within* the game of science –in the sense that they determine how firmly grounded a theory is from an epistemological point of view, for example.

Another real-life example of a non-empirical problem that this *second level* of theoreticity illustrates is given by the *instantaneous action at a distance* included in Newton's theory of gravitation. It was originally criticized and characterized as an epistemological blemish. However, its empirical success was such that the problem was dodged, and the use of instantaneous-forces-at-a-distance turned a standard way of doing science –as the continental approach to electromagnetism in the 19th century shows[3]. Finally, the epistemological problems carried by gravitation as action-at-a-distance got dissolved with the introduction of GR. I think that this example helps to clarify the problem of the ether in the context of this work. Just as action-at-a-distance, the phantasmagoric ether can be considered as a clear and objective a flaw in the theory. Therefore, one could use it as a

not two formulations of the same theory in which one of them has additional superfluous structure. Even if that structure is cut out, a substantial difference remains: the geometry of space-time.

[3] See Harman 1982, chapter IV (especially 103-7).

reason to choose Einstein's theory. However, the empirical success of a theory is a more important aspect for theory evaluation. If we recall the counterfactual history exercise I proposed above –had SR been in conflict with quantum physics instead of the LPT, and had a gravitation theory been formulated that reduced the LPT as a special case–, I think it is rather clear that the problem of the ether would not have counted as a fatal flaw. Analogously to the case of Newton's theory and action-at-a-distance, the LPT theory would have been accepted *in spite* of the phantom ether.

I also stated above that *any* consideration of non-empirical features as a ground to make a decision in a case of EE is always a risky one. My counterfactual history exercise shows that if a choice is made in terms of non-empirical virtues –even though those virtues can be objectively justified–, such a decision is a risky maneuver: the flawed theory might be the more empirically successful of the pair after all.

The question is now what is the epistemological root of this risk. The general answer to this question can be extracted from one of the proposals of Laudan & Leplin (and Boyd). The UD that the EE between two theories brings along is a *contingent* feature. That is, the equivalence, and therefore also the UD, can be broken as science develops. New well-confirmed auxiliary hypothesis that were not available before could be used to obtain further predictions, and these new predictions might break the equivalence. Besides, new theories concerning different realms of phenomena might be incompatible with one of the EE theories but consistent with the other one. Any of these two possibilities can threaten a decision based on non-empirical features, for the breakdown of the UD and/or EE could be such that the more empirically successful theory is the one that contains the (second level) theoreticity flaw. Therefore, we have that certain non-empirical issues regarding theoreticity can be certainly invoked in order to establish objective evaluations of a theory, and also to provide evaluative comparisons between EE theories. They are able to show that, from a theoreticity point of view, one of them is better than the other. However, a decision between them based on some of these non-empirical features is essentially risky, insofar as the empirical UD between the theories can be broken and there is no way to know which side of the contend will prevail. There is no *a priori* reason to believe that the theory which is better from a theoreticity point of view will also be the more empirically successful –if the UD is broken, of course.

The second reason I proposed to favor Einstein's theory over its rival relied on inter-theoretical considerations. Classic electrodynamics formed part of the core of the LPT, and the rise of quantum physics showed that Maxwell's equations, in their classic interpretation, were deeply at odds with the concept of a quantum of energy. SR, on the other hand, did not presuppose any specific model of the ultimate nature of matter, so that quantum physics was not incompatible with it.

I think it is rather clear that this view fits quite neatly with Richard Boyd's argument. As I showed in the first chapter, Boyd states that given two EE theories the inter-theoretic connections of one of them might be problematic, whereas the other one stands in coherence with the rest of the available scientific knowledge. In the particular case of the Einstein *vs.*

Lorentz competition, the temporal variation of the background knowledge was relevant. The conflict between classic electrodynamics and quantum mechanics became clear enough ca. 1908[4], i.e., around three years after Lorentz's theory and SR were formulated and presented. This means that there was a period in which the inter-theoretic connections of the theories could not operate as a criterion to ground a choice –however, these matters were more complicated, for as we saw, it took some time for the scientific community to understand clearly why the theories were different and rivals.

The second reason that I proposed works because the inter-theoretic relationships of scientific theories are a relevant feature when it comes to evaluate the empirical support they have, and because the corpus of background knowledge varies with time, and it can vary in a way such that the resulting inter-theoretic connections may break the UD between two EE theories –even though they remain predictively equivalent. As I stated in the first chapter, just like EE, the UD of the empirical grounds to choose a theory in an equivalent pair is a time-indexed feature. This is exactly what happened in the case of Einstein *vs.* Lorentz. There was a brief period when the choice to be made between the theories *was* empirically underdetermined, so that the only kind of reasons that could be invoked to make a decision were non-empirical –with all the epistemic risks that such decisions carry, as I showed above. However, with the rising of quantum physics, indirect empirical support became available to reject the LPT, even though the theories remained predictively equivalent –and, insofar as I can see, they still are.

The third reason to make a choice between the LPT and SR that I proposed is given by the inclusion of the latter within GR as a special case. Since GR has its own confirming predictions –which SR is not able to entail–, those confirming instances also work as support for SR. On the other hand, even though the LPT has the exact same mathematical structure as SR and is thus *mathematically* derivable from GR, the physical meaning of GR –in its *standard-geometric interpretation*– is incompatible with the physical meaning of the LPT. In GR, the expression $\delta \int_C \sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} g_{ij} dx^i dx^j} = 0$ represents a geodesic in a curved space-time, and consequently –since Minkowskian space-time describes either an empty space-time or an infinitesimal piece of a curved space-time– the expression $\delta \int_C \sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} \delta_{ij} dx^i dx^j} = 0$ describes a geodesic in a Minkowskian space-time –both in the special and in the general theories, of course. Even though the invariant interval $\Delta \boldsymbol{s}^2$ is an element contained in the mathematical structure of the LPT, its physical meaning is not a geometrical one; rather, the invariant is nothing but a mathematical nicety which is the outcome of the dynamical-compensating effects that lead to the Poincarean version of the principle of relativity. Furthermore, the fact that the LPT states that the space-time it describes has a Newtonian metric implies that it is incompatible with GR –for this theory asserts that the metric of

[4] For the *whole* scientific community, I mean; for some scientists were aware of the deep problem ever since Planck's revolutionary paper of 1900 –Einstein and Planck himself, for example.

space-time is given either by $\sqrt{\sum_{i=0}^{3}\sum_{j=0}^{3} g_{ij}dx^i dx^j}$ or by $\sqrt{\sum_{i=0}^{3}\sum_{j=0}^{3} \delta_{ij}dx^i dx^j}$, which define, of course, non-Newtonian metrics.

