

Einstein's Investigations of Galilean Covariant Electrodynamics prior to 1905

John D. Norton¹

Department of History and Philosophy of Science

University of Pittsburgh

jdnorton@pitt.edu

Einstein learned from the magnet and conductor thought experiments how to use field transformation laws to extend the covariance to Maxwell's electrodynamics. If he persisted in his use of this device, he would have found that the theory cleaves into two Galilean covariant parts, each with different field transformation laws. The tension between the two parts reflects a failure not mentioned by Einstein: that the relativity of motion manifested by observables in the magnet and conductor thought experiment does not extend to all observables in electrodynamics. An examination of Ritz's work shows that Einstein's early view could not have coincided with Ritz's on an emission theory of light, but only with that of a conveniently reconstructed Ritz. One Ritz-like emission theory, attributed by Pauli to Ritz, proves to be a natural extension of the Galilean covariant part of Maxwell's theory that happens also to accommodate the magnet and conductor thought experiment. Einstein's famous chasing a light beam thought experiment fails as an objection to an ether-based, electro-dynamical theory of light. However it would allow Einstein to formulate his general objections to all emission theories of light in a very sharp form. Einstein found two well known experimental results of 18th and 19th century optics compelling (Fizeau's experiment, stellar aberration), while the accomplished Michelson-Morley experiment played no memorable role. I suggest they owe their importance to their providing a direct experimental grounding for Lorentz' local time, the precursor of Einstein's relativity of simultaneity, and do it essentially independently of electro-dynamical theory. I attribute Einstein's success to his determination to implement a principle of relativity in electrodynamics, but I urge that we not invest this stubbornness with any mystical prescience.

¹ I am grateful to Diana Buchwald, Olivier Darrigol, Allen Janis, Michel Janssen, Robert Rynasiewicz and John Stachel for helpful discussion and for assistance in accessing source materials.

1. Introduction

Although we have virtually no primary sources, the historical scholarship of the last few decades has painstakingly assembled clues from many places to give us a pretty good sketch of Einstein's route to special relativity. He had a youthful interest in electrodynamics and light with no apparent skepticism about the ether. As a sixteen year old in the summer of 1895, he wrote an essay proposing experimental investigation into the state of the ether associated with an electromagnetic field.² The skepticism emerged later along with a growth of his knowledge of electrodynamics. By the end of 1901, he was writing confidently of work on a "capital paper" on the electrodynamics of moving bodies that expressed ideas on relative motion.³ Later recollections stress the guiding influence of his recognition that the electric field induced by a moving magnet has only a relative existence. His pursuit of the relativity of inertial motion led him to reject Maxwell's theory and its attendant constancy of the velocity of light with respect to the ether in favor of investigation of an emission theory, somehow akin to Ritz' later approach, in which the speed of light was a constant with respect to the emitter. These investigations proved unsatisfactory and Einstein was brought to a crisis in the apparent irreconcilability of the relativity of inertial motion and the constancy of the velocity of light demanded by Maxwell's electrodynamics. The solution suddenly came to Einstein with the recognition of the relativity of simultaneity and a mere five to six weeks was all that was needed to complete writing the paper, which was received by *Annalen der Physik* on June 30, 1905.

My understanding of this episode is framed essentially by the historical researches of John Stachel, individually and in collaboration with the editors of Volume 2 of the *Collected Papers of Albert Einstein*; and by Robert Rynasiewicz and his collaborators. See Stachel (1987, 1989), Stachel *et al.* (1989a), Rynasiewicz (2000) and Earman *et al.* (1983) and the citations therein for their debts to other scholarship. In addition to the arduous scholarship of discovering and developing our present framework, they have supplied particular insights of importance. For example, Rynasiewicz and his collaborators have pointed out that Einstein must have known of field transformations akin to the Lorentz transformation for fields years before he adopted the novel kinematics of the Lorentz transformation for space and time, so that the historical narrative must somehow account for a development from field transformation to the space and time transformations they necessitate. In addition to his work as editor of the Einstein papers in finding source material, Stachel assembled the many small clues that reveal Einstein's serious consideration of an emission theory of light; and he gave us the crucial insight that Einstein regarded the Michelson-Morley experiment as evidence for the principle of relativity, whereas later writers almost universally use it as support for the light postulate of special relativity.⁴

My goal in this paper is not to present a seamless account of Einstein's path to special relativity. That is an ambitious project, hampered by lack of sources and requiring a synthesis with Einstein's other

² *Papers*, Vol. 1, Doc. 5.

³ *Papers*, Vol. 1, Doc. 128.

⁴ Even today, this point needs emphasis. The Michelson-Morley experiment is fully compatible with an emission theory of light that contradicts the light postulate.

research interests at the time.⁵ Rather I seek to extend our understanding of several aspects of Einstein's path to special relativity:

- *The outcome of the magnet and conductor thought experiment.* This thought experiment showed Einstein that electric and magnetic fields might transform between inertial frames under rules that mix both fields and he hoped that this device might somehow enable Maxwell's electrodynamics to be made compatible with the principle of relativity. In Section 2, I will map out the prospects for the Galilean covariance of Maxwell's theory opened by this new device. They are promising but prove not to yield a single theory. A full exploration of the possibilities yields two partial theories with different field transformation laws and I call them the "magnet and conductor partial theory" and the "two charge partial theory". Each is associated with one part of Maxwell's theory and the tension between them reflects an awkwardness that Einstein did not mention, but was mentioned by Föppl, a possible source for Einstein's magnet and conductor thought experiment. It is that the relativity of motion of observables of the magnet and conductor thought experiment is not reflected throughout Maxwell's theory. Föppl illustrated the failure with his two charge thought experiment. That failure, captured formally in the existence of two incompatible partial theories each with its own defects, would have been a pressing problem for Einstein's program of relativizing electrodynamics and, perhaps, fatally discouraging to a less stubborn thinker.
- *Einstein's speculation on an emission theory of light.* In Section 3, I show why Einstein's remarks that he had held to Ritz's view on an emission theory of light cannot be taken literally. Ritz's work depended essentially on a skepticism about fields, which Einstein did not share and which led Ritz to seek action at a distance laws as the fundamental laws of electrodynamics. However a folk version of Ritz's view, articulated most clearly by Pauli, is a good candidate for an emission theory that Einstein might have entertained. It can be grafted directly onto the stronger one of the two partial theories mentioned above ("magnet and conductor partial theory") and would be initially appealing since would promise to preserve the gains of the analysis of the magnet and conductor while also accommodating an emission theory. Since the resulting theory still does not escape the defect of that partial theory, it was at best a brief way station for Einstein as he proceeded to develop quite general objections to any emission theory of light that I outline in Section 4.
- *Einstein's chasing a light beam thought experiment.* In his *Autobiographical Notes*,⁶ Einstein emphasized the importance of this thought experiment, first devised when he was 16 years old. In Section 5, I will argue that its original significance lay in arousing a visceral suspicion towards ether based theories, while not giving any cogent reasons for disbelieving such theories. The fertility of its basic idea—investigating how observers moving with light see the waveform—was proven later in Einstein's work, justifying the prominence Einstein accorded it in his recollections. In Section 6, I will suggest it enables strong

⁵ How could we ignore the possibility of a connection between Einstein's reflections on an emission theory of light and his 1905 postulation of the light quantum hypothesis? But what might that connection be? See Rynasiewicz, 2000, Sections 6 and 7.

⁶ Einstein (1949), pp. 48-51.

arguments against any emission theory of light, giving powerful yet simple grounding for his complaint that no emission theory could be formulated as a field theory.

- *Fizeau's experiment on the velocity of light in moving water and stellar aberration.* Einstein was scarcely able to remember if he knew of the most accomplished of the 19th century experiments on light propagation, the Michelson-Morley experiment, prior to his writing of the 1905 paper. In its place, Einstein singled out Fizeau's experiment and stellar aberration as the more memorable and compelling experiments. In Section 7, I will suggest their importance derives from their giving direct *experimental* foundation to Lorentz' notion of local time without requiring any detailed electrodynamical theory or Lorentz's theorem of corresponding states. I expect this last point to be evident to anyone who has fully understood the relevant section of Lorentz's 1895, *Versuch*, and how directly local time is expressed in the experimental results. Since the notion of local time becomes the relativity of simultaneity, when reinterpreted in the context of the principle of relativity, I suggest that these experiments earned their place in Einstein's thought by providing an experimentally grounded pathway to the relativity of simultaneity.
- In section 8, I remark that what is distinctive about the deliberations reported throughout this paper is that the effect of the motion of an observer on light is investigated in terms of its effect on the *waveform* of the light. While the historical evidence available is small, essentially none of it gives importance to Einstein reflecting on light *signals* used to synchronize clocks. So we must even allow the possibility that these reflections only entered in the last moments of years of work, when the essential results, including the relativity of simultaneity, were already established, but in need of a vivid and compelling mode of presentation. I warn of the danger of illicitly transferring the prominence of light signals and clocks in our thought to Einstein's historical pathway to special relativity.

It might seem perverse to persist in efforts to reconstruct Einstein's path to special relativity when the source material is so scant. However I think the effort is justified by the continuing fascination that Einstein's discovery exerts both inside and outside history of science. It has encouraged all manner of speculation by scholars about the relationship between Einstein's discovery and their special fields of interest, be they modes and methods within science; or Einstein's broader social and cultural context; just about everything in between; and many things that are not in between. As this literature continues to grow, it would seem perverse to me not to persist in efforts to reconstruct what was surely most directly relevant of all to the discovery, Einstein's own antecedent theorizing. And I'd really like to know what Einstein was thinking on the way to special relativity! In these efforts, I am fully aware of the historiographical pitfalls so well described by Stachel (1989, pp. 158-59), so that I need only refer the reader directly to that discussion and to endorse Stachel's analysis.

2. What Einstein Learned from the Magnet and Conductor Thought Experiment

The magnet and conductor thought experiment

Einstein began his celebrated 1905 “On the Electrodynamics of Moving Bodies,” by describing how then current, ether based electrodynamics treated the case of a magnet and conductor in relative motion. The full theoretical account distinguished sharply between the case of the magnet at rest in the ether and the conductor at rest in the ether. In the first case, a simple application of the Lorentz force law yields the measurable current. In the second, the time varying magnetic field of the moving magnet induces, according to Maxwell’s equations, a new entity, an electric field, and this field brings about the measurable current. What is curious is that the currents arising in each case are the same. The theory distinguishes the two cases but there is no observable difference between them; the measurable current depends only on the relative velocity. Cases like these, Einstein suggested, indicate that the ether state of rest is superfluous and that the principle of relativity ought to apply to electrodynamics.⁷

In a manuscript from 1920, Einstein recalled how this simple reflection had played an important role in the thinking that led him to special relativity. The essentially relevant parts of his recollection read:⁸

In setting up the special theory of relativity, the following ... idea concerning Faraday’s magnet-electric induction [experiment] played a guiding role for me.

[magnet conductor thought experiment described].

The idea, however, that these were two, in principle different cases was unbearable for me. The difference between the two, I was convinced, could only be a difference in choice of viewpoint and not a real difference. Judged from the magnet, there was certainly *no* electric field present. Judged from the electric circuit, there certainly was one present. Thus the existence of the electric field was a relative one, according to the state of motion of the coordinate system used, and only the electric and magnetic field *together* could be ascribed a kind of objective reality, apart from the state of motion of the observer or the coordinate system. The phenomenon of magneto-electric induction compelled me to postulate the (special) principle of relativity.

[Footnote] The difficulty to be overcome lay in the constancy of the velocity of light in a vacuum, which I first believed had to be given up. Only after years of [*jahrelang*] groping did I notice that the difficulty lay in the arbitrariness of basic kinematical concepts.

⁷ Einstein (1910, pp. 15-16) gives a slightly elaborated version of the original 1905 statement of the thought experiment.

⁸ Einstein (1920). Einstein’s emphasis.

Einstein's other recollection of the importance of this thought experiment is in a typescript note in English, with handwritten German corrections, in honor of Albert A. Michelson's 100th birthday and dated December 19, 1952.⁹ In the struck out typescript, Einstein discounts the influence of the Michelson-Morley experiment on him "during the seven and more years that the development of the Special Theory of Relativity had been my entire life." The handwritten notation expands and corrects the struck out typescript:¹⁰

My own thought was more indirectly influenced by the famous Michelson-Morley experiment. I learned of it through Lorentz' path breaking investigation on the electrodynamics of moving bodies (1895), of which I knew before the establishment of the special theory of relativity. Lorentz' basic assumption of a resting ether did not seem directly convincing to me, since it led to an [struck out: to me artificial appearing] interpretation of the Michelson-Morley experiment, which [struck out: did not convince me] seemed unnatural to me. My direct path to the sp. th. rel. was mainly determined by the conviction that the electromotive force induced in a conductor moving in a magnetic field is nothing other than an electric field. But the result of Fizeau's experiment and the phenomenon of aberration also guided me.

These recollections leave no doubt of the importance of the magnet and conductor thought experiment in directing Einstein's work towards special relativity. It is significant that Einstein calls it to mind in a tribute to Michelson at a time when the lore held that the Michelson-Morley experiment played a decisive role in leading Einstein to special relativity. Einstein corrects this lore and puts the magnet and conductor thought experiment in its place.

The recollections put no date on when the thought experiment compelled Einstein to postulate the special principle of relativity. The strong suggestion in both is that it was early in Einstein's deliberations. That early timing is made more concrete by the footnote to the 1920 recollection. After the thought experiment, much must still happen. He still faces years of years of groping and will still give serious thought to abandoning the constancy of the velocity of light—presumably referring to Einstein's deliberations on an emission theory of light—before he arrives at the 1905 insight of the relativity of simultaneity.

The transformation of the electric and magnetic field

The magnet and conductor thought experiment not only compelled Einstein to postulate the special principle of relativity, it also gave him an important new device for realizing it: as we transform between inertial frames, the electric and magnetic fields transform by rules that mix the two fields linearly. What might manifest as a pure magnetic field in one frame of reference will manifest as a combination of electric and magnetic fields in another. This device enabled Einstein to see how the

⁹ Document with control number 1 168, Einstein Archive. Available in facsimile at the Einstein Archives Online as <http://www.alberteinstein.info/db/ViewImage.do?DocumentID=34187&Page=1>

¹⁰ Part of translation from Stachel (1989a, p. 262).

relativity of motion in the observables of electrodynamics could be extended to the full theory. The induced electric field surrounding a moving magnet does not betoken the absolute motion of the magnet. It only betokens the motion of the magnet in relation to an observer, who judges the field generated by the magnet to have both magnetic and electric components.

This device of field transformation persists in Einstein's theorizing. It is central to the demonstration of the relativity of motion in electrodynamics in his 1905 "On the electrodynamics of moving bodies," with the full expression for the Lorentz transformation of the electric and magnetic field given in its Section 6.

Which transformation?¹¹

Years before, when Einstein first learned the device of such field transformations from the magnet and conductor thought experiment, upon which transformation did Einstein settle? Surely it was not the full transformation equations of 1905, but something a little less. What was it?

The thought experiment gives us just one special case that is easily reconstructed, as I have done in Appendix A. In the (primed) rest frame of a magnet, we have a magnetic field \mathbf{H}' and no electric field ($\mathbf{E}'=0$). If a charge e moves at velocity \mathbf{v} in this magnetic field, then the Lorentz force law *in vacuo* (L, below) tells us that the force \mathbf{f}' on the charge is $\mathbf{f}'/e = (1/c)(\mathbf{v}\times\mathbf{H}')$. Einstein now expects that this same force must arise in the (unprimed) rest frame of the charge from the transform of \mathbf{E}' , the electric field $\mathbf{E} = (1/c)(\mathbf{v}\times\mathbf{H}')$. That is, the field $\mathbf{E}'=0$ in the magnet rest frame transforms into the field $\mathbf{E} = (1/c)(\mathbf{v}\times\mathbf{H}')$ in a frame moving at \mathbf{v} . Schematically:

$$\mathbf{E}'=0 \rightarrow \mathbf{E} = (1/c)(\mathbf{v}\times\mathbf{H}') \quad (1)$$

The natural linear generalization of this rule is just

$$\mathbf{E} = \mathbf{E}' + (1/c)(\mathbf{v}\times\mathbf{H}') \quad (2)$$

(and I will argue below that this is more than just a natural choice; it is forced in certain circumstances).

What rule should apply to the transformation of \mathbf{H} ? There is a single answer to which modern readers are understandably drawn. Because of the symmetrical entry of \mathbf{E} and \mathbf{H} fields into Maxwell's equation, would not Einstein presume a similar transformation law for \mathbf{H} so that the combined law is

$$\mathbf{E} = \mathbf{E}' + (1/c)(\mathbf{v}\times\mathbf{H}') \quad \mathbf{H} = \mathbf{H}' - (1/c)(\mathbf{v}\times\mathbf{E}') \quad (3)$$

¹¹ What follows is limited to investigation of the prospects of the device of field transformations in the context of Lorentz' version of Maxwell's theory, which is based on just two fields as the basic quantities. This became Einstein's preferred version of Maxwell's theory and he had announced his intention to study it as early as December 28, 1901. (*Papers*, Vol. 1, Doc. 131.) John Stachel has pointed out to me that the two field transformations of Table 1 arise naturally in versions of Maxwell's theory based on four fields, \mathbf{E} , \mathbf{B} , \mathbf{D} and \mathbf{H} , such as Hertz' theory, which we know Einstein had studied earlier. (*Papers*, Vol. 1, Doc. 52.) \mathbf{E} and \mathbf{B} are governed by transformation (5) and \mathbf{D} and \mathbf{H} are governed by transformation (4). For a modern explication of the two transformations, see Stachel (1984). We might also modify Maxwell's theory so that just one field transformation applies. Jammer and Stachel (1980) drop the $\partial\mathbf{H}/\partial t$ term in (M4) to recover a modified theory that (excepting the Lorentz force law (L)) is covariant under (4).

This transformation is the field transformation law Einstein presented in his 1905 paper up to first order quantities in v/c ; and it is the very field transformation law that Einstein would have found when he read Lorentz's (1895) presentation of his theorem of corresponding states.

While it is possible that Einstein may have inferred to this transformation, I do not think that there are good grounds to expect it.¹² The symmetry of \mathbf{E} and \mathbf{H} in Maxwell electrodynamics is only partial. They do not enter symmetrically in the Lorentz force law and the \mathbf{E} field couples to sources whereas the \mathbf{H} field does not, so symmetry is not a compelling reason to proceed from (2) to (3). Of course we know in the long run that cultivation of (3) will bear great fruit. But, to use it in the short run, requires some prescience. Use of the first order Lorentz field transformation (3) requires the use of Lorentz' local time in transforming between frames of reference; otherwise covariance of Maxwell's equations fails even in first order and the whole exercise is for naught. It is one thing to use the first order Lorentz transformation and local time as Lorentz did: as a computational device for generating solutions of Maxwell's equations and, carefully, on a case by case basis, to show that various optical experiments admit no (first order) detection of the earth's motion with respect to the ether. But Einstein's quest is for the transformation that implements the relativity group. That is quite another thing. If he is able to use the first order Lorentz transformation and local time to implement that group, then he would have already to recognize that Lorentz' local time is more than a computational convenience. He must see it is the real time of clocks, the time of an inertial frame, every bit as good as the time of the frame from which he transformed. That requires him already to have his insight into the relativity of simultaneity. Further, since the first order Lorentz transformation preserves the speed of light to first order, there would seem little scope to doubt the constancy of the speed of light and toy with an emission theory of light.

Thus it is unlikely that Einstein inferred directly to the first order Lorentz transformations (3) from the magnet and conductor thought experiment; or, if he did, that he retained them in the core of his theorizing. For his recollections require years of reflection to pass before he arrived at the moment when his insight into simultaneity was decisive; and the above recollections suggest that the time period in which he entertained an emission theory of light was in those intervening years. Curious also is that the 1952 recollection contrasts Lorentz' 1895 work, which is criticized for its treatment of the ether, with Einstein's reflections on the magnet and conductor that provided the "direct path." That is an unlikely contrast if the magnet and conductor thought experiment brought Einstein directly to the essential content of Lorentz' work.