Before offering a general evaluation of the issue, I will detach my position from Elie Zahar's. This author claims that the main reason why Einstein defeated Lorentz was because both the special and the general theory formed part of a *research program*, in the Lakatosian sense, which was shown to be superior to Lorentz's research program. In this context, the road from SR to the GR forms a crucial part of his argument:

> Einstein invented not a theory but a research programme with an immensely powerful heuristic. But research programmes are ultimately judged on their empirical rather than on their heuristic power. No matter how fruitful its heuristic guidelines for the construction of new theories are, the programme will not be successful if these theories are not empirically corroborated. In my view Einstein's relativity programme superseded Lorentz's in the empirical sense in 1915 with its explanation of the precession of Mercury's perihelion. This explanation requires the *general* theory. […]
> My claim that Einstein's programme superseded Lorentz's with the explanation of the perihelion of Mercury raises two difficulties. *First*, since I wish to claim this as a success for the *whole* relativistic programme, I have to establish a continuity between the special and the general theories. *Secondly*, since the behavior of Mercury was well-known, I shall have to show, in line with my definition of empirical support, that the Mercury prediction was an unexpected consequence of the general theory.[5]

In his paper Zahar offers an account of both the difficulties he mentions which, I think, does not work. It is quite correct to claim that there is a continuity between the special and general theories. However, the description of the continuity that Zahar offers is given under the model of a Lakatosian research program, for he argues that the transit from SR to GR can be described accurately by this model. This is a very debatable view. The main tenets that, Zahar argues, form the core part of Einstein's research program are *i)* the quest for general covariance as a generalization of Lorentz-invariance and the principle of relativity; and *ii)* to remove any kinds of 'asymmetries which are not inherent in the phenomena' –in the case of the gravitational theory, the asymmetries involved in the different meaning of inertial and gravitational mass in spite of them being observationally identical. I think that *i)* is problematic. Recent historiography on Einstein's path to GR shows that the quest for general covariance was a constraint that he considered at first, that later he abandoned and argued as non-achievable via the famous *hole argument*; and that he finally resumed. However, the general covariance he achieved in his field equations was not a fulfillment of the goal to extend the principle of relativity. It is rather clear, though, that the quest of general covariance did play an essential role in the continuity of Einstein's path; but all the zigzags related to it make it dubious that its role can be fully grasped by Lakatos' model.

The second part of Zahar's argument is, I think, fatally flawed. Once again, the model of a research program requires him to be able to show that Einstein's program became victorious because it led to *novel* successful predictions than its rival could not make. The definition of 'novelty' that Zahar uses is simply that a prediction is novel if –even though the

[5] Zahar 1973, 249-50.

predicted fact was already known– its prediction was not a specific goal in the formulation of the theory. That is, for Zahar to be right, the correct prediction of the behavior of Mercury's perihelion cannot be a goal that Einstein consciously looked for when he created his theory. Zahar's attempt is fruitless because recent historiography on the issue has shown that a better explanation of Mercury's gravitational behavior than the one that Newton's theory offered was indeed an explicit goal and constraint that Einstein followed in the formulation of GR. In this case, however, we cannot totally condemn Zahar, for this historical information was not available when he wrote his paper[6]. Anyway, this is yet another argument to show that the rationale and continuity that connects the special and the general theories cannot be accurately encompassed by the model of a Lakatosian research program.

This detachment of my position from Zahar's is clarifying when it comes to connect my view with Laudan & Leplin's solution. Recall that these authors state that the class of statements that can provide empirical support for a theory is not equivalent to the class of its empirical consequences. One of the instances of this general view consists in that if a theory is mathematically and conceptually included in a more general one, then the empirical support of the latter flows to the former –including the confirmation given by predictions that the less general theory cannot entail on its own. Moreover, this feature of the dynamics of confirmation can play a role within the problem of UD and EE. If one of the empirically equivalent theories is encompassed by a more general and well-confirmed one, and if its rival is not, then the decision to be made between the theories is not empirically underdetermined anymore. This is exactly what happened in the Einstein *vs.* Lorentz case. SR was encompassed by GR, and GR has empirical consequences which confirm both of them, but that the special theory is not capable to entail on its own –namely, the gravitation of light and the advance of Mercury's perihelion[7]. Therefore, the evidential support that these empirical phenomena gave to GR flowed to SR. I also showed that the LPT is incompatible with GR –in its geometrical interpretation– and thus none of the evidential support of the latter can flow to the former. Consequently, SR's empirical support is larger than the support for the LPT.

I said that the detachment of my position from Zahar's clarifies the fruitfulness of the connection of my view with Laudan & Leplin's model. By this I simply meant that Laudan &

[6] Walter Isaacson, in his biography on Einstein, reports: "Einstein hoped that his new theory of relativity [the *Entwurf*], when its gravitational field equations were applied to the sun, would explain Mercury's orbit. Unfortunately, after a lot of calculations and corrected mistakes, he and Besso came up with a value of 18 seconds of an arc per century for how far Mercury's perihelion should stray, which was not even halfway correct. The poor result convinced Einstein no to publish the Mercury calculations". Isaacson 2007, 199. The failure to explain Mercury's orbit, along with the lack of general covariance and the related fact that the theory did not consider rotation as relative motion, were the motives why Einstein abandoned the *Entwurf* and pursued a theory which did accomplish general covariance and the right explanation of Mercury's perihelion. This information about the period 1912-1915 of Einstein's road to GR became available only with the edition and publication of Einstein's collected papers, which occurred after Zahar's paper appeared.

[7] One could also add gravitational red-shift and gravitational time-dilation. However, to derive these predictions one does not need the full-blown general theory. It is enough to add the principle of equivalence to SR to obtain them. The gravitation of light is also obtainable in this way, however, the correct value of this effect requires to consider the participation of the curvature of space-time, and this can only be done by means of GR fully formulated.

Leplin's scheme is such that the view on the dynamics of confirmation they endorse does not require the previous acceptance of any specific model of the rationality and development of science. This is rather clear if one considers the case of Einstein *vs.* Lorentz. To claim that Einstein's theory is more supported than Lorentz's by means of the role that GR plays in the competition does not require to argue that Einstein's and Lorentz's work can be described by Lakatos model of a research program, or by a Kuhnian paradigm, or what have you. It is simply the outcome of a subtle analysis of the dynamics of the empirical confirmation of scientific theories.

So far, so good, but what about Sorin Bangu's objection? Let us recall that he argues that nothing precludes that another general theory to encompass the other theory in the equivalent pair may be formulated; and that nothing precludes such general theory to be able to entail the empirical consequences that were used to break the UD of the choice. In the case of Einstein *vs.* Lorentz, this would mean that nothing prevents that an alternative general theory of gravitation may be formulated which encompasses the LPT, and that it may be EE to GR –so that it would be able to entail the light-bending effect and the correct description of Mercury's perihelion (if the general theories are not EE, the decision could be made in terms of their different degrees of empirical success).

I stated in the first chapter that Bangu's objection, rather than a denial of Laudan & Leplin's view, is a source of clarification of its real scope. It is not an *algorithmic* criticism, for it does not show that an alternative general theory can be formulated by following a certain procedure, but only that such theory is logically possible. That possible theory has not been formulated. However, Hendrik Lorentz himself made some important contributions and proposed some interpretations regarding GR, and this fact could lead one to think that he was close to offer something like an alternative gravitational theory to encompass his own alternative to SR. I will briefly refer to Lorentz's work on GR in order to show that nothing like that really happened[8].

Lorentz's main contributions on the subject were two. First, he dealt with the quest for general covariance. Einstein imposed himself this goal as a generalization of the principle of relativity to accelerated motion, in the sense that no privileged frames of reference are justified. However, the *Entwurf* theory he created together with Marcel Grossmann in 1913 did not achieve this goal completely, for the equations they formulated were covariant only under a restricted set of linear coordinate transformations. Einstein offered the famous *hole argument* as a justification of the failure to achieve *general* covariance[9]. Lorentz's view on the

---

[8] On this issue I closely follow Illy 1989 and Kox 1988.

[9] "The argument, the first version of which dates from 1913, runs as follows. Consider a finite space-time region Σ, in which no material processes take place, so that the physical happenings within Σ are fully determined by the quantities $g_{\mu\nu}$. In the coordinate system K theses quantities are given as functions of x_a; symbolically, $g_{\mu\nu} = \mathrm{G}(x_a)$. Introduce a new coordinate system K', which coincides with K outside Σ, but deviates from it inside this region, in such a way that the corresponding field $g'_{\mu\nu}$ and its derivatives are everywhere continuous. It may be written as $g'_{\mu\nu} = \mathrm{G}'(x'_a)$. If in G' the argument x'_a is replaced by x_a, a new gravitational field is created that differs from the original one. in the case of generally covariant field equations, both $\mathrm{G}(x_a)$ and $\mathrm{G}'(x_a)$ are solutions of the field equations with respect to K; they describe the same physical situation but are different inside Σ (they coincide on its boundary). Thus in the case of generally covariant field equations the source term (the metrical energy-

matter was that it is always possible to pick a privileged frame of reference, not only on the basis of the mathematical simplicity of the formulas that govern such a frame, but also on physical grounds. He offered an example: a frame in which the expression for Newton's second law does not include terms that refer to centrifugal and Coriolis forces is privileged, mathematically and physically, with respect to one in which those terms do appear. Lorentz argued that something similar occurred in Einstein and Grossmann's first version of GR. Since the theory was not fully covariant, the choice of a privileged frame was possible both mathematically and physically. The target of Lorentz view was rather clear: he simply wanted to make room in the theory for the introduction of an ether-rest frame.

In November 1915 Einstein published his final version of the theory, which was indeed generally covariant. Lorentz's first reaction was somehow stubborn: he wrote a letter to Einstein, dated January 1916, in which he offered an argument that was equivalent to the 'hole problem'. However, he finally accepted Einstein's way out of the difficulty. As Kox poses it, "the only essential elements in physics are coincidences in space-time; coordinates are of secondary importance. The gravitational field does not have to be uniquely determined, as long as all coincidences, such as the formation of a black spot at a certain point on a photographic plate, are described correctly"[10].

After getting convinced of this view, Lorentz set himself to a coordinate-free formulation of the theory, his second main contribution. In the construction of the coordinate-free presentation of GR, he clearly and explicitly interpreted the theory in a thorough geometric sense; he wrote for instance: "I shall conclude now by remarking that, as an immediate consequence of Hamilton's principle, the world-line of a material point which is acted only by a given gravitation field, will be a geodetic line"[11]. These words are relevant here because they are totally clear in showing that Lorentz understood GR in way such that the expression $\delta \int_C \sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} g_{ij} dx^i dx^j} = 0$ represents a geodesic in space-time, and that $\sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} g_{ij} dx^i dx^j}$ describes its metric. Consequently, he also should accept $\sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} \delta_{ij} dx^i dx^j}$ as describing the metric of a Minkowskian space-time. However, and even though Lorentz got convinced of Einstein's view on the achievability of general covariance, he remained attached to an interpretation of the theory in which the ether was needed. Einstein somehow accepted Lorentz's reading in which the general theory included *a sort* of ether, but not quite the one that Lorentz conceived[12]. What Einstein accepted that could be called *ether* in the context of GR was simply the metric field $g_{\mu\nu}$ –that represents the metric of space-time.

momentum tensor) does not uniquely determine the gravitational field. Einstein's (incorrect) conclusion (and justification of the failure of the field equations he has derived to be covariant) is that covariant fields are not allowed. One has to restrict oneself to a limited set of coordinate transformations, determined by the demand that the gravitational field is uniquely fixed by the energy-momentum tensor". Kox 1988, 69.

[10] Kox 1988, 73.

[11] From *On Einstein's Theory of Gravitation*, quoted in Illy 1989, 264.

[12] Einstein presented this view in a lecture he gave in Leiden in October 1920.

Summarizing, Lorentz's understanding of GR accepted its full geometrical significance –with $\delta \int_C \sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} g_{ij} dx^i dx^j} = 0$ representing a geodesic in space-time–; however, his interpretation included an ether –meaning not only *space-time and its metric field*, but also its classical significance as *the medium in which electromagnetic waves propagate*. However, even though empty space does have physical properties in the context of GR and can so be called *ether* in Einstein's sense, this ether cannot be posed as the ground to define a physically privileged reference frame with respect to which motion can be detected, as Lorentz's believed[13]. This is relevant for our subject in the sense that Lorentz's interpretation of GR was not able to make room for the *Lorentzian* ether that is essentially required if one is to make a case for the physical connection between the LPT and GR. Moreover, as I said above, to accept that $\sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} g_{ij} dx^i dx^j}$ expresses the metric of a curved space-time –as Lorentz did in his coordinate-free formulation of the theory– entails to accept that $\sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} \delta_{ij} dx^i dx^j}$ describes the metric of an infinitesimal part of that space-time or the metric of an empty space-time. If this is the case, it also follows that in the space-time or *portion* of space-time described by $\sqrt{\sum_{i=0}^{3} \sum_{j=0}^{3} \delta_{ij} dx^i dx^j}$ there is no distinction between *real* and *local* time, and that simultaneity between events is a frame-relative relation. This in turn means that the interpretation Lorentz made of the general theory was such that he was obliged to conceive the constant factor $\mathbf{\Delta s}^2$ as an expression of the metric of space-time, and not only as the consequence of the conspiracy of dynamical effects. However, until the end of

[13] "The problem of general covariance might have been settled, but Lorentz's ideas about the existence of an ether had not been shaken. For him, admitting generally covariant coordinate transformations did not mean that all coordinate systems were fully equivalent. The possibility always remained to choose a preferred coordinate system, which one might think of as being connected to the 'ether'. In a letter to Einstein, written in June 1916, Lorentz clearly states his point of view. He starts by describing a 'fictional' experiment: in two closed wires that run around the earth along the equator electromagnetic waves are generated in such a way that the waves in the two wires run in opposite directions. In a coordinate system fixed to the earth the waves propagate with different speeds in the two wires; in a system in which the speeds are equal the earth performs a rotation. A convenient way to describe this phenomenon, Lorentz points out, is to introduce an ether as carrier of the waves. he then goes on:

> "If we adopt this standpoint, we may say that the experiment has shown us the motion of the earth relative to the ether. If, then, we have thereby acknowledged the possibility of establishing a relative *rotation*, we should not reject in advance the possibility of also obtaining a relative *translation*, i.e., we should not set up the basic principle of relativity as a *postulate*. We would need, rather, to seek the answer to the question in the observations".