The prospects of a Galilean covariant electrodynamics

So what transformation was the immediate outcome of the magnet and conductor thought experiment for Einstein? We read directly from his recollections that it compelled him to seek an ether free electrodynamics compatible with the principle of relativity and one that may exploit some sort of

¹² The transformation is incomplete; it forms a group only if quantities of second order and higher are ignored. That can be remedied, of course, by the adjustments of 1905; but that presupposes sufficient commitment to the equations to want to remedy the problem.

field transformation law akin to (2) or (3). We know that as early as December 1901, Einstein was hard at work on a paper on a theory of the electrodynamics of moving bodies whose novelty included some ideas on relative motion.¹³ So presumably he was in possession of some sort of novel theory, although evidently it was not sufficiently coherent for him to proceed all the way to attempt publication.

While we have no direct statement of what that theory might have looked like, it is a matter of straightforward calculation to determine what the possibilities were. If we presume that Einstein's kinematics of space and time remain Galilean, then the field transformation laws associated with Maxwell's electrodynamics are given uniquely in Table 1. The table shows the four Maxwell field equations *in vacuo*, in Gaussian units, with charge density ρ and electric current flux $\mathbf{j}=\rho\mathbf{v}$, for a charge distribution moving with velocity \mathbf{v} .

¹³ Einstein wrote to Mileva Maric on December 17, 1901: "I am now working very eagerly on an electrodynamics of moving bodies, which promises to become a capital paper. I wrote to you that I doubted the correctness of the ideas about relative motion. But my doubts were based solely on a simple mathematical error. Now I believe in it more than ever." (*Papers*, Vol. 1, Doc. 128). See also Einstein to Maric, December 19, 1901, for a report by Einstein on discussions with Alfred Kleiner on "my ideas on the electrodynamics of moving bodies" (*Papers*, Vol. 1, Doc. 130). The possessive "my" here seems to have eclipsed Einstein's earlier remark to Maric, March 27, 1901, "How happy and proud I will be when the two of us together will have brought our work on the relative motion to a victorious conclusion!" (*Papers*, Vol. 1, Doc. 94; translations from Beck, 1983.)

$\nabla \cdot \mathbf{E} = 4\pi\rho$ (M1)	$\nabla \cdot \mathbf{H} = 0$ (M2)
$\nabla \times \mathbf{H} = \frac{4\pi}{c} \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}$ (M3)	$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}$ (M4)
	Lorentz force law $\mathbf{f}/e = \mathbf{E} + (1/c)(\mathbf{v} \times \mathbf{H})$ (L)
covariant under	covariant under
Galilean time and space transformation $t=t' \quad \mathbf{r}=\mathbf{r}'-\mathbf{u}t'$ Field transformations $\mathbf{E} = \mathbf{E}' \quad \mathbf{H} = \mathbf{H}' - (1/c)(\mathbf{u} \times \mathbf{E}')$ (4)	Galilean time and space transformation $t=t' \quad \mathbf{r}=\mathbf{r}'-\mathbf{u}t'$ Field transformations $\mathbf{E} = \mathbf{E}' + (1/c)(\mathbf{u} \times \mathbf{H}') \quad \mathbf{H} = \mathbf{H}'$ (5)
<i>The Two Charge Partial Theory</i>	<i>The Magnet and Conductor Partial Theory</i>
Defect <ul style="list-style-type: none"> • A moving magnet does not induce an electric field. • The Lorentz force law is not included, so observable effects of electric and magnetic fields are not deducible. 	Defect <ul style="list-style-type: none"> • A moving charge does not induce a magnetic field.

Table 1. Extent of Galilean Covariance of Maxwell's Electrodynamics

The table divides neatly into two columns. The two equations (M1) and (M3) are Galilean covariant if the field transformation (4) is invoked. The two equations (M2) and (M4) along with the Lorentz force law (L) are Galilean covariant if the field transformation (5) is invoked.¹⁴ (The demonstration of covariance is standard and sketched in Appendix B.) Unlike the first order Lorentz transformation (3), all these covariances are exact; they hold to all orders in v/c and they form a group. There is a lot to be read from the way the table divides.

It is shown in Appendix A that the content of the right hand column-- Maxwell equations (M2) and (M4) and the Lorentz force law (L)—are all that is needed to treat the magnet and conductor thought experiment in a Galilean covariant calculation. Hence I have labeled the equations in the right hand column the “*magnet and conductor partial theory*” since it is *all that is needed to treat the theory of the magnet and conductor thought experiment in a manner compatible with the principle of relativity of inertial motion*. This

¹⁴ I adopt the obvious conventions. The Galilean transformation maps a coordinate system ($t', \mathbf{r}'=(x', y', z')$) to another ($t, \mathbf{r}=(x, y, z)$), moving with velocity \mathbf{u} .

was Einstein's stated goal for all electrodynamics and here it is already for the case he found inspirational.

What mars the success of this partial theory, however, is that forces empirically incorrect results when it is applied to other cases. Take the case of a charge at rest. It is surrounded by an electrostatic field but no magnetic field, so $\mathbf{H}'=0$. If we now view this charge from another frame, the transformation $\mathbf{H}=\mathbf{H}'$ (5) assures us that there is still no magnetic field surrounding the charge. But that contradicts Oersted's famous result that an electric current—charges in motion—are surrounded by a magnetic field.

Föppl's two charges thought experiment

What are we to make of the other column in Table 1? Here is the remainder of Maxwell electrodynamics and it is Galilean covariant, but under a different field transformation law! This difference is the formal expression of a problem that Einstein did not mention in his celebrated discussion of the magnet and conductor thought experiment. In the case of a magnet and conductor, a straightforward application of Maxwell's theory shows that the observables depend only on the relative motion. But one can readily construct other thought experiments in which the observables do depend on absolute motions—or that they actually do not would require exploitation of the full apparatus developed by Lorentz that gets its final expression in Einstein's theory of relativity.

That there were other problematic thought experiments readily at hand had been pointed out clearly by August Föppl (1894) in the first of a venerable lineage of electrodynamics texts. Föppl's (1894, Part 5, Ch.1) text includes a favorable discussion of the relativity of motion and inquires into the extent to which it may be realized in Maxwell's electrodynamics. The magnet and conductor thought experiment is presented (pp. 309-10) as a case in which the relativity of motion is respected. As Holton (1973) shows in his discussion of this aspect of Föppl's work, there is some reason to believe that Einstein had read the Föppl volume, with its version of the thought experiment Einstein would make famous. Föppl immediately proceeded to warn his readers that the relativity of motion was not always respected and one might not always get the same results when systems are set into uniform motion. He made good on the warning with an even simpler thought experiment (pp. 310-11) that I will call the "two charges thought experiment."

One recognizes all the more that such a careful analysis [as given to the magnet and conductor] really was required from the fact that analysis does not yield the same result in all cases. Consider, for example, two electrically charged particles (material points) that move off next to one another in parallel paths with the same speed. They are at rest relative to one another. However they act on one another with quite different forces than they would if they were at absolute rest. Motion through the medium [ether] now leads to electrical convection and displacement currents and, in connection with them, to a magnetic field that is not present in the state of absolute rest. So this will still be true, if we also keep all external, disturbing influences distant and imagine both particles alone in an ether filled space, so that there are absolutely no reference bodies present, against which we could observe motion. Absolute motion already manifests a quite definite influence on

them, whereas that [absolute motion] could not be distinguished at all from a state of rest according to the axiom of kinematics discussed in the previous section. In cases of this type, therefore, the action of the bodies on each other does not depend solely on their relative motion.

The thought experiment is very simple. Consider two charges at rest in the ether. Their interaction is determined by ordinary electrostatics. They exert forces on each other according to Coulomb's inverse square law. Now set them into uniform motion. The interaction becomes very complicated. The moving charge becomes an electric current that will generate a magnetic field; and the time varying electric field around the moving charges will also generate a magnetic field. This magnetic field will act on the charges moving through it. In the case of the magnet and conductor, the analogous induced electric field is almost miraculously of just the right magnitude to obliterate any observable effect that might reveal which of the magnet or conductor is in absolute motion. The same miracle does not happen with the two charges. The extra forces due to the induced magnetic field are simply added to those already due to the electric field. The result is that the forces acting and thus the motions resulting would allow a co-moving observer to distinguish whether the pair of charges is moving through the ether or is at rest.¹⁵

Appendix C gives the calculations needed to show that the principle of relativity fails for the observables in the case of the two charges. The appendix calculates the general case of any static distribution of charges whatever that is then set into uniform motion, since it proves to be no more complicated. In the general case, new forces appear in the moving system as a result of the induced magnetic field, although the forces are second order in v/c small. What is important for our purposes, as Appendix C shows, is that Maxwell's equations (M1) and (M3) are all that is needed to compute the original field and the new magnetic field arising when the charges are set in motion. These equations are used to infer that the \mathbf{E} field of the charge distribution induces a magnetic field $\mathbf{H} = -(1/c)(\mathbf{v} \times \mathbf{E})$ when the system is set into uniform motion with velocity $-\mathbf{v}$. It is easy to see that this very same induced magnetic field could have been inferred directly from the field transformation law (4). The upshot is that the theory of the left hand column of Table 1, the *"two charges partial theory"* is all that is needed to treat the fields of the two charges thought experiment in a manner compatible with the principle of relativity of inertial motion.

The crucial omission is that the treatment extends only to the fields but not to the forces and accelerations associated with them. For the two charges partial theory does not include the Lorentz force

¹⁵ To see that a straightforward analysis will not save the principle of relativity for observables, note that Föppl's case of the two charges is, in its essentials, the same as the problem of determining the behavior of Lorentz's spherical electron when it is set in motion. In Lorentz' case, he must now deal with each of the infinitely many parts of the electron interacting with all the other parts by exactly the interaction that Föppl calls to mind for two point charges. Lorentz (1904, §8) is able to give an account that conforms to the principle of relativity (for observables) only by using the full apparatus of his theorem of corresponding states, including the contraction hypothesis in its generalized form that applies as well to the non-electromagnetic forces that hold the charges of the electron together.

law. Once that law is invoked for the thought experiment of the two charges (or any static charge distribution set into uniform motion) different forces are inferred for the cases of rest and motion and the principle of relativity is violated. This defect cannot be remedied easily. It is shown in Appendix B that field transformation (5) is the *unique* transformation under which the Lorentz force law (L) is covariant. Since Maxwell's equations (M1) and (M3) are not covariant under this transformation, a theory of processes governed by these two equations and the Lorentz force law cannot be given a Galilean covariant formulation.

The two charge partial theory suffers an additional defect analogous to that of the magnet and conductor theory. It precludes the induction of an electric field by a moving magnet. In the magnet's rest frame, we will have $E'=0$. Since its field transformation law (4) requires $E=E'$, there can be no induced electric field associated with a moving magnet, in contradiction with Faraday's experiments on induction.

What the device of field transformation brings

Let us take stock. In the magnet and conductor thought experiment, there are no observable consequences of absolute motion and Einstein reported the importance of this result in his early thinking on relativity theory. What Einstein would surely also have known was that that observable consequences could be recovered from absolute motion in other thought experiments in electrodynamics. Indeed if he read Föppl's account, as we have reason to believe he did, then he would have had just such a thought experiment brought to his attention as failing where the magnet and conductor thought experiment succeeded.

So the magnet and conductor thought experiment does not show satisfaction of the principle of relativity for all observables in electrodynamics. It shows them only in one part of electrodynamics and suggests a device, field transformations, that might bring the principle of relativity to that part of electrodynamics and perhaps more.

We do not know how Einstein applied the device when he first conceived it. However we can map out the space of possibilities that he would have to explore if he began to use the device within Maxwell's electrodynamics. The terrain is quite fixed; it is as described in Table 1. It is what Einstein would find just as long as he was willing to complete the exploration, although he might not present it or conceive it in quite the way I have. Maxwell's electrodynamics can be made Galilean covariant, but only if it is cleaved into two parts, each with its own field transformation law. The two parts complement each other. Each is able to give a Galilean covariant account of processes governed by two of Maxwell's equations; but the field transformation each invokes fails to conform to the processes accommodated by the other partial theory. The tension between the two thought experiments is now reproduced in the tension between the two partial theories.

The device of field transformations has not extended the partial conformity of the observables of Maxwell's theory to the principle of relativity. What it has done, however, is to extend the conformity of the theoretical structures, the fields, to the principle of relativity and that is noteworthy progress. Perhaps it was sufficient progress to figure in what the Einstein of December 1901 thought might become a

“capital paper.” If so, the nagging defects of the two partial theories might also have been sufficient to prevent writing or publishing just such a paper.

The path ahead

How might Einstein proceed with these results in hand? If he had to choose between the two partial theories, the choice would be obvious. The magnet and conductor partial theory was superior in so far as it supplied satisfaction of the principle of relativity for both fields and observables. But why force a choice? The obvious goal would be to unify the two partial theories. However, prior to insights about the relativity of simultaneity, there would be no way to do this. The tension between the two partial theories is readily recognizable as reflecting the most obvious problem facing attempts at a Galilean covariant electrodynamics: Maxwell’s theory entails a constant speed c for light and that result cannot be Galilean covariant. Either of the two parts of Maxwell’s theory alone is insufficient to entail the speed of propagation of waves, so each may admit a Galilean covariant formulation. But once the two parts are combined, the constant speed for light can be derived; a single Galilean covariant formulation will no longer be possible. “The difficulty to be overcome,” as Einstein added in a footnote to his 1920 recollection of the magnet and conductor thought experiment, “lay in the constancy of the velocity of light in a vacuum, which I first believed had to be given up.”

One way to proceed is to attempt to modify Maxwell’s theory in some way to enable Galilean covariance under a single field transformation law. It is obvious that a Galilean covariant electrodynamics must be an emission theory of light, that is, a theory in which the velocity of the emitter is vectorially added to the velocity of the light emitted. If an emitter at rest emits light with velocity c , then Galilean kinematics entails that the emitter, moving at velocity v , must emit the light at velocity $c+v$. So an emission theory of light is necessary in a Galilean covariant electrodynamics. (But it is certainly not sufficient for Galilean covariance of the electrodynamics—and we will see an overlooked failure of sufficiency below in a well know emission theory!) So consideration of an emission theory of light will inevitably arise as long as one’s exploration of Galilean covariant electrodynamics is thorough enough. Thus it is not at all surprising that Einstein would proceed to investigate an emission theory of light in a later phase of his work. In the following section, I will review the little we know of Einstein’s investigations into an emission theory. I will also point out a connection between the partial theories considered here and our best guess for Einstein’s emission theory: if one takes the strongest of the two partial theories, the magnet and conductor partial theory, it turns out it can be extended without modification to this emission theory.

3. Einstein’s Efforts towards an Emission Theory of Light

Even with the insight afforded by the magnet and conductor thought experiment in hand, the footnote to Einstein’s 1920 recollection quoted above shows that years of theoretical groping were still needed to make good on the principle of relativity and that they included serious consideration of abandoning the constancy of the speed of light. Later Einstein (1909, p. 487) remarked that the

abandoning of the ether led naturally to an emission theory of light: “Then the electromagnetic fields that constitute light no longer appear as states of a hypothetical medium, but as independent structures, which are emitted by light sources, just as in Newton’s emission theory of light.”—and it is impossible for modern readers to fail to connect this remark to Einstein’s work on light quanta. As we shall see below, on quite a few occasions Einstein identified his own approach to an emission theory to be akin to that of Walter Ritz.

Ritz’s view.

What was Ritz’s view? It is laid out in a 130 page article (Ritz, 1908), which is summarized in another shorter paper (Ritz 1908a).¹⁶ In one sense, the view laid out is a great deal more than a theory; it is a synoptical view of the present state of electrodynamics, dissenting strongly from some of the mainstream views. But it is also something less than a complete theory. Rather it is an elaborately articulated program for the development of a theory along with quite extensive implementation of the program. Further implementation of the program was terminated by Ritz’s failing health and death in July 1909 of tuberculosis.

Ritz’s (1908) paper was divided into two parts. The first developed a series of skeptical viewpoints, each of which led more or less directly to a revision of then present electrodynamical theory. Those of relevance here included:

Fields and other quantities. Ritz expressed skepticism about the many quantities used in electromagnetic field theory, most notably the electric and magnetic fields. He urged a return to laws expressing the interaction between two charges of the type developed by Weber and others in the 19th century. These action at a distance laws expressed the force exerted by one charge on another in terms of the distance between them and their motions. Ritz even urged that this force could be eliminated in favor of the charges’ motions.

Ether. Ritz laid out objections to the existence of the electrodynamical ether. He urged it should be eliminated from electrodynamics and that the principle of relativity should be restored.

Einstein’s special theory of relativity. Ritz criticized Einstein’s way of implementing the principle of relativity. He felt that Einstein’s insistence on retaining Lorentz’s electrodynamics insufficient grounds to support the strange kinematical notions Einstein introduced.

Retarded potentials. Ritz urged that the presentation of Lorentz’s electrodynamics in terms of Maxwell’s differential equations was incomplete. Instead he favored the more restricted formulation in terms of retarded potentials. This restriction eliminated the advanced solutions of Maxwell’s equations, which, Ritz felt, violated energy conservation, in so far as they represented a never seen convergence of radiation from spatial infinity.

The second part of Ritz’s paper sought to develop the program implicit in the first part through the following strategy. In the first part he had laid out a progression of results in the then standard theory. He now sought to replicate these results, but modifying them in such a way as to bring them into

¹⁶ For discussion and an English translation of the shorter Ritz (1908a), see Hovgaard (1931).

accord with the principle of relativity. The progression of the first part began with a statement of Maxwell's equations (M1)-(M4) and the Lorentz force law (L), drawn from Lorentz' formulation of Maxwell's theory. He then introduced a scalar potential φ and a vector potential \mathbf{A} in the usual way. They are defined implicitly by:

$$\mathbf{E} = -\nabla\varphi - (1/c) \partial\mathbf{A}/\partial t \quad \mathbf{H} = \nabla\times\mathbf{A} \quad (6)$$

Maxwell's equations, reexpressed in terms of these potentials, can then be solved by retarded potentials. These retarded potentials arise when the scalar and vector potentials $\varphi(x,y,z,t)$ and $\mathbf{A}(x,y,z,t)$ at position (x,y,z) and time t are expressed in terms of an integral over what we would now call the past light cone of the event (\mathbf{r},t) ; that is over all events (x',y',z',t') , where the time t' is retarded according to

$$t' = t - r/c \quad (7)$$

with r the distance between points (x,y,z) and (x',y',z') . The integrals are

$$\varphi(x,y,z,t) = \iiint \frac{[\rho']}{r} d\tau' \quad \mathbf{A}(x,y,z,t) = \frac{1}{c} \iiint \frac{[\rho'\mathbf{v}']}{r} d\tau' \quad (8)$$

where the notation "[ρ']" designates that ρ is computed at $(x',y',z',t'=t-r/c)$ and similarly for $[\rho'\mathbf{v}']$.¹⁷

These retarded potentials were only an intermediate for Ritz. He then proceeded to report with approval work of 1903 in which Schwarzschild found the corresponding expression for the retarded force acting on a charge due to the charge distribution in space (Ritz, 1908, p. 326). The expression was so unwieldy that, in the shorter paper, Ritz (1908a, p. 432) stated frankly that "it is a rather complicated expression which we will not write down" and I will follow his good sense. Ritz then proceeded to simplified versions of Schwarzschild's result for special cases, such as two interacting charges with small velocities and accelerations (Ritz, 1908, p. 348, 1908a, p. 433).

The starting point of Ritz's modification in the second part of his paper was the retarded action implicit in the time (7) used in Lorentz's theory. The distance r was measured in a coordinate system at rest in the ether; so this retardation time expresses the constancy of the velocity of light and electromagnetic effects with respect to the ether. In its place, Ritz (1908, p. 373) proposed that electromagnetic action propagates with a velocity c with respect to the source. To make it easy to visualize his proposal, he imagined that electric charges constantly emit infinitely small, fictitious particles in all directions with a radial velocity c with respect to the source. These fictitious particles model the dissemination of the electromagnetic action of one charge onto another and of light. In accord with his skepticism about the ether, he preferred to call it projection, which connotes ballistics, rather than propagation, which connotes transmission by a medium. The clearest comparison between the older view of the propagation of electromagnetic action and his view of its projection came in the following expressions. In Lorentz' theory, in a coordinate system at rest in the ether, the radius of the sphere R at time t emanating from an event at (x',y',z') at time $t'=t-R/c$ is

$$c^2(t-t')^2 = R^2 = [x-x'(t-R/c)]^2 + [y-y'(t-R/c)]^2 + [z-z'(t-R/c)]^2 \quad (9)$$

¹⁷ Ritz (1908, p. 325). I have simplified Ritz's notation slightly by substituting a single boldface vector for the three components Ritz wrote out individually.

This represents an expanding sphere whose center remains at one point at rest in the ether. In its place, Ritz proposed that the radius of the expanding sphere r at time t would be given by¹⁸

$$r^2 = [x - x'(t-r/c) - (r/c)v_x'(t-r/c)]^2 + [y - y'(t-r/c) - (r/c)v_y'(t-r/c)]^2 + [z - z'(t-r/c) - (r/c)v_z'(t-r/c)]^2 \quad (10)$$

The velocity \mathbf{v}' is the velocity of the source and $\mathbf{v}'(t-r/c)$ is the velocity of the source at the moment of emission. Equation (10) describes an expanding sphere whose center is no longer at rest in the coordinate system; if the source were to continue to move uniformly with velocity \mathbf{v}' , it would remain the center of the sphere. This geometrical description was favored by Ritz. We might now just say that the velocity of the source at the instant of emission is added vectorially to the velocity of the action.

With this model in mind, Ritz proceeded directly to the expressions developed by Schwarzschild, such as for the interaction of two charges, and sought ways to eliminate any absolute velocities in them, as required by the principle of relativity. Note that he does not explicitly address the expressions (8) for the retarded potentials, but proceeds directly to modifying expressions for the interaction of two charges. The results are far from simple and not unique. It is helpful to see the expression Ritz presents for the interaction of two charges e and e' , if only to see just how complicated it is. It is given (Ritz, 1908, p. 380) as

$$F_x = \frac{ee'}{r^2 \left(1 - \frac{rw'_r}{c^2}\right)} \left[\cos(r, x) \varphi \left(\frac{u_r}{c}, \frac{u^2}{c^2} \right) - \frac{u_x u_r}{c^2} \psi \left(\frac{u_r}{c}, \frac{u^2}{c^2} \right) - \frac{rw'_x}{c^2} \chi \left(\frac{u_r}{c}, \frac{u^2}{c^2} \right) \right],$$

$$F_y = \dots$$

$$\varphi = 1 + \frac{3-k}{4} \frac{u^2}{c^2} + \frac{3(1-k)}{4} \frac{u_r^2}{c^2} + a_1 \frac{u^4}{c^4} + \dots$$

$$\psi = \frac{k+1}{2} + b_1 \frac{u^2}{c^2} + b_2 \frac{u_r^2}{c^2} + \dots$$

$$\chi = 1 + c_1 \frac{u^2}{c^2} + c_2 \frac{u_r^2}{c^2} + \dots \quad (11)$$

The achievement of this expression is that all the velocities that enter, $u, u_r, u_{x'}$ are relative velocities, although the accelerations, $w_r, w_{x'}$ are not relative and, of course, need not be if all that is sought is the relativity of inertial motion. The expression contains many underdetermined constants (K, a_i, b_i, c_i), which are to be determined by experience. Ritz then proceeded to more specialized cases such as when the speeds and accelerations of the charges are small.

The above gives only a flavor of the range of material in Ritz's paper, which also includes an electromagnetically based theory of gravitation, in which gravitational action also propagates at c .

¹⁸ In the above formulae (9) and (10), " $x'(t-R/c)$ " is to be read as "the value of the x' coordinate of the source at time $t-R/c$ "; and similarly for the remaining terms.

How was Ritz's theory reported?

While Ritz's view could not be described as a theory but was really an elaborate program of research, very little of this entered the literature in which Ritz's name is invoked.¹⁹ The reason, I am presuming, is that this literature was largely devoted to empirical testing of different views about light. What could be tested most directly of Ritz's views was whether the velocity of the emitter is actually added to the velocity of the light emitted. Since that proposition was so central to Ritz's theory and open to test, it needed to be decided before more detailed investigation of Ritz's views was warranted.²⁰ A paper by Ehrenfest (1912) calls attention to Ritz's work on an emission theory. His discussion is devoted essentially to empirical testing and is spare in the details he gives of Ritz's views. He displays assertions (p. 317):

[B] A light pulse emitted by a non-accelerated light source L travels in a concentric sphere, whose radius increases with the constant speed V and whose center remains coincident with L .

...

[C] An observer ascertains a greater speed of light for a light source approaching him than for one at rest with respect to him.

He later glosses Ritz as having a theory in accord with these assertions (p. 318):

It is well known that Ritz developed such an emission theory of light [footnote includes citation of Ritz (1908)]. In this theory, electrons emit retarded potentials according to (B) and (C) and with rejection of the postulate (D) [Einstein's light postulate].

What is striking is that Ehrenfest's gloss encapsulates Ritz's view as a theory concerning retarded potentials, while Ritz at best regarded them as intermediates to Weber-like action at a distance laws. The

¹⁹ The significant exception is the continuing literature that is unconvinced of the necessity of adopting special relativity. O'Rahilly (1938) includes a fairly detailed exposition of Ritz's real views (Ch. XI), urges that they have been slighted in discussions of Einstein's relativity theory (Ch. XIII, §5) and concludes the Epilogue with a provocative "*We therefore reject the false dilemma: Aut Einstein aut nullus!*" [Einstein or nothing!]

²⁰ De Sitter's (1913) very short note reported a disproof of Ritz's proposition by observing the light from double stars, which seemed to be unaffected by the velocities of approach and recession of the stars as they orbited each other. He reported only as much of Ritz's view as was needed for the test: "If a light source has a speed u ...then, according to Ritz's theory, the speed of the emitted light in the same direction is $C+u$, where C is the speed of light emitted from a source at rest." Tolman (1912) includes Ritz's view with discussions of other emission theories of light. He allows (p. 137) that Ritz has proposed "a very complete emission theory of electromagnetism." But he recites just enough of Ritz's views to enable testing, for example (p. 137): "According to this theory, light retains throughout its whole path the component of velocity which it obtained from its original moving source, and after reflection light spreads out in a spherical form around a center which moves with the same velocity as the original source."

idea that Ritz's theory was centrally concerned with retarded potentials was solidified by Pauli's 1921 *Encyklopädie* article, which has become the standard citation for Ritz's theory and the ensuing empirical investigations that refuted it. Pauli (1921, p.6) wrote of efforts to construct a theory of light within electrodynamics that relinquishes the constancy of the velocity of light:

Only Ritz has succeeded in doing this in a systematic theory. He retains the equations

$$\text{curl} \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} = 0 \quad \text{div} \mathbf{H} = 0 \quad [(M4'), (M2')]$$

so that the field intensities can be derived, just as in ordinary electrodynamics, from a scalar and vector potential

$$\mathbf{E} = -\text{grad} \phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \quad \mathbf{H} = \text{curl} \mathbf{A} \quad [(6')]$$

The equations

$$\phi(P, t) = \int \frac{\rho dV_{P'}}{[r_{PP'}]_{t'=t-r/c}} \quad \mathbf{A}(P, t) = \int \frac{(1/c) \rho \mathbf{v} dV_{P'}}{[r_{PP'}]_{t'=t-r/c}} \quad [(8')]$$

of ordinary electrodynamics are now, however, replaced by

$$\phi(P, t) = \int \frac{\rho dV_{P'}}{[r_{PP'}]_{t'=t-[r/(c+v_r)]}} \quad \mathbf{A}(P, t) = \int \frac{(1/c) \rho \mathbf{v} dV_{P'}}{[r_{PP'}]_{t'=t-[r/(c+v_r)]}} \quad [(12)]$$

While Pauli does not identify the variable, presumably v_r is the velocity of the source at time t' in the direction of the point P. With this summary, Ritz's program has now been reduced to the simplest modification of standard electrodynamics. Lorentz's retardation time $t'=t-r/c$ (7) in the retarded potential integrations (8), (8') are simply replaced by the retardation time accompanying Ritz's views on the velocity of propagation of electromagnetic action:

$$t' = t - [r/(c+v_r)] \quad (13)$$

The transition from retarded potentials (8), (8') to (12) is exactly the transition called for by Ritz's proposal that we replace the motions (9) for the propagation of electromagnetic action with (10) for its projection.

Natural as Pauli's formulation of Ritz's theory may be, I have not found its central expression for the projected potentials (12) in Ritz's papers. Ritz proceeded directly to action at a distance expressions such as (11) since they were the fundamental goals of his program; expressions in retarded potentials were at best intermediaries, but as such would be in keeping with his thought. However it is not hard to understand why Pauli and perhaps Ehrenfest would present Ritz's views as they did. While Ritz had a program in electrodynamics, it is very hard to state a simple end point that is the proposed replacement of then standard electrodynamics. Rather Ritz's papers are filled with expressions like (11), valid only for special cases. What Pauli recognized, presumably, is that this difficulty in Ritz's views derives from his insistence that electrodynamics return to Weber like action at a distance laws. The difficulty is not a result of that aspect of Ritz's work that was of interest to Pauli in writing a review article on relativity theory; that is, Ritz's proposal that the velocity of light depend on the velocity of the emitter. So perhaps Pauli felt he was serving his readers well by shielding them from the unnecessary complications of Ritz's other views. Or perhaps he had not sifted Ritz's papers for the final result but had been informed by an

unreliable source. With commendable lack of concern for the quibbles of later historians of science, Pauli reported what Ritz would surely have concluded if only he could suppress his skepticism about fields.

We now have three Ritzes:

The Real Ritz. This is the Ritz of Ritz (1908), enmeshed in an elaborate project to reconfigure electrodynamics.

Pauli-Ehrenfest's Ritz. This is the Ritz who merely sought to reconfigure electrodynamics with retarded potentials that use a projected, retardation time (13) in order to restore Galilean relativity to electrodynamics.

The Experimentalists' Ritz. This is the Ritz that merely proposed that the velocity of the source should be added vectorially to the velocity of light.

Einstein on the similarity between Ritz's and his own emission theory

The earliest remarks we have by Einstein relating his own ideas on an emission theory of light and those of Ritz arise from Einstein's reaction to Ehrenfest's (1912) paper. In a letter from mid 1912 to Ehrenfest responding to the paper, Einstein wrote:²¹

I was not annoyed in the least by your article. On the contrary. Such considerations are quite familiar to me from the pre-relativistic time. I certainly knew that the principle of the constancy of the velocity of light is something quite independent of the relativity postulate; and I considered what would be more probable, the principle of the constancy of c , as was demanded by Maxwell's equations, or the constancy of c , exclusively for an observer sitting at the light source.

In his initial response to Ehrenfest's paper in an earlier letter of 25 April 1912 (*Papers*, Vol. 5, Doc. 384), Einstein allowed that this own thinking on an emission theory was akin to Ritz's:

I believe that there are quite simple experiments to test Ritz' conception, which, incidentally, was also mine before rel. theory.

What is important is the timing and context of Einstein's letter. He wrote less than a decade after his own work on an emission theory and, as remarked in the later (June 1912) letter, still had a comfortable memory of that earlier work. Ehrenfest's paper was raising the question of empirical tests that might distinguish Einstein's theory of relativity from a theory attributed to Ritz. In spite of Einstein's cheer ("not annoyed in the least"), he could not overlook that this was a challenge to his theory. Ehrenfest was apparently standing in for Ritz, whose death in 1909 precluded Ritz defending his own work. Einstein would surely want to be circumspect over claims made in this context and not assert lightly that he had already conceived of Ritz's view.

Although written much later,²² a more revealing statement is in the draft of a response written on the back of a letter dated 1 February 1952 to Einstein from C. O. Hines. (Einstein Archive 12 250, 12 251.)

²¹ Einstein to Ehrenfest, "before 20 June 1912," *Papers*, Vol. 5, Doc. 409. Einstein proceeds immediately to say that he chose the first, so this recollection immediately jumps over the time he spent developing and assessing his emission theory.

Hines reported difficulties in his study of Ritz's treatment of light and pressed Einstein for assistance, hoping that Einstein had had discussions with Ritz on the subject. Einstein replied, now addressing Ritz's ideas directly:

Ritz's ideas on electrodynamics and optics are not so far developed that one can call them a "theory." What is special in them is that there does not exist a definite speed for light propagation at a position and in a given direction, but that this [speed] depends on the state of motion of the light source. Then one cannot trace light propagation back to differential equations, but one must introduce "retarded potentials," which is a kind of action at a distance.

Before setting up the special theory of rel., I had myself thought of investigating such a possibility.

It is revealing that Einstein corrects Hines' supposition that Ritz really had a theory and not merely, as I have suggested above, a well advanced program of research. That shows that Einstein had some real familiarity with Ritz's work. I would also suppose that this familiarity was in place when Einstein responded to Ehrenfest's (1912) paper. Where Ehrenfest (1912) talks of Ritz's "theory," Einstein replies by calling it Ritz's "conception."²³

²² Other later remarks by Einstein affirm the kinship of Einstein's early ideas with Ritz's. With a cover letter dated 21 March 1922, Mario Viscardini sent Einstein an article for his opinion. It was described as providing a new solution to the Michelson experiment. Einstein responded (Einstein Archive 25-302; translation, Rynasiewicz, 2000, p. 168):

The hypothesis articulated in the article, that in free space light has the constant velocity c , not with respect to the coordinate system but relative to the light source, was discussed for the first time in detail by the Swiss physicist W. Ritz and was seriously taken into consideration by myself before the formulation of the special theory of relativity.

Einstein wrote on the back of a letter from A. Rippenbein of 25 August 1952 that once again proposed a novel theory of the motion of light (Einstein Archive, 20-046; translation from Stachel, 1982, p. 189): "Your attempt to replace special relativity with the assumption that the velocity of light is constant *relative to the source of light* was first advocated by Ritz...even before setting up the special theory of relativity, I rejected this way out..." Shankland (1963/73) reported that Einstein "told me that he had thought of, and abandoned the (Ritz) emission theory before 1905."

²³ At the time of Einstein's discussion with Shankland in the 1950s, Einstein seemed to have sufficiently detailed recollection of Ritz's theory to dispute Shankland's (1963, p. 49) praise of the theory:

When I [Shankland] suggested that Ritz's theory was the best of the several emission theories of light, he shook his head and replied that Ritz's theory is very bad in spots. [footnote: citation to Einstein, 1909a] But he quickly added, "Ritz made a great contribution when he showed that frequency differences are the crucial thing in spectral series."

The more revealing remark, however, comes after. Having pointed out that Ritz did not have a definite theory, Einstein extracts the important part of Ritz's program and formulates it as the idea that one *must* introduce retarded potentials. He then proceeds to assert that this was the sort of possibility he had investigated himself.

Which was Einstein's Ritz?

That is, what did Einstein mean when he remarked to Ehrenfest in 1912 that Ritz's conception has been his own? We can immediately rule out the *Real Ritz*. Einstein's sensibilities are well known. He was uninterested in Weber style action at a distance laws as the fundamental laws of electrodynamics. We shall see below that one of the complaints Einstein levels against an emission theory was that he could see no way of converting the theory into a field theory. That is precisely the reverse of Ritz's program, which was to convert field theories into action at a distance laws, even at the cost of extraordinary complications in the laws.

What of the *Experimentalists' Ritz*? Again it is unlikely this is Einstein's Ritz. I have already given grounds for believing that Einstein knew details of Ritz's work—specifically their complicated, programmatic nature. There are more reasons to believe this. By 1912, Einstein had read some of Ritz's work (though not necessarily the relevant papers) and had had some interactions with him. Ritz (1908, Part 1, §3; 1908a, pp. 434-35) had urged that Lorentz's electrodynamics should be restricted to retarded potentials and the advanced potential solutions disallowed. Ritz (1909) then urged that this restriction would resolve the thermodynamic difficulties surrounding thermal radiation (out of which quantum theory emerged). Einstein (1909, pp. 185-86) disputed Ritz's solution, arguing for the admissibility of both retarded and advanced solutions. This attracted a response from Ritz (1909a); and the matter was resolved with a polite statement of their differences in a jointly signed article, Ritz and Einstein (1909). Einstein may never have seen Ritz (1908, 1908a); he may have read Ritz (1909) only because it appeared in a journal in which Einstein sought publication, *Physikalische Zeitschrift*. Perhaps Einstein could engage in all these interactions with Ritz without learning that Ritz (1908, 1908a) contained outspoken even polemical assaults on his special theory of relativity. What makes that blissful ignorance extraordinarily unlikely is that Ritz happened to be the major competing candidate for Einstein's first academic position at the University of Zurich. The committee favored Ritz over Einstein, but chose Einstein only because of Ritz's incurable ill-health. Alfred Kleiner, the professor at Zurich who had fostered the position, initially favored Ritz as well. (See Fölsing, 1997, p. 249 for further details.) While the popular image is of an otherworldly, absent-minded Einstein, the real Einstein of 1909 was eager and ambitious and surely not likely to want to be uninformed of the outspoken criticism of his work from someone who proved to be his principal professional rival. And if Einstein somehow did not know that Ritz was his rival and had even not heard of Ritz's criticism, all this would be likely to change once he was installed at the University of

From his interviews with Einstein, Wertheimer (1959, p.216) reports that work by Einstein on a modified set of Maxwell's equations that might admit a variable speed for light persisted "for years." Ritz was not mentioned.

Zurich. A committee of eleven had voted on the appointment. Might not one of those, perhaps Kleiner himself, ask Einstein to respond to Ritz's criticism? And this would not be the occasion for an uninformed response.

So Einstein's Ritz lay somewhere between the *Real Ritz* and the *Experimentalists' Ritz*. We can certainly imagine many Ritzes in between. They would all be characterized by efforts to use the emission prescription of (10) for the propagation of electromagnetic action to relate the electromagnetic quantities at one point in space and to the distribution and motion of electric sources throughout space. Pauli and Ehrenfest have conveniently supplied us with a description of the intermediate Ritz that seems naturally to have sprung to their minds. That this was also Einstein's Ritz is strongly suggested by Einstein's remarks to Hines quoted above. Einstein explicitly does what Pauli and Ehrenfest do tacitly: reduce and reformulate Ritz's program into a proposal relating to retarded potentials. So I conclude that Einstein's Ritz was the *Pauli-Ehrenfest's Ritz*, or something closely related.

How should we read Einstein's remark to Ehrenfest and others that his conception agreed with Ritz's? We should read it in its context in which Ritz's program had come to be understood as something like Pauli's reduced version. For example, Einstein responded to a paper in which Ehrenfest (1912) characterizes Ritz's theory as one in which "electrons emit retarded potentials according to [emission theory of light]." We should understand Einstein to be saying to Ehrenfest, "Ritz's conception (as you have misdescribed it) was also mine" and to be tactfully reserving the parenthetical reprimand on Ehrenfest's misdescription.

A path from the magnet and conductor to the retarded potentials

One other consideration makes it plausible that Einstein's speculation on an emission theory passed through consideration of Pauli's retarded potentials (12) and perhaps even lingered there. It turns out that there is a natural and direct path to them from the device of field transformations suggested by the magnet and conductor thought experiment. Here is how it arises.

We have seen above that the four Maxwell equations divide uniquely into two pairs, each associated with a Galilean covariant theory with a different field transformation law. The two charge partial theory is based on Maxwell's equations (M1) and (M3). The magnet and conductor partial theory is based on (M2) and (M4). The latter is clearly superior in that it also incorporates the Lorentz force law (L) and accommodates the thought experiment Einstein found so motivating. So, if he conceived these partial theories at all, he might well be tempted to retain the magnet and conductor partial theory and seek to modify the two charge partial theory in an attempt to find a unified theory.