According to Lorentz , the relativity principle is a hypothesis, framed on the basis of experimental results, and always open to refutation. […]
Not surprinsingly, Einstein was not convinced by Lorentz's reasoning. In his reply he admits that the general theory of relativity is closer to an ether hypothesis than the special theory. But the 'ether' he refers to is the metric field, which is something different from the immobile 'substantial' ether Lorentz has in mind. As a consequence, one can distinguish between accelerated and non-accelerated motion: in a part of space where $g_{\mu\nu}$ = constant, a linear coordinate transformation (corresponding to non-accelerated motion) has no influence on $g_{\mu\nu}$, whereas non-linear transformations (accelerated motion) change $g_{\mu\nu}$. Thus non-accelerated motion produces no changes in the gravitational field and cannot be detected". Kox 1988, 73-5.

his life Lorentz remained attached to classical concepts of space and time, a stance which, I think, is inconsistent with the standard-geometric interpretation of GR:

> Lorentz opposes Einstein's views reservedly but obstinately. "Even though I do not wish to enter into a dispute about words", he continues, "I am inclined to associate with 'space' none but purely geometrical concepts, and to find 'vacuum' a bit too empty". Lorentz prefers the hypothesis of motionless ether to relativity "since this does not force us to adopt such radical changes, and enables us, for example, to continue to speak of true time and of true simultaneity"[14].

Lorentz, as we saw, got deeply interested and involved with the development of GR, and he certainly tried to offer an interpretation of it close to his basic views on the ether. However, such interpretation was not enough as to make space for his own ether-theory to be included as a special case of the gravitational theory. Therefore, the possibility that Sorin Bangu notes as *possible*, has not been instantiated in the Lorentz *vs.* Einstein competition. The link between GR and SR provided the latter an evidential support that the LPT cannot have, so that the UD of the choice got broken –and it remains so, for no other alternative to GR that encompasses the LPT has been formulated.

I said that from the standpoint of the *standard-geometric* interpretation of GR it is not possible to encompass the LPT in it. I made that proviso because, at least in principle, it is conceivable to interpret GR in an *alternative-dynamical* way in which it becomes coherent with the LPT. In a nutshell, the trick is done by understanding GR as a theory in *Minkowski space-time*, so that the real geodesics are the Minkowskian ones. Gravitation, therefore, is interpreted as a force rather than as the interaction between matter and the metric field. The LPT can then be included by interpreting Minkowski space-time in the dynamical way, that is, not *really* as a space-time, but as the mathematical description of the conspiracy of dynamical effects. I said that this is a *conceivable* way to interpret GR, but this does not mean right away that by so doing one gets a genuinely scientific *theory*. That is, the alternative-dynamical interpretation, at least at face value, is not able to accomplish important

[14] Illy 1989, 273. The quote of Lorentz is from his 1913 *Nieuwe richtingen in de natuurkunde*. Illy also quotes and comments at length two other works by Lorentz from 1915 in which he defends a similar position. That is, Lorentz remained committed to a classical conception of space and time even *after* he knew GR. The inconsistency that I think underlies Loretz's late years position was then that he understood GR with all its geometrical tenets about space-time, but at the same time he thought that a physical distinction between real and local time exists and that simultaneity is an absolute concept. On Lorentz's attitude towards the ether and the relativity theory in his later years, Kox comments: "In the following years Lorentz inspired several of his students and former students to work in the field of GR and made some further contributions himself. Though he kept insisting on the existence of an ether, he was not dogmatic about it and on many occasions expressed his admiration for Einstein's achievements. His attitude is very clearly illustrated by the statement with which he concluded a series of lectures given at the California Institute of Technology in 1922:

> As to the ether (to return to it once more), though the conception of it has certain advantages, it must be admitted that if Einstein had maintained it he certainly would not have given us his theory, and so we are very grateful to him for not having gone along the old-fashioned roads.

Why Lorentz kept insisting on the existence of an ether is a question not easy to answer. His attitude may show a certain conservatism, perhaps even stubbornness. But it should be kept in mind that from the earliest years of Lorentz's career the concept of an ether had played a fundamental role in his work on electromagnetic theory […]. The concept of the ether must have been very dear to Lorentz, and it does not seem to be out of character for a man like him to remain true to it to the very end". Kox 1988, 75-6.

theoreticity requirements. For example, the status of Minkowski space-time becomes rather awkward: would it be considered just as a mathematical tool or, at least in certain cases, also as having some substantial geometric meaning? On the other hand, the principle of equivalence, a central part of the theory, gets affected, for the equivalence between inertial and gravitational mass remains unexplained. More generally, the fact that this approach has not been taken seriously in the scientific community is at least an indicator that the resulting theory would not be plausible enough as to become a real alternative to the standard GR. Coming back to Bangu's remarks, this logical possibility of conceiving a dynamical version of GR that might encompass the LPT also shows that Laudan & Leplin's solution is a contingent one: *as far as we know*, the alternative interpretation of Einstein's gravitation theory is not plausible enough, but in a different scientific scenario it might become so. The fact that a dynamical interpretation of GR is conceivable does not mean that there is an alternative *theory* of gravitation that restores the UD of the choice between the LPT and SR.

Summarizing, the *most general* conclusion that can be extracted from this work is that the solution to the problem of EE and UD of theory choice that I proposed in the first chapter works. The case of Einstein *vs.* Lorentz gets decided –conceptually speaking– by instances of the general views that Laudan, Leplin and Boyd introduced. I think that the analysis of the real-life case I offered here also illustrates that the qualification I proposed for their general model is correct: I argued that a close consideration of the criticisms directed to Laudan & Leplin (and Boyd) help to clarify the true scope of their solution. They argue that "The thesis of underdetermination, at least in so far as it is founded on presumptions about the possibility of empirical equivalence for theories, stands refuted"[15], and I said that they went beyond the mark by so doing. I think that the solution they offer is not ultimate in the sense that the problem of UD as a consequence of EE between theories remains a possible scenario in real science. Even though all proposed algorithms to formulate, given a certain theory, an EE rival are ineffective; and even though the conditions of theoreticity make it unlikely that predictive equivalence may happen in real science, it *can* happen anyway. It is quite obvious that the case of Einstein *vs.* Lorentz is indeed an example. Moreover, the solutions that Laudan, Leplin and Boyd propose to solve the problem even if the equivalence is the case, are also such that they provide a *contingent* way out of the problem. The development of science *might* be such that the available auxiliary hypotheses *may* break the EE between the theories, or such that the different intertheoretic connections with the corpus of knowledge between the theories *may* break the UD of the decision –even though the equivalence remains–; or a more general and well-confirmed theory that encompasses only one of the competitors *might* be formulated. All these possibilities *might* be the case, but they also might not be the case! Quantum theory might have been such that it was compatible both with SR and the LPT, and Lorentz –or another scientist– might have been successful in providing a version of GR which was coherent with the LPT. This obviously means that UD of theory

[15] Laudan & Leplin 1991, 466.

choice between EE rivals *can* be a real problem in real science. This is the sense in which I stated that the solutions offered are not ultimate.

However, the solutions provided are grounded on a careful examination and clarification of the dynamics of confirmation of theories and of certain procedures which are a common and essential part of scientific practice. The auxiliary hypotheses available change along time because scientists create more and better theories, and theories get encompassed in more general ones because scientists are always looking for bigger theories that explain more. This means that the features which Laudan, Leplin and Boyd show to be capable to break the problem are the outcome of the regular practice of science. No special methods or special models of scientific rationality are necessary. Actually, the case of Einstein *vs.* Lorentz got solved not because scientists set themselves to the specific task of solving the UD or the EE between the theories involved, it got solved because quantum theory and GR were formulated –and these theories were not formulated *in order to* solve the Einstein *vs.* Lorentz case, of course. Moreover, one does not need to be Lakatosian, Kuhnian, Popperian or what have you in order to be able to conceive a way out of the problem. What scientists do is –or more precisely, might be– enough.