This path would lead directly to the emission theory Pauli ascribed to Ritz. The important fact about (M2) and (M4) is that field sources--charge density ρ and flux \mathbf{j} --do not appear in them. The sources appear only in the other two equations (M1) and (M3). Since emitters are merely accelerating charges and light the propagating waves they generate, these two equations (M1) and (M3) are the natural candidates for modification if an emission theory of light is sought. But can such a modification of (M1) and (M3) be found that would not compromise (M2) and (M4)? It can. Einstein merely needs to reformulate the theory in terms of retarded potentials as in (6), (6') and (8), (8') above. As Pauli suggests, one immediately

incorporates the emission theory by merely altering the retardation time accordingly. Yet Maxwell's equations (M2) and (M4) are left untouched, for, by familiar theorems, those two equations are equivalent to the assertion of the existence of the scalar and vector potentials.²⁴ That is, assuming the magnet and conductor partial theory entails the existence of the potentials φ and \mathbf{A} ; this is the path from magnet and conductor partial theory to the retarded potentials. The path back is just as easy; by assuming a retarded potential formulation of electrodynamics that uses potential φ and \mathbf{A} defined by (6), (6') entails Maxwell's two equations (M2) and (M4).

The fatal defect of the theory Pauli attributed to Ritz

This is an harmonious extension of the magnet and conductor theory. If Einstein had entertained extending the magnet and conductor partial theory towards an emission theory of light, we could readily imagine him finding it. Unfortunately we could equally well imagine him finding the fatal defect in the theory, a defect that Pauli did not mention. The original tension between the two partial theories, as captured by the conflicting field transformation laws (4) and (5), remains unresolved.

To see the problem we need to determine the transformation laws for the potentials φ and \mathbf{A} . As shown in Appendix B, the field transformation (5) can be extended to the potentials φ and \mathbf{A} by assuming that they transform according to

$$\varphi = \varphi' - (1/c)\mathbf{u}\cdot\mathbf{A}' \quad \mathbf{A} = \mathbf{A}' \quad (14)$$

Unfortunately these transformations are incompatible with the covariance of the retarded potential integrals (12). The quickest way to see the incompatibility is to take the case of a source charge distribution that is at rest. Since we have $\mathbf{v}=0$ everywhere, it follows from (12) that $\mathbf{A}=0$. If we now transform to another frame using the above transformation law, we have $\mathbf{A}'=0$. That is a disaster. In the new frame, the charges will no longer be at rest and thus by (12) will produce a non-vanishing \mathbf{H}' field, so that \mathbf{A}' cannot vanish. The transformation (14) for the potentials φ and \mathbf{A} simply replicates the defect of the magnet and conductor partial theory and its field transformation $\mathbf{H}=\mathbf{H}'$, which also precludes a moving charge from inducing an electric field.

There is no simple repair. One might wonder whether the alternative field transformations (4) might be called upon in some way, since they do not include $\mathbf{H}=\mathbf{H}'$. Or one might inspect the retarded potential integrals (12) and notice that they would be covariant under the transformation $\varphi' = \varphi$ and

²⁴ Maxwell's equation $\nabla\cdot\mathbf{H} = 0$ (M2) asserts that \mathbf{H} is divergenceless; so there must exist another vector field—let us call it \mathbf{A} —such that $\mathbf{H} = \nabla\times\mathbf{A}$. Substituting this expression for \mathbf{H} into Maxwell's equation (M4), we learn that $\nabla\times[\mathbf{E} + (1/c)(\partial\mathbf{A}/\partial t)]=0$. That is $\mathbf{E} + (1/c)(\partial\mathbf{A}/\partial t)$ is irrotational, so there must exist a scalar field φ , such that $\mathbf{E} + (1/c)(\partial\mathbf{A}/\partial t) = -\nabla\varphi$. These are equivalent to the expressions for \mathbf{E} and \mathbf{H} in (6). This shows that the equations (M2) and (M4) entail the existence of the scalar and vector potentials; the converse entailment follows just by reversing the above argumentation.

$\mathbf{A}' = \mathbf{A} + (1/c)\varphi\mathbf{u}$. All this is to no avail.²⁵ We must recall that the mere existence of the potentials as defined in (6) entails Maxwell's equations (M2) and (M4). As Appendix B shows, the transformation (14) is the one associated with transformation (5), under which (M2) and (M4) are covariant. Any other transformation for φ and \mathbf{A} would be incompatible with the covariance of (M2) and (M4) and thus with the covariance of the definitions of φ and \mathbf{A} themselves.

Thus, if Einstein followed this path to its end, he would have been disappointed. Indeed what we have found is that the theory Pauli attributed to Ritz in (6') and (12) is not Galilean covariant after all! The definitions (6') require field transformations (5) and (14); but the integrals (12) are not covariant under these field transformations. Presumably Pauli (1921, p.8) overlooked this since he proceeded to aver that "the relativity principle is automatically satisfied by all such [emission] theories." Conformity to an emission theory of light is necessary for a Galilean covariant electrodynamics; but, as this example shows, that conformity is not sufficient to assure Galilean covariance.

In sum...

In sum, it is not so easy to recover a clear statement of just what Einstein is claiming for his earlier view when he equates it with "Ritz's conception." Setting the remark in its context, the view claimed would use an emission prescription akin to (10) for the propagation of electromagnetic action using some sort of non-local field law. The most plausible, concrete formulation at hand is Pauli's retarded potentials (12) with a projected retardation time. While we certainly cannot preclude other formulations of an emission theory by Einstein, Pauli's proposal fits well with Einstein's remark to Hines that a Ritz inspired emission theory *must* be formulated in terms of retarded potentials. There is also a natural path to this same formulation directly from the magnet and conductor partial theory, but, contrary to appearances, it turns out not to be Galilean covariant.

4. Einstein's Objections to Emission Theories

Einstein's analysis of 1912

Einstein abandoned his efforts to find an emission theory of light. The reasons seem not to relate to defects in one or another particular emission theory. Rather they derive from a concern that an emission theory of light must contradict some quite secure properties known empirically for light. Einstein gives us his most extensive expression of these concerns in 1912 in his correspondence with Ehrenfest and also in a long unpublished manuscript on special relativity. Our problem will be to try to

²⁵ Might an escape lie in the fact that \mathbf{A} and φ need only be determined up to a gauge transformation, so that we can be more lenient in the transformations allowed, as long as the measurable \mathbf{E} and \mathbf{H} fields conform to Galilean covariance? The escape fails since whatever transformation we might envisage for the potentials, it must return observable fields that conform to the transformations (5) for the fields \mathbf{E} and \mathbf{H} used in their definition and those transformations includes the fatal transformation $\mathbf{H}=\mathbf{H}'$.

disentangle which of the objections to an emission theory played a role in Einstein's thought prior to his 1905 paper and which were now being advanced by Einstein in 1912 as a contribution to the then current debate over emission theories—although I will conclude it cannot be done cleanly.

In a long unpublished exposition of special relativity written in 1912-1914, Einstein (1912-14, pp. 35-35; translation Beck, 1996, p. 26) Einstein explained why an emission theory of light would be unsatisfactory:

[In one possibility] the velocity of light in [the medium of Fizeau's experiment, which measures the velocity of light in moving water] M depends on the velocity of motion of the light source with respect to M (Ritz [deleted: and Ehrenfest]). This being so, light rays of all possible propagation velocities, arbitrarily small or arbitrarily large, could occur in M . Intensity, color, and polarization state would not suffice to define a plane light wave; one would have also to add the determinative element of *velocity*, which, however, should not make itself felt in any effects of the first order (which would be proportional to the first power of velocity of the light source). For the light coming from stars that are in motion relative to the Earth has—as far as our experience extends—the same properties as the light from terrestrial sources of light. To do justice to that, one is forced to make the most peculiar assumptions if one pursues this point of view, as for example the following: if light of velocity $c+v$ strikes a mirror perpendicularly, then the reflected light has the velocity $c-v$. These complications make it seem understandable why it has not proved possible so far to set up differential equations and boundary conditions that would do justice to this conception.--

The concerns described here are a digest of issues raised in the 1912 exchange between Einstein and Ehrenfest following the latter's publication of Ehrenfest (1912). The main import of Einstein's first reaction (Einstein to Ehrenfest, 25 April 1912, *Papers Vol. 5*, Doc. 384) was to suggest to Ehrenfest that Ritz's conception was open to simple experimental test. The test depended on which of two cases was assumed.

In the first case, one might assume that light from a moving source retains the motion of the source when it passes through a medium at rest or is reflected from a substance at rest. For this case, Einstein observed, the wavelength of the light would be unaffected by the motion of the source, but the frequency would be affected. So a Doppler shift would not be experimentally detected by devices that measure wavelength directly (such as diffraction gratings); but it would be detected by processes that measure the frequency directly. Here he named dispersion processes that depend on resonance.²⁶ In the

²⁶ Einstein's intent is clear. If light emitted from a source at rest has the waveform $f(\mathbf{k}\cdot\mathbf{r}-\omega t)$, then the effect of a velocity \mathbf{v} of the source is to boost the waveform according to the Galilean rule $\mathbf{r} \rightarrow \mathbf{r} - \mathbf{v}t$, so that the waveform becomes $f'(\mathbf{k}\cdot\mathbf{r}-(\omega+\mathbf{k}\cdot\mathbf{v})t)$. The boost has left the wave number \mathbf{k} unaffected, but the frequency has been altered (Doppler shifted) from ω to $\omega+\mathbf{k}\cdot\mathbf{v}$. Diffraction gratings form interference patterns by reassembling light that has followed paths of different lengths to the observing screen, so the resulting patterns depend only on the wavelength of the light and will not reveal the Doppler shift in this case

usual understanding, such as supplied by relativity theory, since wavelength λ and frequency ν are always related by $c=\lambda\nu$, with c constant, a Doppler shift in frequency can only arise if there is a corresponding Doppler shift in the wavelength.

In the second case, light that interacts with matter is transformed so that it loses the motion imprinted by the moving source; presumably it adopts the velocity c with respect to the intervening matter. Einstein proposed an experiment that would reveal this effect. One of two coherent light rays from a moving star would pass through a foil. The effect of the foil on the velocity of one ray would be evident in a phase difference between the two rays.

In the exchanges that followed, misunderstandings were resolved. To get the result Einstein reported in the first case, it was essential that, if light from a source moving at v attains velocity $c+v$, it must retain this velocity in all directions upon deflection. This is *not* the case described above in the 1912-14 manuscript in which light with velocity $c+v$ reflecting as light with velocity $c-v$. So when Ehrenfest mistook this last rule for the first of the two cases, it took a few exchanges to resolve the matter. (Einstein to Ehrenfest, 2 May 1912, Doc. 390; Ehrenfest to Einstein, after 16 May, 1912, Doc. 394; Einstein to Ehrenfest, 3 June 1912, Doc. 404; all in *Papers*, Vol. 5.) In his June 3 letter, Einstein explained that Ehrenfest's alternative would still have untenable, observable consequences: the different velocities of incidence and reflection would now mean that angles of incidence and reflection would no longer be equal. However Einstein also needed to concede to Ehrenfest (Doc. 409, *Papers*, Vol. 5 "before June 1912") that there would be no first order effect in some experiment involving reflection.²⁷

Einstein's letter of June 1912 identifies one further problem for an emission theory of light. Einstein wrote:

In support of the independence of the speed of light from the state of motion of the light source one can, of course, quote its simplicity and ease of realization. As soon as one gives up this hypothesis, then, even to explain shadow formation, one must introduce the ugly assumption that light emitted from a resonator depends on the type of excitation (excitation through "moving" radiation or excitation of another kind).

Einstein's point is hard to interpret. Shadow formation is usually the province of simple, geometric optics, with diffraction at hard edges handled by Huygens constructions. It is hard to see how re-radiation from resonators could be involved unless Einstein is considering the shadows cast by semi-transparent bodies. Their transparency depends upon the frequency of the incoming light not arousing resonant responses in the atoms of the bodies, with these atoms modeled as resonators. Perhaps the point is that, in an emission theory of light, whether a pane of glass is transparent to light or casts a shadow

(assuming the speed of light remains isotropic). Light interacts with suitable resonators, such as bound charges, according to the light's frequency. So dispersion phenomena that depend on the latter will be sensitive to the changes in frequency and will reveal this Doppler shift.

²⁷ It is not clear to me to which experiment he referred. In the proposed experiment of the second case of his letter of April 25, Einstein had claimed an effect that depends on the foil-screen distance in quantities of first order, but this experiment did not involve reflection.

would no longer depend on just the wavelength and polarization of the light, but the relative velocity between the source and the glass as well.

How much of these considerations played a role in Einstein's evaluation of emission theories prior to the 1905 paper? How much were later elaborations for the debate of 1912? This last letter gives the answer. In the full passage quoted partially above in Section 3, Einstein wrote:

I was not annoyed in the least by your article. On the contrary. Such considerations are quite familiar to me from the pre-relativistic time. I certainly knew that the principle of the constancy of the velocity of light is something quite independent of the relativity postulate; and I considered what would be more probable, the principle of the constancy of c , as was demanded by Maxwell's equations, or the constancy of c , exclusively for an observer sitting at the light source. I decided in favor of the first, since I was convinced that each light [ray] should be defined by frequency and intensity alone, quite independently of whether it comes from a moving or a resting light source. Moreover it did not occur to me to consider whether the radiation deflected at a point could behave differently in propagation compared to newly emitted radiation from the point concerned. Such complications seemed to me far less justified than those brought by the new concept of time.

The decisive consideration, Einstein tells us, that spoke to him against an emission theory prior to his 1905 paper was his conviction that light should be characterized by frequency and intensity (and polarization) alone. He was then rather uninterested in the fussy details of how a variety of distinct emission theories might be devised to accommodate to various sorts of processes of deflection or reflection. They seem to have come to the fore in the literature emerging around 1912 that sought to test an emission theory experimentally, for just those details decide how the experiments are to be done. (See Tolman, 1912, for example.)

Later remarks

Later remarks augment the comments from 1912, but not always univocally. The earliest of them contradicts Einstein's 1912 assertion of lack of interest in specific hypotheses about deflected radiation. He wrote to Mario Viscardini (April 1922, Einstein Archive, 25-301; translation in part from Rynasiewicz, 2000, p. 182) in direct continuation of the part quoted above:

I rejected this [emission] hypothesis at the time, because it leads to tremendous theoretical difficulties (e.g., the explanation of shadow formation by a screen that moves relative to the light source). However the Dutch astronomer de Sitter has given the most convincing refutation of this hypothesis, in that he pointed out that the light from a component of a double star must be emitted with a time changing speed that is absolutely not in agreement with what is given by observation.

Another repeats earlier remarks. In a 1952 draft written on the back of a letter from A. Rippenbein, Einstein wrote (Einstein Archive, 20-040; translation based on Stachel, 2002, p. 189):

Your attempt to replace special relativity with the assumption that the velocity of light is constant *relative to the source of light* was first advocated by Ritz. This assumption is

compatible with Michelson's experiment and with aberration...[Einstein then refers to De Sitter's refutation of Ritz's theory]. In addition this theory requires that light waves with different velocities of propagation shall be possible everywhere and in each definite direction. It would be impossible to set up any sort of reasonable electromagnetic theory that accomplishes this. This is the principal reason why, even before setting up the special theory of relativity, I rejected this way out, although it is intrinsically conceivable."

In his Shankland (1963, p. 49) report of discussion with Einstein in the early 1950s, Shankland described a new objection by Einstein to an emission theory. It would allow light phases to get mixed up and for light even to reverse itself:

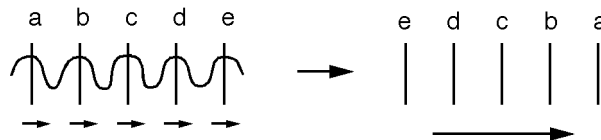
... he told me that he had thought of, and abandoned the (Ritz) emission theory before 1905. He gave up this approach because he could think of no form of differential equation which could have solutions representing waves whose velocity depended on the motion of the source. In this case, the emission theory would lead to phase relations such that the propagated light would be all badly "mixed up" and might even "back up on itself." He asked me, "Do you understand that?" I said no and he carefully repeated it all. When he came again to the "mixed up" part he waved his hands before his face and laughed, an open hearty laugh at the idea!

Then he continued, "The theoretical possibilities in a given case are relatively few and relatively simple, and among them the choice can often be made by quite general arguments. Considering these tells us what is possible but does not tell us what reality is."

Presumably mere repetition along with some gymnastic hand waving did not help Shankland figure out what Einstein intended. Fortunately another letter by Einstein from the same time adds the missing piece that makes sense of it all. Einstein's above quoted response to Hines from February 1952 continues (Einstein Archive, 12-250, 12-251):

...Then one cannot trace light propagation back to differential equations, but one must introduce "retarded potentials," which is a kind of action at a distance.

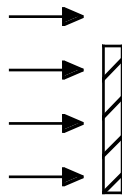
Before setting up the special theory of rel., I had myself thought of investigating such a possibility. At that time I had only a weighing of the plausibility of theoretical arguments at my disposal. I did not then think of the use of the evidence on double stars. I deliberated as follows: If a suitably accelerated light source emits light in one direction (e.g., the direction of the acceleration), then the planes of equal phase move with different speed, and one can set it up like this,



so all the surfaces of equal phase coincide at a particular place, so that the wavelength there is infinitely small. Moreover the light will be so turned around that the rear part overtakes the front.

If we imagine a light source to be accelerating sufficiently rapidly in the direction of the emitted light, then the increasing phase velocity may allow light emitted later to catch up with light emitted earlier and then overtake it, so the light emitted later arrives before that emitted earlier. If one programs the acceleration just right, all the waves will catch up at just the same moment, producing a superposition of continuously many waves at one point. That some sort of singular behavior arises is evident, but it is not clear to me why Einstein characterized it as an infinitely small wavelength. My supposition is that the emitted light waves all have the same wavelength in an emission theory but different frequencies, so the singularity would be in the frequency at the point in question. Einstein continued, repeating a description of the experiment he proposed to Ehrenfest in 1912 and returning to familiar themes:

Further an ever so thin, diaphanous film will change the speed of “moving” light by a finite contribution, so that interference, e.g. in the case



would give rise to quite incredible phenomena.

But the strongest argument seemed to me: If there is no fixed velocity for light at all, then why should it be that all light emitted by “stationary” bodies has a velocity *completely independent of the color*? This seemed absurd to me. Therefore I rejected this possibility as a priori improbable.

De Sitter’s argument concerning emission from double stars is—as far as I can see—sufficient by itself as contrary evidence.

The last remark was apparently responding to Hines’ remark that double star observations provided the only objections properly raised against Ritz’s optical theories and, Hines felt, “even that may be explained away.”

A puzzle: why no differential field equations for an emission theory?

This collection of remarks by Einstein on the inadmissibility of an emission theory of light clearly mixes objections conceived prior to his 1905 paper with those developed later when he discussed the issue of empirical testing of emission theories. I find it quite plausible that Einstein’s deliberations prior to the 1905 paper did not depend much on considerations of particular hypotheses on how deflected light might move (just as Einstein writes to Ehrenfest in June 1912, above). The mere fact that many velocities for light were possible seemed incompatible with observations: accelerating sources might lead to light reversing itself, a phenomenon never seen; and Einstein concluded from it that there could be no field theory based on differential equations for light. But I also find it plausible that Einstein may not have had a perfect memory of deliberations that were undertaken years before and possibly never committed to writing; and that in his later recollections and reports in letters he might be less concerned to lay out an

accurate report as a resource for historians of what he had thought and when, rather than to convince a possibly argumentative correspondent of the untenability of an emission theory.

Instead of continuing to try to sort out just what might have belonged to which time, I want to point out the puzzling character of the principal thread. Einstein's repeated concern with an emission theory is that there seems to be no way to formulate it in a field theory based on differential equations.²⁸ One reason given is that an emission theory allows waves of different velocities. A light wave of velocity $c+v$ can reflect as one of $c-v$. Light from accelerating sources can overtake and even form singular points. This "many velocities" argument is simply not cogent. It takes very little effort to find differential equations that admit just this behavior for waves. They are now quite familiar to us from quantum theory, for example. Both the ordinary Schrödinger equation and the Lorentz covariant Klein-Gordon equation admit waves with many different phase velocities. Since they are linear equations, we can readily construct fields consisting of the superposition of many waves propagating at different velocities. We will even find Einstein's example of overtaking waves with velocities contrived so that a singular point momentarily forms. We should not be so troubled by such points. They are otherwise known in analogous cases in optics as caustics and are not regarded as fatal to our present wave theories of light.

So perhaps Einstein was just hasty and blundered. Before we accept that possibility, I want to recall the other remark he makes repeatedly about emission theories: that such a theory makes it impossible to characterize light solely by the usual parameters of intensity, color and possibly polarization. That remark, which has so far been uninterpreted, seems decisive to me. For I believe that it is impossible to give an electromagnetic field theory specified by differential equations of the type familiar to us that is: (a) an emission theory of light; (b) Galilean covariant, even with field transformation laws; and (c) characterizes light waves by intensity, color and polarization alone. This claim might seem to need some significant computation for support. It turns out not to. There is a very simple thought experiment that makes it transparent. That thought experiment is primally attached to Einstein's name and to the discovery of special relativity.

5. Einstein Chases a Light Beam

A thought experiment of unclear import

After yielding to the "invitation and earnest request" and "quite some persuasion" of its editor (Einstein, 1949, preface), in 1946, Einstein put his autobiographical reminiscences of a life in science on paper for the volume *Albert Einstein: Philosopher-Scientist*. On other occasions, he had stressed the

²⁸ This objection has entered the standard lore. Becker (1964, p. 313) writes: "...the [Ritz hypothesis of an emission theory] is completely untenable from the *theoretical* standpoint of a field concept which describes the motion of light by a differential equation, because it cannot be understood how the velocity of propagation of light from a source located at a point of space should be related to the condition of the light source." Alas, no justification is given.

importance of the magnet and conductor thought experiment in bringing him to special relativity. In this account, however, it is not mentioned. Instead another thought experiment is given the central role. Einstein reformulated the problem as the search for a universal formal principle akin the principles of thermodynamics. He continued (p. 49-50)

After ten years of reflection such a principle resulted from a paradox upon which I had already hit at the age of sixteen: If I pursue a beam of light with the velocity c (velocity of light in a vacuum), I should observe such a beam of light as an electromagnetic field at rest though spatially oscillating. There seems to be no such thing, however, neither on the basis of experience nor according to Maxwell's equations. From the very beginning it appeared to me intuitively clear that, judged from the standpoint of such an observer, everything would have to happen according to the same laws as for an observer who, relative to the earth, was at rest. For how should the first observer know or be able to determine, that he is in a state of fast uniform motion?

One sees in this paradox the germ of the special relativity theory is already contained. Today everyone knows, of course, that all attempts to clarify this paradox were condemned to failure as long as the axiom of the absolute character of time, or of simultaneity, was rooted unrecognized in the unconscious.

This thought experiment has proven immensely popular in accounts of the discovery of special relativity. Who could not fail to be charmed by the image of a precocious sixteen year old whose innocent imaginings lay the groundwork for a great discovery?

What is rarely mentioned, however, is that the thought experiment does not quite make sense.²⁹ Usually Einstein's thought experiments are models of lucid argumentation, quite compelling to the

²⁹ Banesh Hoffmann (1982, pp. 93-97), a former collaborator and biographer of Einstein, is one of the few to discuss the problem. His explanation is that the sixteen year old Einstein tacitly applied the principle of relativity to rule out frozen light and that some kind of subconscious, psychological block precluded him conceiving that light might even be slowed by the motion of an observer. Darrigol (1996, pp. 289-90) resolves the problem by suggesting that we simply doubt the veracity of Einstein's recollection since Einstein believed in the existence of the ether as late as 1901. He concludes, "we should therefore regard the wide-spread belief that Einstein had an inborn trust in the relativity principle as a myth." This reaction seems to me over eager to dismiss a recalcitrant historical datum. Einstein's recollection does not require an inborn trust in the relativity principle at age 16. It merely requires that he, as an ether theorist, be disturbed by one consequence of his theory, that it allows a moving observer to catch light. Surely we must allow a precocious 16 year old a visceral reaction that something is wrong with one consequence of his theory without demanding that he then discard the theory or even have a cogent argument behind the reaction. However I do agree with Darrigol that we should not allow the story of the premonitions of a 16 year old to invest him with mysteriously prescient powers, as is the wont of popularizers. We cannot say what significance is to be accorded to such anticipations until we know what other anticipations were felt

reader. But this one, if it intended to be an objection to an ether theory of light, does not bear scrutiny. Using ordinary kinematical notions, if we sped fast enough through the ether after a beam of light, we would bring it to rest. The fact that we do not experience such a thing merely reflects the fact that we happen not to be moving very fast through the ether. What of Maxwell's equations? Using the standard kinematical notions associated with an ether theory, it is trivial to show that rapid motion would bring the light to rest. Maxwell's equations (M1)-(M4) admit plane wave solutions of frequency ω propagating in the direction of the wave number vector \mathbf{k}

$$\mathbf{E} = \mathbf{E}_0 \sin(\omega t - \mathbf{k} \cdot \mathbf{r}) \quad \mathbf{H} = \mathbf{H}_0 \sin(\omega t - \mathbf{k} \cdot \mathbf{r}) \quad (15)$$

where the field vectors are transverse to the direction of propagation ($\mathbf{k} \cdot \mathbf{E} = \mathbf{k} \cdot \mathbf{H} = 0$) and orthogonal to each other ($\mathbf{k} \times \mathbf{E} / |\mathbf{k}| = \mathbf{H}$); and the frequency ω and wave number vector \mathbf{k} relate as $\omega = |\mathbf{k}|c$. We transform from the unprimed ether frame (t, \mathbf{r}) to a primed frame (t', \mathbf{r}') using $t = t'$ and $\mathbf{r} = \mathbf{r}' + \mathbf{v}t'$, where the condition that the velocity \mathbf{v} matches the velocity of propagation of the wave entails³⁰ $\mathbf{k} \cdot \mathbf{v} = \omega$. The wave in the new frame is

$$\mathbf{E}' = \mathbf{E}'_0 \sin(-\mathbf{k} \cdot \mathbf{r}') \quad \mathbf{H}' = \mathbf{H}'_0 \sin(-\mathbf{k} \cdot \mathbf{r}') \quad (16)$$

since the argument in the waveform is $\omega t - \mathbf{k} \cdot \mathbf{r} = \omega t' - \mathbf{k} \cdot (\mathbf{r}' + \mathbf{v}t') = (\omega - \mathbf{k} \cdot \mathbf{v})t' - \mathbf{k} \cdot \mathbf{r}' = -\mathbf{k} \cdot \mathbf{r}'$. (The primes on the field vectors allow for the possibility of some sort of transformation of the field quantities. We need not concern ourselves too much with that possibility since we merely seek to describe the waveform in another frame but not establish any sort of Galilean covariance for the laws governing the descriptions.) The wave has been brought to rest; it is a frozen sine wave ("spatially oscillating").

The doubts of a sixteen year old

We can quickly resolve another apparent problem with the thought experiment. While Einstein attaches the thought experiment to himself as a sixteen year old, the mention of Maxwell's equations is unlikely to come from that time. He was then a high school student and the formal study of Maxwell's theory did not come until his university studies, around 1898 (see "Einstein on the Electrodynamics of Moving Bodies," pp. 223-35 in *Papers*, Vol. 1). Einstein's wording in his *Autobiographical Notes*, clearly allows the original thoughts of the 16 year old Einstein to gestate and develop over a ten year period and we should surely imagine Maxwell's equations entering the reflection in the course of this process. Or if Maxwell's equations played any part in the 16 year old's reflections, they were hesitant and incomplete, reflecting his own incomplete knowledge of Maxwell's equations.

That the original thought experiment was more a visceral reaction of disbelief to a striking possibility, frozen light, is suggested by the two other accounts we have of the thought experiment. Both are more clearly localized to a young Einstein at age 16 in the year 1895-96 he spent at a high school (gymnasium) in the Canton of Aarau, Switzerland. In 1916, the psychologist Max Wertheimer met with

by the young Einstein but came to nothing. A poor predictor can be made into a profound seer by selective reporting of his few successes.

³⁰ A point of constant phase moves according to $\mathbf{k} \cdot \mathbf{r} = \omega t$. Differentiate with respect to t to recover $\mathbf{k} \cdot \mathbf{v} = \omega$.

Einstein in order to probe the psychology of his work. Later, in a work first published posthumously in 1945 after his death in 1943, Wertheimer (1959, pp. 214-15) reported:

The problem began when Einstein was sixteen years old, a pupil in the Gymnasium (Aarau, Kantonschule)...

The process started in a way that was not very clear, and is therefore difficult to describe—in a certain state of being puzzled. First came such questions as: What if one were to run after a ray of light? What if one were riding on the beam? If one were to run after a ray of light as it travels, would its velocity thereby be decreased? If one were to run fast enough, would it no longer move at all? ...[W's ellipses] To young Einstein this seemed strange.

...When I asked him whether, during this period, he had already had some idea of the constancy of light velocity, independent of the movement of the reference system, Einstein answered decidedly: "No, it was just curiosity. That the velocity of light could differ depending upon the movement of the observer was somehow characterized by doubt. Later developments increased that doubt."

The same thought experiment is recounted in an autobiographical sketch written shortly before Einstein's death. While reminiscing of his year at the gymnasium in Aarau, he wrote (Einstein, 1956, p. 10)

During this year in Aarau the following question came to me: if one chases a light wave with the speed of light, then one would have before one a time independent wave field. But such a thing appears not to exist! This was the first child-like thought experiment related to the special theory of relativity. Discovery is not a work of logical thought, even if the final product is bound in logical form.

Both these reports portray the thought experiment in its original form as much less of the well reasoned and polished display pieces that characterize Einstein's scientific writing. Rather it was more an account of a psychological event in his personal history, the instinctive sense of a sixteen year old Einstein that something was not right in a theory that allows light to be brought to rest. This is surely the import of Einstein's insertion of a remark ("Discovery is not a work of logical thought..."), in the midst of cozy recollections of a happy year of his youth. They recall a familiar theme in his more philosophical writing, that discovery in science inevitably involves creative steps that cannot be reconstructed logically.

6. The Importance of the Thought Experiment

A proposal

The instincts of a sixteen year old Einstein were just the first step. Einstein's recollection in his *Autobiographical Notes* clearly accords the light chasing thought experiment an importance that endures well into the work that eventually yields special relativity. The mention of Maxwell's equations tells us that it was invoked when he knew Maxwell's equations and that part of its use involved Maxwell's equations. Further, its function in the narrative of *Autobiographical Notes* is to bridge the gap between the

early sensing of a problem to do with relative motion through to his insight into the relativity of simultaneity.

My goal in this section is to propose an account of how such broad importance might be possible so that thought experiment merits pride of place in an autobiographical statement that he recognized would be definitive. (He called it, “my obituary”(p.3).) Beyond sowing the initial seed of doubt in his sixteen year old mind, I do not think we will be able to identify a single result in which the primary importance of the thought experiment can be found. If that were the case, Einstein would likely have mentioned it in his own accounts; and, even if such an import were there, we have too meager resources to locate it. Instead, what we can do is map out a period we know occurred in Einstein’s research in which the thought experiment could be peculiarly effective in advancing his investigations. This was the time when Einstein weighed emission theories of light and found them wanting. Below I will propose three ways in which the thought experiment could be used to cast doubt on an emission theory of light in ways compatible with Einstein’s remarks.

More generally, however, I suggest that the thought experiment reflects the standard way that Einstein investigated the interaction of light and the motion of an observer in the years leading up to the 1905 paper: the effects of that motion were to be understood in terms of how it affected the waveform of the light, just as the 16 year old Einstein had done. Such deliberations gave mounting reasons to doubt ordinary kinematical assumptions. I will suggest in Section 8 that our present fascination with light as a clock synchronizing signal as opposed to light as a propagating electromagnetic waveform may reflect only a very small part of the thinking that led him to special relativity.

Problems for an emission theory of light

Let us place Einstein at a time in which he is entertaining the possibility of an emission theory of light that conforms to the principle of relativity. Any such theory would allow a sufficiently rapidly moving observer to catch light.

Observational problem: “There seems to be no such thing, however, neither on the basis of experience...”

The ether theorist’s rejoinder was that we do not see such a thing since we happen not to be moving that fast through the ether. In an emission theory that is no longer as viable an objection. All that is required for us to see light frozen or significantly slowed in its propagation is for there to be a light source moving rapidly enough away from us. The ether theorist’s explanation required slow motion for just one entity, our earth. The emission theorists’ explanation requires slow motion for many more: every object that transmits light to us. If we allow independence in their motions, the chances that this can be the case decrease exponentially with the number of objects considered.

Theoretical problem: “...nor according to Maxwell’s equations.”

An emission theory of light must still harbor propagating light waves, such as the sinusoidal waves (15) or something with similar undulations—that is one of our most familiar observations concerning light. In an emission theory of light, an inertially moving observer would be able to observe the frozen fields (16) or something with similar undulations. Because the theory is supposed to satisfy the principle of relativity, these fields must also satisfy the basic equations of the theory, an additional

condition the ether theorist did not need to meet. That is, the two sinusoidal field (16) or something like them must be added to the repertoire of fields admissible in electrostatics and magnetostatics. But that is a very dubious outcome. While we may not know the precise form that an emission theory might take, it is surely reasonable to expect it to agree quite closely with older theories in domains in which the older theories have been extensively explored. One of the most secure of these domains is electrostatics and magnetostatics. But such fields are simply unknown in standard 19th century electrostatics and magnetostatics. All known static fields are well captured by the condition that they be time independent and solutions of Maxwell's equations (M1) and (M3) and they do not admit the sort of spatially oscillating fields characterized by (16). In short, the fact that Maxwell's equations do not admit a spatially oscillating field at rest is a serious problem for an emission theory of light that satisfies the principle of relativity.

Failure of differential equations: "For how should the first observer know or be able to determine, that he is in a state of fast uniform motion?"

Let me ask Einstein's question in a slightly narrower fashion. Imagine that you are given the instantaneous state of a waveform. How are you to determine whether it belongs to a waveform speeding past you (i.e. you are "at rest"); or whether you share the waveform's motion (i.e. you are "in a state of fast uniform motion")? The answer is routinely given by the field equations of the theory in question. Take for example Maxwell's equations (M1)-(M4). They are first order partial differential equations. As a result, the state of fields at some instant specified how they will develop in time. You feed in the \mathbf{E} and \mathbf{H} fields at some instant; use the field equations to read off their first derivatives; and finally integrate the first derivatives to find the time development of the fields.

There is no reason to doubt that a similar procedure would be possible for an emission theory until we add in one condition repeatedly mentioned by Einstein as decisive in his evaluation of emission theories: a plane light wave must be fully specified by its intensity, color and, possibly, polarization. What that means is that the waveform is given by (15). The intensity is given by magnitude of the vectors \mathbf{E}_0 and \mathbf{H}_0 . The color is given by one or another of the frequency ω or (scalar) wave number k , where fixing one fixes the other, since they are related by $\omega=kc$. The polarization is determined by the direction of the vectors \mathbf{E}_0 and \mathbf{H}_0 . (The waveform (15) is the simplest case of linear polarization; circular and elliptical polarization arises when waveforms of the same frequency but different phases are superimposed.) It is of course entirely understandable why Einstein would insist on these few parameters being all that is needed to characterize a light wave. As Einstein explains in the 1912-14 manuscript (as quoted above), it is an experimental matter; these are the only properties that have been found, even after light from both celestial and terrestrial sources has been subject to extensive experimental investigation.

Now the problem is acute. For the two waveforms are identical at an instant. The waveform that speeds past, (15), evaluated at $t=0$, is identical with the frozen waveform, (16). If the theory is deterministic, given that waveform as its initial state, only one future time development will be possible according to the field equations. So the field equations will be unable to determine correctly whether the waveform is destined to propagate rapidly as in (15) or whether it is frozen as is (16) and represents the wave for all time. For field laws expressed as differential equations to capture an emission theory of light, the waveform must have some distinguishing characteristic encoded within it, so that the initial states

presented to the field equations are different in the case of propagating or frozen light. That distinguishing characteristic would enable the field equations to return different time developments. The motion of its source must somehow be imprinted onto light as an additional property. But just such an additional property is what Einstein denies exists on the basis of experimental evidence.

This argument for the failure of any formulation in terms of differential equations is quite general and quite robust. I have stated it as an objection to any set of first order differential equations, they being the natural choice if one seeks laws modeled after Maxwell's equations. One might well wonder if higher order differential equations would be able to capture an emission theory, such as second order differential equations. The idea would be that the initial state supplied to the second order differential equations would be the fields \mathbf{E} and \mathbf{H} along with their first time derivatives, $\partial\mathbf{E}/\partial t$ and $\partial\mathbf{H}/\partial t$. While the propagating and frozen waveforms may agree on \mathbf{E} and \mathbf{H} at the instant in question, the motion of the first would be encoded in non-zero time derivatives not manifested by the second. But the very idea of fields that agree in their intensities but not their derivatives is precluded, exactly because these first derivatives encode an additional property that makes a physical difference. The first derivatives would add an additional physical parameter exactly of the type denied by Einstein. This problem could not be escaped with differential equations of still higher order.

Or perhaps as a last effort, one might look to field transformation laws of the type (4) and (5). The idea would be that these transformations alter the relative magnitude of the \mathbf{E} and \mathbf{H} vectors and that the velocity of the wave's source might somehow be imprinted in differences of those ratios. If the transformation of the wave from (15) to (16) is accompanied by field transformation (5), then the effect is to annihilate the transformed wave's \mathbf{E}' field entirely, while leaving the \mathbf{H}' field unchanged. Might we use this to distinguish propagating wave (15) from the frozen wave (16)? Once again, the idea founders on the limited number of properties allowed a light wave. In addition to intensity, color and polarization, one would have to add another property representing different ratios of the magnitude of the \mathbf{E} and \mathbf{H} fields. (Such difference would surely be evident in experiments measuring interference; waves with different ratios of field intensities would not annihilate in the same way.)

Finally, it is instructive to review an example that shows that the difficulty is not merely forming differential equations that satisfying the principle of relativity and that admit waves with their velocity of propagation somehow encoded within. That by itself is easily achieved. Take the Lorentz covariant Klein Gordon equation for some particle of imaginary mass, im . Setting the units conveniently ($c=1$, $\hbar/2\pi=1$) and taking the simplest case of one spatial dimension, the equation is

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - m^2 \right) \varphi = 0 \quad (17)$$

Its plane wave solutions are $\varphi = \exp i(\omega t - kx)$ where $m^2 = k^2 - \omega^2$. Since the speed of propagation is just $v = \omega/k$, waves with speeds from zero to arbitrarily close to one are admissible. The case of the wave of zero velocity concerns us. It is just $\varphi = \exp i(mx)$ or $\varphi = \exp i(-mx)$. If this is supplied as an initial condition to the Klein Gordon equation (17), the equation will tell us that this wave does not propagate. That can happen because the wave number $k=m$ encodes the crucial velocity information. The condition

$m^2 = k^2 - \omega^2$ guarantees that any wave with exactly the wave number $k = m$ is a frozen wave. No propagating wave can bear that wave number. No analogous escape is possible for an emission theory of light. What is essential is that the particle of imaginary mass m of this Klein Gordon equation (17) has the characteristic parameter m for which there is no analog in light.

This example might also explain a remark Einstein makes in his draft letter to Hines in 1952: "...why should it be that all light emitted by "stationary" bodies has a velocity *completely independent of the color*? This seemed absurd to me...." It follows immediately from the above that the speed v of the waves admitted by this Klein Gordon equation satisfies $v^2 = (\omega/k)^2 = 1 - (m/k)^2$. That is, there is a fixed relationship between the speed of the wave and its wavelength k , which also fixes its color. Presumably Einstein had hoped that the color of a light wave in an emission theory would encode its velocity, analogously. Evidently light does not admit such encoding since, empirically, there seems to be no connection between the colors possible for light and the velocity of its source.

This completes the argument, associated with the chasing the light thought experiment, that demonstrates that no field law expressed in differential equations can be (a) an emission theory of light; (b) Galilean covariant, even with field transformation laws; and (c) characterize light waves by intensity, color and polarization alone.

7. Stellar Aberration and Fizeau's Experiment

The optical experiments that mattered to Einstein

In spite of the lore that places the Michelson-Morley experiment at the center of Einstein's thoughts as he worked on the problems that would lead to special relativity, it has long been known that the experiment played only a small role. See Stachel (1987) for a recent appraisal that draws on the correspondence between Einstein and Mileva Maric brought to light in the 1980s. Einstein did however give a clear statement of the optical experiments that were significant to him. In the note written in Michelson's honor and quoted from in Section 2 above, Einstein had first written (Einstein Archive 1-168):

I am not sure when I first heard of the Michelson experiment or its more precise repetition by Michelson and Morley. I was not conscious that it influenced me directly during the seven and more years that the development of the Special Theory of Relativity had been my entire life; for I had taken it for granted as being true.