At this point a question concerning the epistemological status of the reasons to make the choice I invoked comes up: are they truth-leading or merely pragmatic? I stated in the first chapter that my standpoint is not committed to a realist conception of science. The reasons favoring Einstein's theory are not necessarily an indication that SR is true or that it is closer to the truth than the LPT. However, the fact that I deny that the results of this work could be used, at least at face value, as an argument for a realist conception of science does not mean that I am attacking such a position. I am not arguing that the rationality of the reasons I proposed is *merely* pragmatic, for example. My position on the ultimate epistemological status of such reasons is *neutral*, in the sense that I think that their rationality would be accepted both by realists and anti-realists, for I am not dealing with the *general semantics* of scientific theories.

The first reason I invoked is *non-empirical* and is grounded in a general conception of the nature and goals of science. In this sense, it is a pragmatically grounded reason. However, it is also true that the nature and goals of science, even though historical and pragmatic issues, are grounded on epistemic considerations. The reason why we do not want (quasi) metaphysical entities and superfluous terms in scientific theories is epistemic. On the other hand, the other two reasons that I offered are related to the way in which the *evidential confirmation* of theories works –they count as *empirically grounded* reasons to decide: SR has more evidential support than the LPT. However, to claim that the inter-theoretic connections of a given theory can work as (indirect) empirical support for its acceptance or as (indirect) empirical evidence for its rejection does not mean that a realist or anti-realist conception of the ultimate status of empirical (dis)confirmation is being defended or attacked –the realist conceives that empirical confirmation is an indication of the (approximate) truth of a theory, the anti-realist does not. In a word, the inter-theoretical connections of a theory –as a feature which is relevant for its evidential support– can be considered as rational and objective grounds to make a choice *both by realists and anti-realists*[16] –they will disagree with respect to the ultimate epistemological status they assign to evidential support, of course.

More generally, I do not think that the results of this work can be used, *at face value*, as arguments in the debate between realists and anti-realists[17]. That the reference to quasi-metaphysical entities and terms in science is not welcome, and that inter-theoretic connections are relevant regarding the empirical support of theories are, I think, features of science which are rationally justified both for a realist and an anti-realist. In simple words, the main result of this thesis is that UD and EE between theories is more a *scientific* problem than an *epistemological* one.

[16] Note that the inter-theoretic connections of theories, and the related evidential import, are features which that sense relative to specific state of scientific knowledge. *So far as we know*, Einstein's theory is better supported by evidence than Lorentz's theory. However, the latter has not been directly refuted, so that a scenario in which the situation regarding empirical confirmation gets reverted is possible. This is yet another indication of the general conclusion that I endorse: the solution of the problem of EE and UD is more a matter of *science* than a matter of *epistemology*.

[17] It is true that an important anti-realist argument against realism is based on EE and UD, and I proposed a *solution* of the problem that makes possible a rational choice between the theories based on evidential grounds. However, it is also true that this solution leaves open the possibility that EE and UD can be the case, and the anti-realist could argue that is just that very possibility what he needs in order to make his point against the realist.

BIBLIOGRAPHY

Achinstein, Peter (ed.) (1983). *The Concept of Evidence*. Oxford University Press.

Adlam, Emily (2011). "Poincare and special relativity", in ArXiv Cornell University Library. Available at http://arxiv.org/pdf/1112.3175.pdf.

Bangu, Sorin (2006). "Underdetermination and the argument from indirect confirmation". *Ratio* **19**: 269-277.

Bell, John. S. (1987). "How to teach special relativity". In: Bell, John S., *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press, 1987. Chapter 9, pp. 67–80.

Bergström, Lars (1993). "Quine, underdetermination and skepticism". *The Journal of Philosophy* **90**: 331-358.

Bird, Alexander (2007). "Underdetermination and evidence". Chapter V of Monton, B. (ed.) *Images of Empiricism: essays on science and stances, with a reply from Bas C. van Fraassen*. Oxford University Press.

Bonk, Thomas (2008). *Underdetermination: an essay on evidence and the limits of natural knowledge*. Dordrecht: Springer.

Boyd, Richard (1973). "Realism, underdetermination, and a causal theory of evidence". *Noûs* **7**: 1-12.

Brown, Harvey (2001). "The origins of length contraction: I. The FitzGerald-Lorentz deformation hypothesis". *American Journal of Physics* **69**: 1044-1054.

Brown, Harvey (2003). "Michelson, FitzGerald and Lorentz: the origins of special relativity revisited". *Bulletin de la Société des Sciences et des Lettres de Lódz, Volume LIII; Série: Reserches sur les Déformations*, Volume XXXIX; 23-35.

Brown, Harvey (2005). *Physical Relativity: space-time structure from a dynamical perspective*. Oxford University Press.

Brown, Harvey & Pooley, Oliver (2000). "The origin of the spacetime metric: Bell's 'Lorentzian pedagogy' and its significance in general relativity", in Callender, C. & Huggett, N., *Physics Meets Philosophy at the Planck's Scale*. Cambridge University Press, 2000.

Brown, Harvey & Pooley, Oliver (2006). "Minkowski space-time: a glorious non-entity"; in *The Ontology of Spacetime*, Dennis Dieks (ed.), Elsevier B. V. (2006), 67-89.

Brush, Stephen (1999). "Why was relativity accepted?". *Physics in Perspective* **1**: 184-214.

Bunge, Mario (1961). "The weight of simplicity in the construction and assaying of scientific theories". *Philosophy of Science* **28**: 120-141.

Busch, Jacob (2009). "Underdetermination and rational choice of theories". *Philosophia* **37**: 55-65.

Carrier, Martin (2011). "Underdetermination as an epistemological test tube: expounding hidden values of the scientific community". *Synthese* **180**: 189-204.

Clendinnen, F. John (1989). "Realism and the underdetermination of theory". *Synthese* **81**: 63-90.

Cuvaj, Camilo (1968). "Henri Poincare's mathematical contributions to relativity and the Poincare stresses". *American Journal of Physics* **36**: 1102-1113.

Darrigol, Olivier (1994). "The electron theories of Larmor and Lorentz: a comparative study". *Historical Studies in the Physical and Biological Sciences* **24**: 265–336.

Darrigol, Olivier (1996). "The electrodynamic origins of relativity theory". *Historical Studies in the Physical Sciences* **26**: 241-312.

Darrigol, Olivier (2005). "The genesis of the theory of relativity", in T. Damour, O. Darrigol, B. Duplantier, V. Rivasseau (eds.), *Einstein 1905-2005: Poincaré seminar 2005*. Basel: Birkhäuser, 2006, 1-31.

De Regt, Henk W. and Dieks, Dennis (2005). "A contextual approach to scientific understanding". *Synthese* **144**: 137-170.

Dieks, Dennis. "Bottom-up versus top-down: the plurality of explanation and understanding in physics", in H. de Regt, S. Leonelli, K. Eigner (eds), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh: University of Pittsburgh Press 2009. Available at http://www.projects.science.uu.nl/igg/dieks/Understandingrevised.pdf.