That English typescript had been struck out and replaced by the handwritten remark (quoted above) that he learned of the experiment through reading Lorentz' 1895 *Versuch*. It led up to the reaffirmation of the importance of the magnet and conductor thought experiment and ended in a statement of the optical experiments that *did* matter to him:

But the result of Fizeau's experiment and the phenomenon of aberration also guided me. Similar remarks were also reported by Shankland, with a little dismay. He had first approached Einstein "to learn from him what he really felt about the Michelson-Morley experiment, and to what degree it had

influenced him in his development of the Special Theory of Relativity.” (Shankland, 1963, p.47) Instead he found Einstein telling him that he only learned of the experiment after 1905:

“Otherwise,” he said, “I would have mentioned it in my paper.” He continued to say the experimental results which had influenced him most were the observations of stellar aberration and Fizeau’s measurements on the speed of light in moving water. “They were enough,” he said. ... when I added that it seemed to me that Fizeau’s original result was only qualitative, he shook his pipe and smiled, “Oh, it was better than that!”

In a follow up report, Shankland (1973, p. 896) once again recounted how Einstein sought to direct discussion away from ether drift experiments to stellar aberration and Fizeau’s experiment:

...Prof. Einstein volunteered a rather strong statement that he had been more influenced by the Fizeau experiment on the effect of moving water on the speed of light, and by astronomical aberration, especially Airy’s observations with a water filled telescope, than by the Michelson-Morley experiment.

It is not hard to understand the small role Einstein accorded to the Michelson-Morley experiment. Stachel (1982, p.179) made the decisive remark that, in all later accounts in which it is mentioned, Einstein offered the experiment as evidence for the principle of relativity, not the light postulate, as is now the routine practice. If this was Einstein’s attitude prior to 1905, then the experiment would be relevant only at the earliest stages of Einstein’s explorations. Einstein had decided that the principle of relativity must hold by the time of the magnet conductor thought experiment. From then on, the Michelson-Morley experiment could be of little assistance. Its null result was a foregone conclusion, entailed by the principle of relativity. It could not assist Einstein in deciding between “the principle of the constancy of c , as was demanded by Maxwell’s equations, or the constancy of c , exclusively for an observer sitting at the light source [emission theory],” the options he recalled to Ehrenfest that he had weighed. In either case, a null result is expected. An emission theory would only predict a positive result if the experiment used a light source moving with respect to the interferometer.

But what of stellar aberration and Fizeau’s experiment? How could these long known first order experiments be more informative than the newer Michelson-Morley second order experiment? This is the answer I will develop here: If one looks carefully at how both experimental results are accommodated in Maxwell’s electrodynamics by the then standard analysis of Lorentz, one finds that they supply direct experimental evidence for Lorentz’s local time and that they do this essentially independently of Maxwell’s theory. I urge that this is precisely what Einstein realized.

*Lorentz’s theorem of corresponding states*³¹

Lorentz’s (1895) *Versuch* is a *tour de force*. It develops a unified theoretical framework in which to treat electromagnetic and optical processes in moving bodies and to demonstrate that, even though Maxwell’s electrodynamics depends essentially on a resting ether, optical processes will depend only on

³¹ For a account of nineteenth century ether drift experiments and their analysis in Lorentz’s work, see Janssen and Stachel, forthcoming.

relative velocities at least to first order in v/c . The centerpiece of Lorentz's work is his theorem of corresponding states. (pp. 85-86) It amounts to a computationally simple prescription for constructing new solutions of Maxwell's equations from old solutions, where the new solutions represent the old, set into uniform motion. More exactly, it enabled a new solution as close to a uniformly moving copy of the old solution as Maxwell's theory allowed and, in the early versions of the theorem of 1895, the new solution would satisfy Maxwell's equations only up to quantities of the first order in v/c . The theorem reduced a difficult problem, finding solutions of Maxwell's equations for systems in motion, to a much easier one, finding solutions of Maxwell's equations for systems at rest. The basis of the theorem was a transformation for field and other quantities and the coordinates $(t, \mathbf{r}=(x,y,z))$ of the ether rest frame that included:

$$\mathbf{E}' = \mathbf{E} + (1/c)(\mathbf{u} \times \mathbf{H}) \quad \mathbf{H}' = \mathbf{H} - (1/c)(\mathbf{u} \times \mathbf{E}) \quad t' = t - \mathbf{u} \cdot \mathbf{r} / c^2 \quad \mathbf{r}' = \mathbf{r} - \mathbf{u}t \quad (18)$$

where the term $\mathbf{u} \cdot \mathbf{r} / c^2$ gave an artificial-looking place dependence to the time variable t' , so Lorentz (p. 50) called it "local time." To use the theorem, one finds a solution of Maxwell's equations in the ether frame, designating the field and coordinates with primes; the primed quantities are replaced with unprimed quantities by means of the transformation (18); and the resulting formulae are assured by the theorem of being solutions of Maxwell's equations up to first order quantities *in the ether frame*. They represent the corresponding state of the original solution.

This gloss greatly simplifies Lorentz's own development, which included a multistep transformation to arrive at (18). For a thorough treatment, see Janssen (1995, Ch. 3) and also Rynasiewicz (1988). I will continue with this simplified treatment since it contains all that is needed for the point to be made. In similarly reduced fashion, I will review how Lorentz (1895, §37, §§60-78) treated three optical processes: Doppler shift, stellar aberration and the velocity of light in moving media. In each, the transformations of the field quantities turn out to play no role in the final result. All that matters to track the velocity of a wave are the locations at which the field intensity drops to zero; these are unaffected by the field transformations. As a result, I will simply represent a propagating wave as $f(\omega t - \mathbf{k} \cdot \mathbf{r})$, where ω is the frequency and \mathbf{k} the wave number vector and f stands for the multivalued field intensities. I will also suppress accompanying transformations of the field intensities.

Doppler shift. If a source emits a plane light wave in the $+x$ direction, its waveform is $f(\omega t - kx)$. The wave propagates with speed $V = \omega/k$, which need not be c . Setting $\mathbf{u} = (u, 0, 0)$ so we "boost" in the x -direction, the corresponding state is a light wave emitted by a source moving at u in the x -direction. Its waveform is $f(\omega t' - kx') = f(\omega(t - \mathbf{u} \cdot \mathbf{r} / c^2) - k(x - ut))$. That is, the transformation is

$$f(\omega t - kx) \rightarrow f(\omega(1 + u/V)t - k(1 + uV/c^2)x) \quad (19)$$

From it, we read that the velocity boost has increased the frequency from ω to $\omega(1 + u/V)$.

Stellar Aberration. The simplest case arises for starlight arriving from a star lying in a direction perpendicular to the earth's motion. To arrive at the result, assume we have light propagating in the $+y$ direction. Its waveform is $f(\omega t - ky)$. We presume propagation in a vacuum, so we have $c = \omega/k$. We arrive at the same result if we imagine the earth boosted in the $-x$ direction or the star boosted by velocity

$\mathbf{u}=(u,0,0)$ in the $+x$ direction. Take the latter case. The waveform now becomes $f(\omega t' - ky')$
 $= f(\omega(t - ux/c^2) - ky)$. That is, the transformation is

$$f(\omega t - ky) \rightarrow f(\omega t - kux/c - ky) \quad (20)$$

One sees that this transformation has deflected the direction of propagation from the perpendicular toward the direction of motion of the star by the small angle of (u/c) radians.³² This result illustrates the general result for stellar aberration: for first order quantities in (u/c) , the *direction* of the light propagating from a star is given by the *direction* of the vector sum of the velocity of the light and the relative velocity between the earth and the star.

Fizeau's Experiment. In this experiment, Fizeau measured the speed of light in moving water. When the water is at rest, light propagates through it with a speed c/n , where n is the refractive index of the water. When the water moves at speed u in the direction of the light, the question was how much of that speed would be added to that of the light. If the moving water was conceived as fully dragging the ether along with it, then the resulting speed would be $c/n + u$. If the ether was not dragged along at all, the speed would remain at c . Fizeau found experimentally a result compatible with partial dragging of the ether: $c/n + u(1 - 1/n^2)$. This was the result originally proposed by Fresnel for just these circumstances and the coefficient $(1 - 1/n^2)$ is the Fresnel drag coefficient.

While Lorentz's theory supposed a fixed ether that was not dragged at all, he could still recover the Fresnel drag from an application of his theorem of corresponding states. The waveform for light propagating in the $+x$ direction for water at rest is $f(\omega t - kx)$, where $c/n = \omega/k$. For the case of the water moving at $\mathbf{u}=(u,0,0)$, the theorem of corresponding states gives the transformed waveform $f(\omega t' - kx')$
 $= f(\omega(t - ux/c^2) - k(x - ut))$. That is, the transformation is

$$f(\omega t - kx) \rightarrow f(\omega(1 + un/c)t - k(1 + u/cn)x) \quad (21)$$

We read from it that the frequency is Doppler shifted to $\omega(1 + un/c)$ and the wave number to $k(1 + u/cn)$. The speed of propagation is given by their ratio:

$$\frac{\omega(1 + un/c)}{k(1 + u/cn)} \approx \frac{c}{n} + u \left(1 - \frac{1}{n^2}\right) \quad (22)$$

where the equality holds to quantities first order in u/c . This last result is an astonishing triumph for Lorentz's theory. The velocity that light adopts in moving water results from a massively complicated interaction between charged ions of the moving water and the light wave propagating in the spaces in between. Yet all Lorentz needs to recover the experimentally observed effect is his theorem of corresponding states and a small amount of algebra on the space and time variables in the argument of the waveform.

³² To see this, note that a surface of constant phase for the deflected wave is given by the constancy of $(\omega t - kux/c - ky) = k(ct - (u/c)x - y) = k(ct - \mathbf{b} \cdot \mathbf{r})$, where $\mathbf{b}=(u/c, 1, 0)$ is a vector normal to the surface. For $(u/c) \ll 1$, this vector \mathbf{b} is deflected (u/c) radians from the $+y$ direction towards the $+x$ direction. To see that \mathbf{b} is normal to the surface, note that for any two points on the surface \mathbf{r} and \mathbf{s} , we have $\mathbf{b} \cdot (\mathbf{r} - \mathbf{s}) = 0$, which just asserts that \mathbf{b} is orthogonal to any vector $(\mathbf{r} - \mathbf{s})$ lying in the surface.

Einstein abandons an emission theory of light. What now?

The decision to abandon the search for an emission theory of light would have created great difficulties for Einstein in his efforts to realize the principle of relativity in electrodynamics. Without later insights into simultaneity, a simple consequence of the principle of relativity is that the velocity of a source must be added to the velocity of the light. Yet no adequate electrodynamics could accommodate it exactly because it also meant that a suitable velocity between the emitter and observer would allow the observer to see light slowed or even frozen. Once again, let us chase a beam of light. Assume the waveform of the light, propagating at c in the $+x$ direction, is $f(\omega t' - kx')$, with $c = \omega/k$. Then a Galilean transformation to an observer moving at u in the $+x$ direction is $t' = t$ and $x' = x + ut$ and the transformation between the two observers is:

$$f(\omega t' - kx') = f(\omega(1 - u/c)t - kx) \quad (23)$$

The light frequency is Doppler shifted to $\omega(1 - u/c)$ at the same time as the light is slowed to a speed of $(\omega/k)(1 - u/c) = c - u$. Einstein's original idea of field transformations could be of no further assistance in avoiding this slowing of light. Transformations of the field, such as (4) and (5), did not affect the speed of the wave, since they did not change the position of points of zero field intensity.

In retrospect it is so clear that Lorentz had precisely the theoretical machinery needed to solve this problem. A quick comparison of the Galilean transformation (23) with Lorentz's analogous (19) shows it. Had Einstein used Lorentz's local time in the transformation, then, he would have retained the correct result for the Doppler shift and the velocity of the light wave would have remained c . For $V = c$, Lorentz's waveform (19) is $f(\omega(1 + u/c)t - k(1 + u/c)x)$, so its speed of propagation is simply $c = \omega/k$. In a 1907 review article, Einstein (1907, p. 413) made just this point, writing

One needed only to realize that an auxiliary quantity that was introduced by H. A. Lorentz and that he called "local time" can simply be defined as "time."

Yet at the time it would be anything but an obvious connection. Indeed, even after Einstein showed the connection, Lorentz never felt it was the correct view. Lorentz's local time was an artificial mixture of time and space coordinates justified solely by its utility as an auxiliary quantity in the process of generating new solutions from old by means of the theorem of corresponding states; it simply was not the true time of the ether. The time that Einstein sought was the real time of his observers in relative motion. The recognition that local time might well just be this time would be easier for someone who approached Lorentz's work with a goal of realizing the principle of relativity. Even then the connection proved hard to make. Einstein read Lorentz's *Versuch* well before he had the crucial insight into the correct interpretation of local time.³³ He also had just the ability to adopt the sort of unorthodox views needed, for he had been quick to conclude that a magnetic field transforms to a mixture of magnetic and electric field merely under a change of state of motion of the observer.

³³ On 28 December 1901, Einstein wrote to Mileva Maric of his plan to read what both Lorentz and Drude had written on the electrodynamics of moving bodies. (*Papers*, Vol. 1, Doc. 131). Drude's (1900, Ch. 8) also contained an account of aberration, Fizeau's experiment and the theorem of corresponding states.

We may never know the precise path that led Einstein from his rejection of emission theories of light to Lorentz's local time and whether that path led through a rereading of parts of Lorentz's *Versuch* or an independent recreation of their content. However I want to point out one quite plausible possibility.

First note that the phenomenon of stellar aberration raises no special problem for an ether-based, wave theory of light. Take the simple case considered above of a star located in a direction perpendicular to the earth's motion through the ether. The light from the distant star arrives on earth as plane waves propagating perpendicular to the earth's motion. Earth-borne astronomers, however, must compensate for the earth's motion if they wish to center the star's image in their telescopes. Their telescopes must be tilted forward slightly by the very small angle of u/c radians to accommodate the earth's velocity u . The result is captured in an analogy given routinely in textbooks even today. The starlight is analogous to a rain shower with the raindrops falling vertically. If we run across the shower with a narrow, deep hat and try to catch the raindrops, we will need to tilt the hat slightly in the direction of our motion, so that the trailing side of the hat does not intercept the raindrops before they strike the hat's base. What holds for the raindrops entering the hat is also true for the portion of the starlight wave front that enters the telescope. The telescope must now be tilted if the wave front is to remain centered in the telescope barrel as it propagates along its length.

Stellar aberration does, however, raise serious problems for a wave theory of light if we ask *in addition* that the theory conforms to the principle of relativity. If the direction of starlight from a resting star is deflected by the earth's motion, then the principle of relativity demands a second effect, arrived at merely by redescribing this result in the earth's frame of reference. If the earth is at rest, then the apparent direction of starlight from a distant star must be affected by the star's motion u and by the usual angular deflection of u/c , for small u/c . It is not at all obvious how a wave theory of light can accommodate this result. To see the problem, begin with the star at rest. Its light will arrive at the earth as plane waves propagating perpendicular to the wave fronts. Let us say the waveform is $f(\omega t' - ky')$ so that it propagates along the y -axis to the earth. Now use a Galilean transformation to set the star and its light into uniform motion u in any direction perpendicular to the direction to the earth. Define the x -axis so that this transformation is $t' = t$, $x' = x + ut$, $y' = y$. The transformed waveform is just $f(\omega t - ky)$, which still propagates in exactly the same direction. The Galilean transformation is unable to turn the direction of the normal to the wave fronts and thus unable to replicate the change of direction of propagation demanded by a relativized stellar aberration.³⁴ In short, the principle of relativity demands that stellar aberration depends

³⁴ One might try to escape the problem by supposing that the direction of propagation is not always given by the normal to the wave front. We might, as Born (1962, p.141) proposed, identify the direction of propagation with the direction of energy propagation, supposing the latter to transform differently from the wave normal under Galilean transformation. Whatever may be the merits of such proposals, they are unavailable to someone trying to implement a principle of relativity. If the direction of propagation of a plane wave is normal to the wave fronts in one inertial frame—say the “ether” frame—then that must be true in all inertial frames. Or again: a simple Huygens construction tells us that plane waves propagate in a direction normal to their wavefronts. If that construction works in one inertial frame of reference, it

only on the relative velocity of the earth and the star; but the simple analysis suggests that the effect only arises when the earth is moving. (For an informal account of this problem, see Norton, manuscript.)

An emission theory of light would hold the promise of solving this problem, since it would seem to be well adapted to stellar aberration already. Here is the empirically validated rule for computing stellar, expressed in terms of relative velocities to conform to the principle of relativity. We add vectorially the velocity of the star with respect to the earth to the velocity of the starlight with respect to the star; the *direction* of the resulting vector is the apparent direction of the starlight on earth. An emission theory of light merely adds that this resulting vector gives not just the direction of the starlight, but its speed as well. It would seem that the observed phenomenon of stellar aberration almost directly expresses the central principle of an emission theory of light.

With his rejection of emission theories, if he had not already done so, Einstein would have to solve the problem of developing an account of stellar aberration that depends only on the relative velocity of the earth and the star. But then he might recall that Lorentz had recovered the essential result in his *Versuch*. Even in his ether-based theory, it turned out that stellar aberration did depend only on the relative velocity and not the absolute velocities of the earth and the star. It would have been a simple matter to consult the *Versuch* or Drude's *Lehrbuch der Optik* to recall how this was possible.

What stellar aberration and Fizeau's experiment show

Consider how Lorentz's transformation (20) may be read by someone seeking to realize a principle of relativity in electrodynamics. The waveform $f(\omega t - ky)$ represents light emitted by a star at rest. The waveform $f(\omega t - kux/c - ky)$ represents light emitted by the star when that star moves at u perpendicular to the direction of the emission of the light. For a relativist, that same motion could arise merely through an observer moving at u in the opposite direction. So the two waveforms could equally be just the same wave but viewed by observers in relative motion. Then one simply recovers the transformation between the two frames by reversing Lorentz's calculation $f(\omega t - kux/c - ky) = f(\omega(t - ux/c^2) - ky) = f(\omega t' - ky')$. Identifying the two arguments in the waveforms we have

$$\omega(t - ux/c^2) - ky = \omega t' - ky'$$

From this we read off the partial transformation

$$t' = t - ux/c^2 \quad y' = y.$$

This last step involves a small complication since we must divide one equation into two. The $\omega ux/c^2$ term could in principle be incorporated in the y transformation equation. That possibility proves to be unphysical.³⁵

must work in all. Stellar aberration must be accommodated by turning the wave fronts, if the principle of relativity is to be respected.

³⁵ If the term is located in the y transformation equation, that equation would be $y' = y + (u/c)x$. Thus a boost in the x direction would take an arbitrary point $(x,0)$ on the x -axis and shift it in the $+y$ direction,

What is important about viewing the calculation in reverse is that it requires no electrodynamical theory and no theorem of corresponding states to recover the above transformation. On the contrary, it is driven empirically. The observations associated with stellar aberration tell us that a waveform $f(\omega t - ky)$ will be deflected by an angle (u/c) if there is a relative motion $u \ll c$ of source and observer, perpendicular to the propagation. That fixes the deflected waveform as $f(A(t - (u/c)x - y))$, for some constant A . The value of A is then determined empirically by knowing that there is no transverse Doppler shift, at least to first order in u/c . That sets $A = \omega$ and we have recovered the full waveform. We now read the partial transformation between two inertial frames directly from empirically given waveforms.

The most important part, however is this: the deflection of the waveform is due entirely to the term kux/c in the waveform (20). In inferring back to the transformation, that deflection term simply becomes the local time term. In short the characteristic deflection of stellar aberration is direct empirical evidence for a local time term in the transformations; and, since we are reading backwards, it supplies that support independently of electrodynamical theory.

The analysis is essentially the same for Fizeau's experiment, with the algebra only slightly more complicated. The two waveforms of (21) can be thought of as the one wave viewed by two observers in relative motion. That is, if an observer moving with the water sees the waveform $f(\omega t' - kx')$, where $c/n = \omega/k$, then an observer moving at $-u$ in the x direction sees the waveform $f(\omega(1 + u/c)t - k(1 + u/cn)x)$. We recover the transformation between the two frames by identifying the arguments of the waveforms

$$\omega t' - kx' = \omega(1 + u/c)t - k(1 + u/cn)x = \omega(t - ux/c^2) - k(x - ut)$$

From this we read off the partial transformation

$$t' = t - ux/c^2 \quad x' = x - ut$$

As before, the last step involves a complication since we must divide one equation into two. The term $\omega ux/c^2$ could in principle be incorporated into the x transformation equation. That possibility proves to be unphysical.³⁶

The recovery of this partial transformation equation does not require electrodynamical theory or the theorem of corresponding states. It proceeds from two empirical observations: Fresnel's formula for the velocity of light in a moving medium, $c/n + u(1 - 1/n^2)$, and the Doppler shift. Note that the Doppler shift formula would be one in which Einstein could have considerable confidence. It is an observational result and the same formula is returned by transformations on waveforms using Lorentz's theorem of corresponding states (as shown in (19) above) or by a Galilean transformation of the waveform (as shown in (23)). The observer moving with the water sees a waveform $f(\omega t - kx)$, with $c/n = \omega/k$. That is simply a

which would violate the physically mandated symmetry of the transformation under reflection over the x axis.

³⁶ If the term is located in the x transformation equation, that equation becomes $x' = (1 + u/cn)x - ut$. It is unphysical because it retains the refractive index for water n in an equation that merely relates inertial frames; and it corresponds to a length expansion or contraction according to whether the velocity is directed in $+x$ or $-x$ direction, in contradiction with the isotropy of space.

wave moving at c/n . Recovering the waveform for the other observer is a little more complicated. From Fresnel's formula, we know that the waveform for light propagating in moving water is $f(\omega''t - k''x)$ where $\omega''/k'' = c/n + u(1 - 1/n^2)$. From the Doppler shift formula, we know that $\omega'' = \omega(1 - un/c)$. Combining the two, using a calculation essentially contained in (22), we find that $k'' = k(1 + u/cn)$. Hence the waveform is $f(\omega(1 + un/c)t - k(1 + u/cn)x)$. We proceed as above to recover the partial transformation.

My proposal is that Einstein recognized the essential import of the above reversed calculations: that one can read from the empirical phenomena of stellar aberration and Fizeau's experiment back to the transformation equations that relate the coordinates of inertial frames in relative motion, at least to first order quantities in u/c , and that these equations contain essentially a local time term. I do not say that Einstein necessarily went through the steps of the inverse calculation explicitly. A competent physicist running the calculation in the forward direction rapidly senses the close contact between the start and end. All I propose is that Einstein saw the closeness of the connection so that the empirical correctness of stellar aberration and the result of Fizeau's experiment made acceptance of a local time term in the transformation equation unavoidable and that this inevitability was largely independent of electrodynamical theory. This would bring Einstein to the recognition that, to first order in u/c , the equations relating inertial coordinates were just the space and time transformations of (18). All this could happen purely as a matter of manipulating equations formally and without recognition of the physical result of the relativity of simultaneity. But once the form of the equations was secured, their physical interpretation would be a pressing problem so that the relativity of simultaneity could not be far away.³⁷

Einstein also mentioned the importance of Airy's observations with a water filled telescope as an experiment having an important role in his thought. The analysis of Airy's experiment is implicit in the above; the result is accommodated by a combination of stellar aberration and the Fresnel drag and need not be elaborated here. See Miller (1981, p. 20).

Other writings

Einstein's other writings on stellar aberration and Fizeau's experiment offer little to assist us in deciding just how he used them in preparation for the 1905 paper beyond affirming their importance as experimental results. The repeated theme is their importance as experiments in deciding between an electrodynamics with a fully dragged ether or one with an immobile ether, specifically Lorentz's. Earlier discussions name only Fizeau's experiment in this context (Einstein 1909, p. 485; 1910, §2; 1911, pp. 3-4; 1915, pp. 703-704). Later discussion name both Fizeau's experiment and stellar aberration (Einstein, 1918; 1920, §1; 1920a, p.5; 1922, p. 18). Einstein (1954, p. 147) names stellar aberration and the Doppler effect, but not Fizeau's experiment. Fizeau's experiment is even sometimes described as an "*experimentum crucis*" (Einstein, 1910, p. 7, and below)

³⁷ The problem that the transformation holds only to first order quantities can be resolved by solving the straightforward mathematical problem of finding the second and higher order terms that must be added to make the transformation a group, which would give the familiar transformation of the 1905 paper. It is equivalent to the problem of constructing a continuous Lie group from its generators.

Einstein made clear that, at least in the case of Fizeau's experiment, in supporting Lorentz's theory, it also supported his relativity theory. He wrote (Einstein, 1917, §13; translation R. W. Lawson) of the success of Lorentz's theory with Fizeau's experiment:

This theory [of Lorentz] was of a purely electro-dynamical nature, and was obtained by the use of particular hypotheses as to the electromagnetic structure of matter. This circumstance, however, does not in the least diminish the conclusiveness of the experiment as a crucial test [German text: *experimentum crucis*] in favor of the theory of relativity, for the electro-dynamics of Maxwell-Lorentz, on which the original theory was based, in no way opposes the theory of relativity. Rather the latter has been developed from electro-dynamics as an astoundingly simple combination and generalization of the hypotheses, formerly independent of each other, on which electro-dynamics was built.

Similarly, in his manuscript on relativity, Einstein (1912-14, p.15) found the success of Lorentz's theory in accommodating Fizeau's experiment "a brilliant confirmation of Lorentz's theory and thereby one of the principal supports of the theory of relativity." In this manuscript and two sets of teaching notes (Einstein, 1914-1915; 1918-1919), Einstein gave a fully electro-dynamical derivation of the Fresnel drag within Lorentz's electro-dynamics. It is similar to an 1892 derivation of Lorentz (according to *Papers*, Vol. 7, p. 99, fn. 14) and essentially similar to Becker (1964, §72). Why Einstein would do this is puzzling at first. The derivation is quite unilluminating, demonstrating only that a rather cumbersome and opaque application of Maxwell's equations to the propagation of electromagnetic waves in moving media yields the Fresnel drag. Why give such a derivation when a much simpler, essentially kinematical derivation is already available within Lorentz's theory? My supposition is that its obscurity really is the point. Einstein wants to contrast the spirit of Lorentz's theory with relativity theory. The former accounts for the behavior of light in moving media by constructing a full account of the fields comprising the light; whereas the latter uses the principle of relativity to give a simple kinematical analysis to recover the same result.

We must also consider another possibility that Michel Janssen (private communication) has pointed out to me. When Einstein mentioned the importance of stellar aberration and Fizeau's experiment in his path to special relativity, he may merely have been recalling their historical importance in deciding between a resting and dragged ether, just as his other writings describe. Unfortunately Einstein's remarks are too brief for us to decide between the possibilities. I lean away from this new possibility, however, since I think it gives a poor account of the relevant remark in Einstein's tribute to Michelson (as quoted in Section 2 above). It leaves unexplained the remarkable coincidence that Einstein names *just* the two experiments whose analysis in Lorentz's 1895 *Versuch* depends entirely on local time and that he does it within three sentences of recalling his reading of Lorentz's *Versuch* prior to the 1905 paper. It also leaves unexplained the noteworthy juxtaposition of their mention with the magnet and conductor thought experiment. He wrote: "My direct path to the sp. th. rel. was mainly determined by the conviction that the electromotive force induced in a conductor moving in a magnetic field is nothing other than an electric field." And we know that he learned from this thought experiment that the principle of relativity requires the novel device of field transformations. He continues immediately "But

the result of Fizeau's experiment and the phenomenon of aberration also guided me." And I propose that the learned from these experiments that the principle of relativity requires a novel time transformation.

8. How Important was Clock Synchronization by Light Signals?

Waveforms or light signals?

In his 1905 "On the Electrodynamics of Moving bodies," Einstein considered the use of light signals to synchronize clocks as a means of establishing the relativity of simultaneity—perhaps the most famous conceptual analysis of modern science. The pervasiveness of this analysis in later writings has fostered a tacit assumption that Einstein's path to the relativity of simultaneity must have depended essentially on reflecting on light signals and how they might be used to synchronize clocks. So the literature in history of science looks to earlier analyses of light signals used to synchronize clocks and asks whether Einstein's possible contemplation of them led him to his essential insight of the relativity of simultaneity. A quite concrete candidate for such an earlier analysis is Poincaré's use of a light signaling protocol to interpret Lorentz's local time. See Darrigol (1996, p. 302).

What I would like to suggest is that it is entirely possible that thoughts of clocks and their synchronization by light signals played no essential role in Einstein's discovery of the relativity of simultaneity. A plausible scenario is that Einstein was compelled to the Lorentz transformation for space and time as a formal result, but needed some way to make its use of local time physically comprehensible. Thoughts of light signals and clock synchronization would then briefly play their role. It is also entirely possible that these thoughts entered only after Einstein had become convinced of the relativity of simultaneity; that is, they were introduced as an effective means of conveying the result to readers of his 1905 paper and convincing them of it. In both cases, thoughts of light signals and clock synchronization most likely played a role only at one brief moment, some five to six weeks prior to the completion of the paper, at the time that Einstein brought his struggle with him to a celebrated meeting with his friend Michele Besso. (Stachel, 1982, p. 185) We should not allow the excitement of this moment to obscure the fact that its place in Einstein's pathway is momentary in comparison to the years of arduous exploration that preceded.

With the near complete lack of direct evidence on the final steps Einstein took to the discovery, it is difficult to say anything with great confidence. However we can say this much: all of Einstein's significant remarks on how light entered into his deliberations prior to 1905 pertain to light as a waveform and not a signal (that is, a spatially localized point moving at c). Light, in his chasing a light beam thought experiment, for example, is a propagating waveform, for he immediately remarks that the resulting frozen light would be a frozen electromagnetic field, incompatible with experience and Maxwell's equations. The optical experiments of stellar aberration and Fizeau's experiment, if they follow the sort of analysis Lorentz pioneered, are also analyses of waveforms.

Waveforms in 1905

The Fizeau experiment can be reanalyzed in terms of the speed of propagation of light signals. The experimental result turns out simply to be an application of the relativistic rule of velocity composition. It is applied to the velocity of the light c/n in the water composed with v , the velocity of the water:

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

where the approximation holds up to quantities of first order in v/c . This analysis of Fizeau's result appeared in Einstein's developments of relativity theory starting in 1907. See Einstein (1907, p. 426) and Einstein (1917, §13). We are assured, however, that this was *not* the analysis Einstein used prior to his 1905 paper by a remark in the introduction of Einstein (1907, p. 413-14), where Einstein thanks Laue for alerting him to the possibility of the analysis both orally and through a paper, which is cited later when the analysis is given.

It might seem surprising that Einstein could devise and publish the relativistic rule of velocity composition in his 1905 paper (§5) without recognizing that the result of the Fizeau experiment is a vivid implementation of the rule. If however, we assume that Einstein's analysis of light propagation was largely conducted in term of waveforms and their Lorentz transformations, then it ceases to be surprising. As the last section showed, the result follows without any invocation of velocity composition. The situation with stellar aberration is similar. The result can be arrived at rapidly by means of the relativistic rule of velocity composition.³⁸ Yet Einstein (1905, §7) derives the result from the same transformation of the waveform that gives the Doppler shift without mention of velocity composition.

We know that Einstein thought of light in the context of stellar aberration and Fizeau's experiment in terms of waveforms even at the time of the writing of the 1905 paper and that he later singled them out as experimental results of greater importance in his thought than the Michelson-Morley experiment. We know that an analysis of the waveforms involved in these two results is sufficient to return the local time term responsible for relativity of simultaneity in the first order Lorentz transformations. Do we have any comparable positive evidence that shows that deliberations on light signals and clocks played any role in his discovery of the relativity of simultaneity beyond the question of how to present the result in its most convincing form?³⁹

³⁸ Following the notation of Einstein (1905, §5), if a light signal has velocity $(0, w_{\eta}=c, 0)$ in system k , then its velocity in system K is $(v, c(1-v^2/c^2)^{1/2}, 0)$, which is $(v, c, 0)$ to first order quantities, so that the signal is deflected by an angle of v/c radians.

³⁹ The closest to evidence that I know for a further role is in the transcript of an impromptu talk Einstein gave in Kyoto in 1922. Einstein recounted the importance of a visit to a friend (presumably, Besso) some 5 weeks before completion of the theory:

The very next day, I visited him again and immediately said to him: "Thanks to you, I have completely solved my problem."

9. Conclusion

Einstein recalled “the seven and more years that the development of the Special Theory of Relativity had been my entire life.”⁴⁰ The few clues he left can give us no more than glimpses of the intellectual struggles of these years, like momentary glances of a distant land through a train window. Yet they reveal a lot. Throughout, we see an Einstein stubbornly determined to realize the principle of relativity in electrodynamics. There were clues that he read well—the curious failure of all optical experiments to yield a demonstration of the earth’s motion and the apparent replication of this curiosity in some parts of electrodynamics. There were also strong signs that the quest was mistaken. Relative motion did fix the observables in the magnet and conductor thought experiment. But, as Föppl pointed out, that dependence solely on relative motion did not extend to all of electrodynamics and equally simple thought experiments did not manifest it. The thought experiment gave Einstein the device of field transformations and the expectation that this device would lead him to an implementation of the principle of relativity in electrodynamics. That expectation would surely look suspect if Einstein had explored the possibilities mapped out in Section 2. A Galilean covariant theory using his device of a field transformation law could only be made adequate to one part of Maxwell’s electrodynamics that also happened to accommodate the magnet and conductor thought experiment (“magnet and conductor partial theory”); a different field transformation was needed for the remaining part.

Einstein persisted. If Maxwell’s electrodynamics could not be made compatible with the principle of relativity by the device of field transformations, then the electrodynamics must be changed. The principle of relativity, if implemented in Galilean kinematics, dictated that the modified theory must embody an emission theory of light. We know that Einstein entertained such a theory, that it was akin to the theorizing of Ritz and that it probably used retarded potentials. I have suggested that we have a strong candidate for the theory: it is the one Pauli incorrectly attributed to Ritz, as described in Section 3. That theory could be grafted directly on the “magnet and conductor partial theory” without any alteration of the partial theory. It would give Einstein both the relativistic treatment of the magnet and

My solution actually concerned the concept of time. Namely time cannot be absolutely defined by itself, and there is an unbreakable connection between time and signal velocity.

Using this idea, I could now resolve the great difficulty that I had previously felt.

(Revised translation from Stachel, 1982, p. 185.) It is unclear to me whether the formulation of the relativity of simultaneity that mentions “signal velocity” pertains to the way Einstein actually first conceived it; or whether the result has already been redescribed in an awkwardly oversimplified form for a non-technical audience. (A more careful statement would speak of simultaneity of spatially separated events, not just “time,” and make clear that it is not just signal velocity, but signal velocity only if the signal happens to be light.)

⁴⁰ Einstein Archive 1-168. Shankland (1962, p. 56) also reported: “I asked Professor Einstein about the three famous 1905 papers and how they all appeared to come at once. He told me that the work on special relativity ‘had been his life for over seven years and that this was the main thing’.”

conductor thought experiment using a field transformation law and also an emission theory of light. As outlined in Section 4, Einstein leveled objections against all theories of this type. Some were technical complications. The most fundamental, however, was that these emission theories admitted no field theory. To accept some action at a distance formulation, as had Ritz, was a compromise Einstein was unwilling to make. The principle of relativity was to be realized in electrodynamics and it had to be done in the right way.

Einstein's stubbornness was reflected in the memorable thought experiment first conducted at age 16 in which he imagined chasing a beam of light. In Section 5, I have described how the thought experiment could provide no truly cogent reason for a 16 year old Einstein to doubt ether theories and, following remarks by Einstein, suggest that its initial import was more visceral than logical. Yet Einstein found the notion of chasing light sufficiently characteristic of his labors that this is the thought experiment given pride of place in his famous autobiographical reflections. In Section 6, I suggest how Einstein might have later turned the original thought experiment into logically compelling grounds for rejecting all emission theory of light. I also believe that this thought experiment is characteristic of how Einstein deliberated on the interaction between light and the motion of the observer for most of the preparatory work for special relativity. He looked to the effect of that motion on the *waveform* of the light. As I suggest in Section 8, there is little evidence of Einstein pondering at any length how the motion of the observer might affect light signals used to synchronize clocks; or that such analysis was more than a convenient way to present a result achieved by other means. Our present obsession with finding precursors for such analysis seems to be more a reflection of the powerful effect this analysis has had on us than any encouragement offered by Einstein's autobiographical remarks.

The fertility of Einstein's stubbornness surely owes a lot to his tempered respect for experiment. Later he could barely recall whether he knew of the Michelson-Morley experiment, instead calling to mind stellar aberration and Fizeau's experiment on the speed of light in moving water. I have suggested in Section 7 why these particular experiments may have been so memorable. They are the experimental results recovered with great ingenuity by Lorentz in his 1895 *Versuch* by means of the novel conception of local time. My proposal is that these experiments can be analysed in reverse, so that one arrives at the necessity of local time on the basis of these two experimental results independently of any detailed electro-dynamical theory. Local time, in Einstein's hands, transforms into the celebrated result of the relativity of simultaneity. But that transformation is only possible if one comes to Lorentz's formalism and asks how it could be used to realize a principle of relativity, concluding that all inertial observers have their own distinct times, with none preferred. Since Lorentz did not share Einstein's conviction that the principle of relativity must be realized unconditionally, he never found Einstein's reinterpretation compelling.

Einstein's determination was rewarded. The realizing of the principle of relativity in electrodynamics yielded a new theory of space and time that sped Einstein towards the pantheon of science. We should, however, resist the temptation of investing Einstein's determination with a mystical prescience. He had no extraordinary power to divine that this was the right path. All we can really attest to is a persistence that was both fertile and, at times, bordered on unmoving dogmatism. Before we invest

any more into it, we should recall the pattern of the research to come. Starting in 1907, Einstein developed a determination to realize an extension of the principle of relativity to acceleration through a relativistic theory of gravity. No one can doubt the fertility of these efforts over the years that follow; they gave us his general theory of relativity. No one can doubt the dedication of Einstein's pursuit in the face of daunting mathematical obstacles. (Norton, 1984) What we should doubt is his prescience. For we remain divided on the question of whether he achieved the goal single-mindedly pursued, a generalized principle of relativity. (Norton, 1993) With general relativity completed, Einstein refocused his unbending resolve on the idea that the quantum riddle was to be solved by a unified field theory that extended the spacetime methods of his general theory of relativity to electrodynamics. While Einstein's dedication in over three decades of work remains beyond doubt, a half century after his death, what must be doubted is both its fertility and success.

Appendices

The following identity of vector calculus will be used frequently in the calculations of the appendices. For any vector field \mathbf{F} and any *constant* vector field \mathbf{v} we have

$$\nabla \times (\mathbf{v} \times \mathbf{F}) \equiv -(\mathbf{v} \cdot \nabla) \mathbf{F} + \mathbf{v} (\nabla \cdot \mathbf{F}) \quad (\text{I})$$

It is most easily verified by simply computing the components of each expression directly.