Dorling, Jon (1968). "Length contraction and clock synchronization: the empirical equivalence of Einsteinian and Lorentzian theories". *The British Journal for the Philosophy of Science* **19**: 67-69.

Douven, Igor (2000). "The anti-realist argument for underdetermination". *The Philosophical Quarterly* **50**: 371-375.

Duhem, Pierre (1906). "Physical theory and experiment" (chapter VI of *The Aim and Structure of Physical Theory*), in Harding, Sandra (ed.) (1976).

Einstein, Albert (1905a). "Zur Elektrodynamik bewegter Körper". *Annalen der Physik* **17**: 891-921. Reprinted in Lorentz et al. 1952.

Einstein, Albert (1905b). "Ist die Trägheit eines Körpers von seinem Energieinhalt abhänging?" *Annalen der Physik* **18**: 639–641. Reprinted in Lorentz et al. 1952.

Einstein, A. (1919), "My theory", *The London Times*, November 28, 13. Reprinted as "What is the theory of relativity?" in Einstein (1954), 227-232.

Einstein, Albert (1906). "Das Prinzip von der Erhaltung der Schwerpunktsbewegung und die Trägheit der Energie". *Annalen der Physik* **20**: 627–633. Reprinted in Stachel, J., Cassidy, D., Renn, J., and Schulmann, R. (eds.). *The Collected Papers of Albert Einstein, Vol. 2, The Swiss years: writings, 1900-1909*. Princeton: Princeton University Press, 1989.

Einstein, Albert (1920). *Äther und Relativitätstheorie. Rede gehalten am 5. Mai 1920 an der Reichs-Universität zu Leiden*. Berlin: Springer. English translation: "Ether and the theory of relativity". In G. B. Jeffery and W. Perrett (transl.), *Sidelights on Relativity*. London: Methuen; New York: E. P. Dutton, 1922, 1–24.

Einstein, A. (1954), *Ideas and Opinions*. New York: Crown Publishers.

Einstein, Albert (1949). "Autobiographical notes", in Schilpp 1949, 1–95.

French, Steven (2011). "Metaphysical underdetermination: why worry?" *Synthese* **180**: 205-221.

Galison, Peter L. (1979). "Minkowski's space-time: from visual thinking to the absolute world". *Historical Studies in the Physical Sciences* **10**: 85–121.

Giedymin, Jerzy (1982). *Science and Convention: essays on Henri Poincare's philosophy of science and the conventionalist tradition*. Pergamon Press.

Goldberg, Stanley (1967). “Henri Poincare and Einstein’s theory of relativity”. *American Journal of Physics* **35**: 934–944.

Goldberg, Stanley (1969). “The Lorentz theory of electrons and Einstein’s theory of relativity”. *American Journal of Physics* **37**: 498–513.

Grünbaum, Adolf (1959). “The falsifiability of the Lorentz-FitzGerald contraction hypothesis”. *The British Journal for the Philosophy of Science* **37**: 48–50.

Grünbaum, Adolf (1960). “The Duhemian argument”. *Philosophy of Science* **27**. Reprinted in Harding, Sandra (ed.) 1976.

Grünbaum, Adolf (1976). “*Ad hoc* auxiliary hypotheses and falsificationism”. *The British Journal for the Philosophy of Science* **27**: 329–362.

Harding, Sandra (1976). *Can Theories be Refuted? Essays on the Duhem-Quine thesis.* Dordrecht: Reidel.

Harman, P. M. (1982). *Energy, Force and Matter: the conceptual development of nineteenth century physics.* Oxford University Press.

Hempel, Carl (1965). “Studies in the logic of confirmation”. Reprinted as chapter 1 in Achinstein (ed.) 1983.

Hirosige, Tetu (1976). “The ether problem, the mechanistic worldview, and the origins of the theory of relativity”. *Historical Studies in the Physical Sciences* **7**: 3–82.

Hoefer, Carl and Rosenberg, Alexander (1994). *Philosophy of Science* **61**: 592-607.

Holton, Gerald (1969). “Einstein, Michelson and the “crucial” experiment”. *Isis* **60**: 133–197. Reprinted in Holton 1988.

Holton, Gerald (1988) [1973]. *Thematic Origins of Scientific Thought: Kepler to Einstein*. Cambridge: Harvard University Press.

Hunt, Bruce (1988). “The origins of the Fitzgerald contraction”. *The British Journal for the History of Science* **21**: 67-76.

Illy, József (1981). “Revolutions in a revolution”. *Studies in History and Philosophy of Science* **12**: 173-210.

Illy, József (1989). “Einstein teaches Lorentz, Lorentz teaches Einstein: their collaboration in general relativity, 1913–1920”. *Archive for History of the Exact Sciences* **39**: 247-288.

Isaacson, Walter (2007). *Einstein: his life and universe*. New York: Simon & Schuster.

Ives, Herbert (1952). “Derivation of the mass-energy relation”. *Journal of the Optical Society of America* **42**: 540-543

Janssen, Michel (1995). *A Comparison Between Lorentz's Ether Theory and Special Relativity in the Light of the Experiments of Trouton and Noble*. Dissertation. University of Pittsburgh, 1995. Available at http://www.mpiwg-berlin.mpg.de/litserv/diss/janssen_diss/TitleTOC.pdf.

Janssen, Michel (2002). “Reconsidering a scientific revolution: the case of Einstein versus Lorentz”. *Physics in Perspective* **4**: 421-446. Available at https://netfiles.umn.edu/users/janss011/home%20page/HALvsAE.pdf.

Janssen, Michel (2003). "The Trouton experiment, E = mc2, and a slice of Minkowski space-time"; in Abhay Ashtekar et al. (ed.), *Revisiting the Foundations of Relativistic Physics: Festschrift in honor of John Stachel*. Dordrecht: Kluwer, 2003. Available at https://netfiles.umn.edu/users/janss011/home%20page/trouton.pdf.

Janssen, Michel & Mecklenburg, Mathew (2007). "From classical to relativistic mechanics: electromagnetic models of the electron"; in: V. F. Hendricks, K. F. Jørgensen, J. Lützen, and S. A. Pedersen (eds.), *Interactions: Mathematics, Physics and Philosophy 1860–1930.* Dordrecht: Springer, 2007. Available at https://netfiles.umn.edu/users/janss011/home%20page/electron.pdf.

Janssen, Michel & Stachel, John (2004). "The optics and electrodynamics of moving bodies", preprint, Max Planck Institute for the History of Science. Available at http://www.mpiwg-berlin.mpg.de/Preprints/P265.PDF.

Katzir, Shaul (2005). "Poincare's relativistic physics: its origins and nature". *Physics in Perspective* **7**: 268-292.

Kox, A. J. (1988). "Hendrik Antoon Lorentz, the ether, and the general theory of relativity". *Archive for History of Exact Sciences*. **38** (1988): 67–78.

Kragh, Helge (1999). *Quantum Generations: a history of physics in the twentieth century.* Princeton: Princeton University Press.

Kukla, Andre (1993). "Laudan, Leplin, and underdetermination". *Analysis* **53**: 1-7.

Kukla, Andre (1996). "Does every theory have empirically equivalent rivals?" *Erkenntnis* **44**: 137-166.

Kukla, Andre (1994). "Non-empirical theoretical virtues and the argument from underdetermination". *Erkenntnis* **41**: 157-170.

Kukla, Andre (2001). "Theoreticity, underdetermination, and the disregard for bizarre scientific hypotheses". *Philosophy of Science* **68**: 21-35.

Lakatos, Imre (1978). *The Methodology of Scientific Research Programmes. Philosophical Papers.* Vol. 1. Cambridge: Cambridge University Press, 1978

Laudan, Larry (1965). "Grünbaum on the Duhemian argument". *Philosophy of Science* **32**. Reprinted in Harding, Sandra (ed.) (1976).

Laudan, Larry and Leplin, Jarret (1991). "Empirical equivalence and underdetermination". *The Journal of Philosophy* **88**: 449-472.

Laudan, Larry and Leplin, Jarret (1993). "Determination underdeterred: reply to Kukla". *Analysis* **53**: 8-16.

Leplin, Jarrett (1975). "The concept of an *ad hoc* hypothesis". *Studies in History and Philosophy of Science* **5**: 309–345.

Leplin, Jarret (1997a). *A Novel Defense of Scientific Realism*. Oxford: Oxford University Press.

Leplin, Jarret (1997b). "The underdetermination of total theories". *Erkenntnis* **47**: 203-215.

Lorentz, Hendrik Antoon (1886). "Over den invloed, dien de beweging der aarde op de lichtverschijnselen uitoefent". *Koninklijke Akademie van Wetenschappen* (Amsterdam). *Afdeeling Natuurkunde. Verslagen en Mededeelingen* **2**: 297–372. French translation: "De l'influence du mouvement de la terre sur les phénomènes lumineux". *Archives Néerlandaises des Sciences Exactes et Naturelles* **21**: 103–176. This translation is reprinted in Lorentz 1934–39, Vol. 4, pp. 153–214.

Lorentz, H. A. (1892a). "La Théorie électromagnétique de Maxwell et son application aux corps mouvants". *Archives Néerlandaises des Sciences Exactes et Naturelles* **25**: 363–552. Reprinted in Lorentz 1934–39, Vol. 2, pp. 164–343.

Lorentz, H. A. (1892b). "De relatieve beweging van de aarde en den aether". *Verslagen van de gewone vergaderingen der wis– en natuurkundige afdeeling, Koninklijke Akademie van Wetenschappen te Amsterdam* **1** (1892–1893): 74–79. English translation in Lorentz 1934–39, Vol. 4, pp. 219–223.

Lorentz, H. A. (1895). *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern.* Leiden: Brill. Reprinted in Lorentz 1934–39, Vol. 5, pp. 1–138.

Lorentz, H. A. (1899). "Simplified theory of electrical and optical phenomena in moving bodies". *Proceedings of the section of sciences, Koninklijke Akademie van Wetenschappen te Amsterdam* **1**: 427–442. English version of Lorentz 1899a. Reprinted in Schaffner 1972, pp. 255–273.

Lorentz, H. A. (1902). "Théorie simplifiée des phénomènes électriques et optiques dans des corps en mouvement". *Archives Néerlandaises des Sciences Exactes et Naturelles* **7**: 64–80. Translation of Lorentz 1899a. Reprinted in Lorentz 1934–39, Vol. 5, pp. 139–155.

Lorentz, H. A. (1904a). "Weiterbildung der Maxwellschen Theorie. Elektronentheorie". *Encyclopädie der mathematischen Wissenschaften, mit Einschlussiher Anwendungen*, vol.5, *Physik*, part 2. Arnold Sommerfeld (ed.), Leipzig: Teubner, 1904–1922. Pp. 145–288.

Lorentz, H. A. (1904b). "Electromagnetische verschijnselen in een stelsel dat zich met willekeurige snelheid, kleiner dan die van het licht, beweegt". *Verslagen van de gewone vergaderingen der wis– en natuurkundige afdeeling, Koninklijke Akademie van Wetenschappen te Amsterdam* **12**: 986–1009. English translation: "Electromagnetic phenomena in systems moving with any velocity less than that of light". *Proceedings of the section of sciences, Koninklijke Akademie van Wetenschappen te Amsterdam* **6**: 809–831. Reprinted in Lorentz 1934–39, Vol. 5, pp. 172–197, and (without the final section 14) in Lorentz *et al.* 1952.

Lorentz, H. A. (1934–39). *Collected Papers*. 9 Vols. The Hague: Nijhoff.

Lorentz, H. A., Einstein, A., Minkowski, H., and Weyl, H. (1952). *The Principle of Relativity*. New York: Dover. Translation by W. Perrett and G.B. Jeffery of: Blumenthal, Otto (ed.) *Das Relativitätsprinzip: eine Sammlung von Abhandlungen*. Fourth edition. Leipzig: Teubner, 1922.

Magnus, P. D. (2003). "Underdetermination and the problem of identical rivals". *Philosophy of Science* **70**: 1256-1264.

McCormmach, Russell (1970a). "Einstein, Lorentz, and the electron theory". *Historical Studies in the Physical Sciences* **2**: 41–87.

McCormmach, Russell (1970b). "H.A. Lorentz and the electromagnetic view of nature". *Isis* **61**: 459–497.

Miller, Arthur I. (1973). "A study of Henri Poincare's *Sur la dynamique de l'électron*". *Archive for History of Exact Sciences* **10**: 207–328. Reprinted in Miller 1986.

Miller, Arthur I. (1974). "On Lorentz's methodology". *The British Journal for the Philosophy of Science* **25**: 29-45. Reprinted in Miller 1986.

Miller, Arthur I. (1980). "On some other approaches to electrodynamics in 1905", in Miller 1986.

Miller, Arthur I. (1996). "Why did Einstein not formulate special relativity in 1905?", in Greffe, J.-L. et al. *Henri Poincare: Science and Philosophy*. Ed. Akademie Verlag, Berlin, and Albert Blanchard, Paris

Miller, Arthur I. (1998) [1981]. *Albert Einstein's Special Theory of Relativity: emergence (1905) and early interpretation (1905–1911).* New York: Springer.

Miller, Arthur I. (1986). *Frontiers of Physics: 1900–1911*. Boston: Birkhäuser.

Miller, Richard (1987). *Fact and Method: explanation, confirmation and reality in the natural and the social sciences*. Princeton University Press.

Minkowski, Hermann (1909). "Raum und Zeit". *Physikalische Zeitschrift,* **10**, pp. 104–111. Reprinted in Lorentz et al. 1952.

Mormann, Thomas (1995). "Incompatible empirically equivalent theories: a structural explication". *Synthese* **103**: 204-249.

Muller, F. A (1997). "The equivalence myth of quantum mechanics –part I". *Studies in history and Philosophy of Modern Physics* **28**: 35-61.

Nersessian, Nancy J. (1984). "Aether/or: the creation of scientific concepts". *Studies in History and Philosophy of Science* **15**: 175-212.

Nersessian, Nancy J. (1986). "Why wasn't Lorentz Einstein? An examination of the scientific method of H. A. Lorentz". *Centaurus* **29**: 205–242.

Norton, John D. (2008). "Must evidence underdetermine theory?", chapter 1 in Carrier, M., Howard, D., and Kourany, J., *The Challenge of the Social and the Pressure of Practice: science and values revisited*, pp. 17-44. Pittsburgh: University of Pittsburgh Press, 2008.

Okasha, Samir (1997). "Laudan and Leplin on empirical equivalence". *The British Journal for the Philosophy of Science* **48**: 251-256.

Okasha, Samir (2002). "Underdetermination, holism and the theory/data distinction". *The Philosophical Quarterly* **52**: 303-319.

Pais, Abraham (1982). *'Subtle is the Lord...': the science and the life of Albert Einstein*. Oxford: Oxford University Press.

Park, Seungbae (2009). "Philosophical responses to underdetermination in science". *Journal for General Philosophy of Science* **40**: 115-124.

Poincare, Henri (1900a). "Sur les rapports de la physique expérimentale et de la physique mathématique". In: *Rapports Presentés au Congrès International de Physique Réuni à Paris en 1900.* Vol. 1, pp. 1–29. Paris: Gauthier–Villars. Translated as chs. 9–10 in Poincare 1952.

Poincare, Henri (1900b). "La théorie de Lorentz et le principe der réaction". In: Bosscha 1900, pp. 252–278. Reprinted in Poincare 1934–54, Vol. 9, pp. 464–488.

Poincare, Henri (1904). "L'état actuel et l'avenir de la physique mathématique". *Bulletin des Sciences Mathématiques* **28**: 302–324. Translated as chs. 7–9 in Poincare 1958.

Poincare, Henri (1906). "Sur la dynamique d'électron". *Rendiconti del Circolo Matematico di Palermo* **21**: 129–175. Reprinted in Poincare 1934–54, Vol. 9, pp. 494–550. English translation of sections 6–8 in Miller 1973.

Poincare, Henri (1934–54). *OEvres de Henri Poincare*. 11 Vols. Paris: Gauthier-Villars.

Poincare, Henri (1952a). *Science and Hypothesis*. New York: Dover.

Poincare, Henri (1952b). *Science and Method*. New York: Dover.

Poincare, Henri (1958). *The Value of Science*. New York: Dover.

Poincare, Henri (1963). *Mathematics and Science: last essays*. New York: Dover.

Psillos, Statis (1999). *Scientific Realism: how science tracks truth*. London: Routledge.

Quine, W. V. O. (1951). "Two dogmas of empiricism". *The Philosophical Review*: **60**. Reprinted in Harding, Sandra (1976).

Quine, W. V. O. (1975). "On empirically equivalent systems of the world". *Erkenntnis* **9**: 313-328.

Reichenbach, Hans (1958) [1928]. *The Philosophy of Space & Time.* New York: Dover.

Reignier, Jean (2004). "Poincare synchronization: from the local time to the Lorentz group", in *Proceedings of the Symposium Henri Poincare (Brussels, 8-9 October 2004).* International Solvay Institutes for Physics and Chemistry. Available at http://www.ulb.ac.be/sciences/ptm/pmif/ProceedingsHP/Reignier.pdf.

Rindler, Wolfgang (1991). *Introduction to Special Relativity.* Oxford: Oxford University Press.

Sarkar, Husain (2000). "Empirical equivalence and underdetermination". *International Studies in the Philosophy of Science* **14**: 187-197.

Sartori, Leo (1996). *Understanding Relativity.* University of California Press.

Schaffner, Kenneth F. (1969). "The Lorentz electron theory of relativity". *American Journal of Physics* **37**: 498-513.

Schaffner, Kenneth F. (1972). *Nineteenth Century Aether Theories*. Oxford, New York: Pergamon Press.

Schaffner, Kenneth F. (1974). "Einstein versus Lorentz: research programmes and the logic of comparative theory evaluation". *The British Journal for the Philosophy of Science*, **25**: 45–78.

Schaffner, Kenneth F. (1976). "Space and time in Lorentz, Poincare, and Einstein: divergent approaches to the discovery and development of the special theory of relativity". In: Machamer, Peter K., and Turnbull, Robert G. (eds.). *Motion and Time, Space and Matter: interrelations in the history of philosophy of science*. Ohio State University Press.

Schilpp, Paul Arthur, ed. (1949). *Albert Einstein: philosopher-scientist*. Evanston, IL: Library of Living Philosophers.

Scribner, Charles, Jr. (1964). "Henri Poincare and the principle of relativity". *American Journal of Physics* **32**: 672–678.

Sklar, Lawrence (1974). *Space, Time, and Spacetime*. Berkeley, Los Angeles, London: University of California Press.

Stachel, John (1982). "Einstein and Michelson: the context of discovery and the context of justification". *Astronomische Nachrichten* **303**, I, 47–53.

Stachel, John (2002). "What song the syrens sang?: how did Einstein discover special relativity", in Stachel, John, *Einstein from "B" to "Z"*. Boston: Birkhäuser, 2002.

Stanford, Kyle (2001). "Refusing the devil's bargain: what kind of underdetermination should we take seriously?" *Philosophy of Science* **68** (Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers): S1-S12.

Stanford, Kyle (2009). "Underdetermination of scientific theory". *Stanford Encyclopedia of Philosophy*, http://stanford.library.usyd.edu.au/entries/scientific-underdetermination/.

Szabó, László. "Lorentzian theories vs. Einsteinian special relativity –a logico-empiricist reconstruction" in András, M., Rédei, M. & Stadler, F. (eds); *The Vienna Circle in Hungary*. Springer, 2011. Available at http://philsci-archive.pitt.edu/5339/.

Torretti, Roberto (1996) [1983]. *Relativity and Geometry*. New York: Dover.

Torretti, Roberto (2003). *Relatividad y Espaciotiempo.* Santiago: RIL.

Van Dongen, Jeroen (2009). "On the role of the Michelson-Morley experiment: Einstein in Chicago". *Archive for History of Exact Sciences* **63**: 655-663.

Van Fraassen, Bas C. (1980). *The Scientific Image.* Oxford: Clarendon Press.

Walter, Scott. "Minkowski's modern world", in Petkov, V. (ed), *Minkowski Spacetime: a hundred years later*. Springer, 2010. Available at
http://www.univ-nancy2.fr/DepPhilo/walter/papers/2010mink100.pdf.

Warwick, Andrew (1995). "The sturdy protestants of science: Larmor, Trouton and the earth's Motion through the aether". Chapter 11 in Buchwald, J. Z. *Scientific Practice: theories and stories of doing physics.* The University of Chicago Press, 1995.

Whittaker, Edmund T. (1953). *A History of the Theories of Aether and Electricity*. 2 Vols. London: Nelson.

Yalçin, Ümit D. (2001). "Solutions and dissolutions of the underdetermination problem". *Noûs* **35**: 394-418.

Zahar, Elie (1973). "Why did Einstein's programme supersede Lorentz's?" *The British Journal for the Philosophy of Science* **24**: 95-123, 223-262. Reprinted in Zahar 1989.

Zahar, Elie (1977). "Mach, Einstein and the rise of modern science". *The British Journal for the Philosophy of Science* **28**: 195–213. Reprinted in Zahar 1989.

Zahar, Elie (1978). "Einstein's debt to Lorentz: a reply to Feyerabend and Miller". *The British Journal for the Philosophy of Science* **29**: 49–60. Reprinted in Zahar 1989.

Zahar, Elie (1989). *Einstein's Revolution: a study in heuristic*. La Salle, IL: Open Court.