Appendix A. The Magnet and Conductor Thought Experiment

Einstein's result—that the observable current depends only on the relative motion—can be derived in a fully Galilean covariant analysis using only two of the four Maxwell equations (M2, M4) and the Lorentz force law (L):

Case I. The magnet is at rest and charge e in the conductor moves at \mathbf{v} . By direct application of the Lorentz force law (L), we have that the current generating force on the charge is

$$\mathbf{f}/e = (1/c)(\mathbf{v} \times \mathbf{H}) \quad (\text{A1})$$

Case II. The charge e and the conductor are at rest and the magnet moves past at $-\mathbf{v}$. We compute the current generating force on the charge when the charge and magnet, judged from the magnet rest frame, have the same relative position and relative velocities as in Case I. The force will be due to an electric field induced by the time dependent magnetic field of the passing magnet. The primed coordinate system ($t', \mathbf{r}' = (x', y', z')$) and field \mathbf{H}' pertain to the magnet rest frame; unprimed quantities pertain to the conductor rest frame ($t, \mathbf{r} = (x, y, z)$). They are related by a Galilean transformation

$$\mathbf{H} = \mathbf{H}' \quad t = t' \quad \mathbf{r} = \mathbf{r}' - \mathbf{v}t' \quad (\text{A2})$$

so that $\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \frac{\partial \mathbf{r}}{\partial t'} \cdot \nabla = \frac{\partial}{\partial t} - \mathbf{v} \cdot \nabla$. In the rest frame of the magnet, the magnetic field is independent of

time, so that $\frac{\partial \mathbf{H}'}{\partial t'} = 0$; which entails that $\frac{\partial \mathbf{H}}{\partial t} = (\mathbf{v} \cdot \nabla) \mathbf{H}$ in the ether frame. Using the identity (I), we

recover $\frac{\partial \mathbf{H}}{\partial t} = -\nabla \times (\mathbf{v} \times \mathbf{H}) + \mathbf{v} (\nabla \cdot \mathbf{H}) = -\nabla \times (\mathbf{v} \times \mathbf{H})$, where the last equality follows after application of

Maxwell's equation $\nabla \cdot \mathbf{H} = 0$ (M2). We have from Maxwell's equation (M4) that $\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}$. Hence

$\nabla \times \mathbf{E} = (1/c) \nabla \times (\mathbf{v} \times \mathbf{H})$. If two vector fields agree in their curls, then by a standard theorem, they agree up to an additive term in the form of a gradient of an arbitrary scalar field φ . Hence

$$\mathbf{f}/e = \mathbf{E} = (1/c)(\mathbf{v} \times \mathbf{H}) - \nabla \varphi \quad (\text{A3})$$

The additive term $\nabla \varphi$ makes no contribution to the steady current in a closed conductor. It contributes a term $\oint \nabla \varphi \cdot d\mathbf{r} = 0$ to the emf; the term vanishes by an application of Stokes' theorem using the fact that $\nabla \times \nabla \varphi = 0$. Thus the two forces (A1) and (A3) on charges in the conductor will yield the same current in a closed conductor in the two cases, provided the field \mathbf{H} is the same when the magnet and conductor have the same relative positions and velocities. That sameness is assured by the transformation $\mathbf{H}' = \mathbf{H}$.

This last transformation $\mathbf{H}' = \mathbf{H}$ is the weak point of the calculation. At first it seems too obvious to be troublesome. It merely asserts that a moving magnet carries with it, in the co-moving frame, a clone of the field it carries when at rest in the ether. Moreover this assumption then leads directly to the result that the forces of (A1) and (A3) agree. However the transformation $\mathbf{H}' = \mathbf{H}$ is not something to be assumed. Maxwell's theory is sufficiently complete to specify the field of a moving magnet. It is something to be derived from Maxwell's equations, not posited independently. We now know using Lorentz's theorem of corresponding states that this transformation only holds to first order quantities and fails if there is an electric field somehow also associated with the magnet at rest in the ether. We could proceed on this path, but that would lead us into the depths of a Lorentz covariant analysis that would include the assumption that the force \mathbf{f} does not transform by a Galilean transformation (as tacitly supposed here) but by a Lorentz transformation. I will set all this aside. My concern is how the calculation would have first appeared to Einstein and at a time when he did not use the Lorentz transformation for forces. He tells us his result: the two currents are the same; that is, the two forces of (A1) and (A3) are the same. So we can immediately infer back that he must have assumed the transformation $\mathbf{H}' = \mathbf{H}$. More cautiously, to get agreement in (A1) and (A3), he need merely assume that \mathbf{H} and \mathbf{H}' agree up to an undetermined component parallel to \mathbf{v} , which would make no contribution to the force when the vector product of (A1)/(A3) is taken.

For completeness, I note the outcome of applying the remaining two of Maxwell's equations. The result is augmented comfortably by Maxwell's equation $\nabla \cdot \mathbf{E} = 4\pi\rho$ (M1), for charge density $\rho=0$. To apply it, we need to note that the operator $\nabla \times$ is an invariant under a Galilean transformation so that $\nabla \times \mathbf{H} = \nabla' \times \mathbf{H}'$. Since $\nabla' \times \mathbf{H}' = 0$ in the magnet rest frame, it follows that $\nabla \times \mathbf{H} = 0$ in the conductor frame. Applying (M1) to the \mathbf{E} field of (A3) yields

$$0 = \nabla \cdot \mathbf{E} = (1/c) \nabla \cdot (\mathbf{v} \times \mathbf{H}) - \nabla \cdot \nabla \varphi = (1/c) (\mathbf{H} \cdot (\nabla \times \mathbf{v}) - \mathbf{v} \cdot (\nabla \times \mathbf{H})) - \nabla \cdot \nabla \varphi = -\nabla \cdot \nabla \varphi;$$

so that Maxwell's equation (M1), requires that the field φ be harmonic, satisfying $\nabla \cdot \nabla \varphi = 0$. Applying the remaining Maxwell equation $\nabla \times \mathbf{H} = \frac{4\pi}{c} \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}$ (M3) is disastrous, however. Since we have both $\nabla \times \mathbf{H} = 0$ and $\mathbf{j} = 0$ (outside the conductor), it immediately follows that $\partial \mathbf{E} / \partial t = 0$ so the \mathbf{E} field is constant in time and no \mathbf{E} field can be brought into being by the passage of the magnet. If (M3) is invoked, the existence of the induced electric field (A3) is contradicted and the analysis fails.

In retrospect, it is not at all surprising that the analysis fails when all four of Maxwell's equations are invoked, for these equations are Lorentz covariant, not Galilean covariant. What is surprising is that so much of the analysis can be given in a Galilean covariant account, compatible with three of Maxwell's equations and the Lorentz force law. For comparison, we can see how the Lorentz covariant analysis proceeds by replacing the Galilean transformation (A2) by the *first order* Lorentz transformation, which, for the case of $\mathbf{E}' = 0$, is

$$\mathbf{H} = \mathbf{H}' \quad t = t' - \mathbf{v} \cdot \mathbf{r}' / c^2 \quad \mathbf{r} = \mathbf{r}' - \mathbf{v} t'$$

Under this transformation, the $\nabla \times$ operator is not invariant. Instead we have $\nabla' \times = \nabla \times - (1/c^2) \mathbf{v} \times (\partial / \partial t)$ with the additional term in $\partial / \partial t'$ arising directly from Lorentz' local time or Einstein's relativity of

simultaneity, depending on the view taken. The field of the magnet is irrotational in its rest frame: $\nabla' \times \mathbf{H}' = 0$. This transforms directly to $\nabla \times \mathbf{H} = (1/c^2) \mathbf{v} \times (\partial \mathbf{H} / \partial t)$ in the conductor rest frame. Using the formerly troublesome Maxwell equation (M3) to substitute for $\nabla \times \mathbf{H}$, we now recover $\frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c} \frac{\partial}{\partial t} (\mathbf{v} \times \mathbf{H})$. (It turns out that the calculation repeated with the exact Lorentz transformation yields this last equation as well.) Integrating with respect to t we have $\mathbf{E} = (1/c) (\mathbf{v} \times \mathbf{H}) + \mathbf{E}_{\text{constant}}$ where $\mathbf{E}_{\text{constant}}$ is an \mathbf{E} field constant in time only. We can readily set this time-constant field to zero by noting that it is, by presumption, zero in the vicinity of the conductor prior to the approach of the magnet; thus it must vanish there for all time. Hence the invocation of (M3) in conjunction with the Lorentz transformation gives us the \mathbf{E} field of (A3) as well as the means to set the additive field to zero.

Appendix B: Galilean Covariance Properties of Maxwell's Electrodynamics

Uniqueness of field transformation (5)

We can see that the field transformation $\mathbf{E} = \mathbf{E}' + (1/c)(\mathbf{v} \times \mathbf{H}')$, $\mathbf{H} = \mathbf{H}'$ (5) is the unique transformation preserving covariance of the Lorentz force law (L) as follows. First, the transformation must be linear if it is to respect the linearity of Maxwell's theory. To see this, represent the combined states of the field \mathbf{E} and \mathbf{H} by the six component vector $\mathbf{F} = (\mathbf{E}, \mathbf{H})$ and write the transformation sought as mapping \mathbf{F} to $\mathbf{T}(\mathbf{F})$. The linearity of Maxwell's theory entails that any linear sum $\mathbf{F} = a\mathbf{F}_1 + b\mathbf{F}_2$ of two fields \mathbf{F}_1 and \mathbf{F}_2 (for any reals a and b) is also a licit field and that this summation is an invariant fact; that is, it does not depend on the coordinate system employed for the description. This means that the transform of the summed field $\mathbf{T}(\mathbf{F}) = \mathbf{T}(a\mathbf{F}_1 + b\mathbf{F}_2)$ must be the same field as would be recovered if we transformed the fields first and then summed them; that is, $\mathbf{T}(\mathbf{F}) = a\mathbf{T}(\mathbf{F}_1) + b\mathbf{T}(\mathbf{F}_2)$. Combining we recover

$$\mathbf{T}(a\mathbf{F}_1 + b\mathbf{F}_2) = a\mathbf{T}(\mathbf{F}_1) + b\mathbf{T}(\mathbf{F}_2),$$

which just expresses the linearity of the transformation. Breaking \mathbf{F} into its two field parts, we can now write the linear transformation in its most general form as a transformation from a primed to an unprimed frame moving at \mathbf{u} :

$$\mathbf{E} = \mathbf{A}(\mathbf{u})\mathbf{E}' + \mathbf{B}(\mathbf{u})\mathbf{H}' \quad \mathbf{H} = \mathbf{C}(\mathbf{u})\mathbf{E}' + \mathbf{D}(\mathbf{u})\mathbf{H}' \quad (\text{B1})$$

$\mathbf{A}(\mathbf{u})$, $\mathbf{B}(\mathbf{u})$, etc. are linear operators that map vectors to vectors (i.e. tensor operators) and functions of \mathbf{u} alone. Since force \mathbf{f} is an invariant under the Galilean transformation, we must have $\mathbf{f}/e = \mathbf{f}'/e$. Therefore, if the Lorentz force law is covariant under transformation (B1), we must have $\mathbf{f}'/e = \mathbf{E}' + (1/c)(\mathbf{v} \times \mathbf{H}') = \mathbf{f}/e = \mathbf{E} + (1/c)((\mathbf{v}-\mathbf{u}) \times \mathbf{H})$. Substituting for \mathbf{E} and \mathbf{H} , we have

$$\mathbf{E}' + (1/c)(\mathbf{v} \times \mathbf{H}') = (\mathbf{A}(\mathbf{u})\mathbf{E}' + \mathbf{B}(\mathbf{u})\mathbf{H}') + (1/c)((\mathbf{v}-\mathbf{u}) \times (\mathbf{C}(\mathbf{u})\mathbf{E}' + \mathbf{D}(\mathbf{u})\mathbf{H}')) \quad (\text{B2})$$

For the case of $\mathbf{H}'=0$, (B2) reduces to $\mathbf{E}' = \mathbf{A}(\mathbf{u})\mathbf{E}' + (1/c)((\mathbf{v}-\mathbf{u}) \times (\mathbf{C}(\mathbf{u})\mathbf{E}'))$. Since \mathbf{v} is an arbitrary vector, this equality is only assured to hold if $\mathbf{C}(\mathbf{u})=0$, the zero operator, and $\mathbf{A}(\mathbf{u})=\mathbf{I}$, the identity. For the case of $\mathbf{E}'=0$, equality (B2) reduces to

$$(1/c)(\mathbf{v} \times \mathbf{H}') = \mathbf{B}(\mathbf{u})\mathbf{H}' + (1/c)(\mathbf{v}-\mathbf{u}) \times (\mathbf{D}(\mathbf{u})\mathbf{H}') \quad (\text{B3})$$

For the case of $\mathbf{v}=0$, this reduces to $\mathbf{B}(\mathbf{u})\mathbf{H}' = (1/c)\mathbf{u} \times (\mathbf{D}(\mathbf{u})\mathbf{H}')$. Substituting this last equality back into (B3) yields $\mathbf{v} \times \mathbf{H}' = \mathbf{v} \times (\mathbf{D}(\mathbf{u})\mathbf{H}')$. Hence $\mathbf{D}(\mathbf{u})=\mathbf{I}$. Therefore finally, $\mathbf{B}(\mathbf{u})\mathbf{H}' = (1/c)\mathbf{u} \times \mathbf{H}'$. Combining the expressions recovered for \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{D} , we have $\mathbf{E} = \mathbf{IE}' + (1/c)\mathbf{u} \times \mathbf{H}' = \mathbf{E}' + (1/c)\mathbf{u} \times \mathbf{H}'$ and $\mathbf{H} = \mathbf{0E}' + \mathbf{IH}' = \mathbf{H}'$, which is just (5).

Galilean covariance of Maxwell's equations

To demonstrate the Galilean covariance stated in Table 1, note that the Galilean transformation $t=t'$, $\mathbf{r}=\mathbf{r}'-\mathbf{u}t'$ entails the variable and operator transformations

$$\nabla' = \nabla \quad \nabla' \times = \nabla \times \quad \partial / \partial t' = \partial / \partial t - \mathbf{u} \cdot \nabla \quad \mathbf{v}' = \mathbf{v} + \mathbf{u} \quad \rho' = \rho \quad \mathbf{j}' = \mathbf{j} + \rho \mathbf{u}$$

Covariance of (M2) and (M4) under $\mathbf{E} = \mathbf{E}' + (1/c)\mathbf{u} \times \mathbf{H}' \quad \mathbf{H} = \mathbf{H}'$ (5)

For $\nabla \cdot \mathbf{H} = 0$ (M2), the covariance is automatic since $\nabla' \cdot \mathbf{H}' = \nabla \cdot \mathbf{H}$. For (M4), using the above substitutions, we have that $\nabla' \times \mathbf{E}' = -(1/c)\partial \mathbf{H}' / \partial t'$

becomes $\nabla \times (\mathbf{E} - (1/c)\mathbf{u} \times \mathbf{H}) = -(1/c)(\partial \mathbf{H} / \partial t - (\mathbf{u} \cdot \nabla)\mathbf{H})$,

which is $\nabla \times \mathbf{E} - (1/c)(\mathbf{u}(\nabla \cdot \mathbf{H}) - (\mathbf{u} \cdot \nabla)\mathbf{H}) = -(1/c)\partial \mathbf{H} / \partial t + (1/c)(\mathbf{u} \cdot \nabla)\mathbf{H}$

using identity (I). Invoking (M2) and canceling like terms, we recover

$\nabla \times \mathbf{E} = -(1/c)\partial \mathbf{H} / \partial t$ and the covariance is shown.

Covariance of (M1) and (M3) under $\mathbf{E} = \mathbf{E}' \quad \mathbf{H} = \mathbf{H}' - (1/c)(\mathbf{u} \times \mathbf{E}')$ (4)

For $\nabla \cdot \mathbf{E} = 4\pi\rho$ (M1), the covariance is automatic, since $\nabla' \cdot \mathbf{E}' = \nabla \cdot \mathbf{E}$ and $\rho' = \rho$. For (M3), using the above substitutions we have that $\nabla' \times \mathbf{H}' = (4\pi/c)\mathbf{j}' + (1/c)\partial \mathbf{E}' / \partial t'$

becomes $\nabla \times (\mathbf{H} + (1/c)\mathbf{u} \times \mathbf{E}) = (4\pi/c)(\mathbf{j} + \rho\mathbf{u}) + (1/c)(\partial \mathbf{E} / \partial t - (\mathbf{u} \cdot \nabla)\mathbf{E})$,

which is $\nabla \times \mathbf{H} + (1/c)(\mathbf{u}(\nabla \cdot \mathbf{E}) - (\mathbf{u} \cdot \nabla)\mathbf{E}) = (4\pi/c)\mathbf{j} + (4\pi/c)\rho\mathbf{u} + (1/c)\partial \mathbf{E} / \partial t - (1/c)(\mathbf{u} \cdot \nabla)\mathbf{E}$,

using identity (I). Invoking (M1) and canceling like terms, we recover

$\nabla \times \mathbf{H} = (4\pi/c)\mathbf{j} + (1/c)\partial \mathbf{E} / \partial t$ and the covariance is shown.

Covariance of scalar and vector potentials under $\varphi = \varphi' - (1/c)\mathbf{u} \cdot \mathbf{A}' \quad \mathbf{A} = \mathbf{A}'$ (14)

The potentials φ and \mathbf{A} are defined by (6), (6') and we need to show the covariance of these definitions.

For $\mathbf{H} = \nabla \times \mathbf{A}$, the covariance is automatic, since $\mathbf{H}' = \nabla' \times \mathbf{A}' = \nabla \times \mathbf{A} = \mathbf{H}$. For \mathbf{E} , we have

$$\mathbf{E}' = -\nabla' \varphi' - (1/c) \partial \mathbf{A}' / \partial t' = -\nabla \varphi - (1/c)\nabla(\mathbf{u} \cdot \mathbf{A}) - (1/c) \partial \mathbf{A} / \partial t + (1/c)(\mathbf{u} \cdot \nabla)\mathbf{A}$$

$$= -\nabla \varphi - (1/c) \partial \mathbf{A} / \partial t - (1/c)\mathbf{u} \times (\nabla \times \mathbf{A}) \text{ using } \mathbf{u} \times (\nabla \times \mathbf{A}) = \nabla(\mathbf{u} \cdot \mathbf{A}) - (\mathbf{u} \cdot \nabla)\mathbf{A}, \text{ which is a vector identity for}$$

constant \mathbf{u} . Hence $\mathbf{E}' = \mathbf{E} - (1/c)\mathbf{u} \times \mathbf{H}$, which is a form of the field transformation (5). Note that this

demonstration depends upon the field quantities \mathbf{E} and \mathbf{H} transforming according to (5), under which

(M2) and (M4) are covariant.

Appendix C: Föppl's Two Charges Thought Experiment

Föppl considers two charges at rest in the ether. When they are set into uniform motion together, he recalls, the forces between them change as a result of an induced magnetic field, so the cases of rest

and joint common motion through the ether are observationally distinguishable. Föppl's thought experiment is a special case of one in which we consider any distribution of charges at rest in the ether, acted upon by their own electrostatic fields. Of course, if the charges are to remain at rest in the ether, there must be other forces present, whose nature lies outside the present consideration. We imagine that charge distribution is set into uniform motion through the ether and we compute the forces between the charges to see if a change in the forces would allow a co-moving observer to detect the uniform motion.

Select a test charge e . When it is at rest in the ether along with the remaining charge distribution ρ , the force acting on it is just

$$\mathbf{f}/e = \mathbf{E} \quad (C1)$$

where \mathbf{E} is the field due to the charge distribution ρ . Now take the case of this same charge distribution moving at velocity $-\mathbf{v}$ in the ether. Using the primed coordinate system ($t', \mathbf{r}' = (x', y', z')$) for the charge distribution rest frame and the unprimed coordinate system for the ether frame, we have the transformations

$$\mathbf{E} = \mathbf{E}' \quad t = t' \quad \mathbf{r} = \mathbf{r}' - \mathbf{v}t' \quad \frac{\partial}{\partial t'} = \frac{\partial}{\partial t} - \mathbf{v} \cdot \nabla \quad (C2)$$

The charge distribution is static in its rest frame and the \mathbf{E}' field time (t') independent, so we have

$$0 = \frac{\partial \mathbf{E}'}{\partial t'} = \frac{\partial \mathbf{E}}{\partial t} - (\mathbf{v} \cdot \nabla) \mathbf{E}. \text{ Hence, using identity (I), we have } \frac{\partial \mathbf{E}}{\partial t} = (\mathbf{v} \cdot \nabla) \mathbf{E} = -\nabla \times (\mathbf{v} \times \mathbf{E}) + \mathbf{v}(\nabla \cdot \mathbf{E}). \text{ Using}$$

Maxwell's equation (M3) to substitute for $\partial \mathbf{E} / \partial t$ and using Maxwell's equation (M1), with $\mathbf{j} = -\rho \mathbf{v}$, to substitute for $\nabla \cdot \mathbf{E}$, we recover $c \nabla \times \mathbf{H} - 4\pi \mathbf{j} = -\nabla \times (\mathbf{v} \times \mathbf{E}) - 4\pi \mathbf{j}$, so that $\nabla \times \mathbf{H} = -(1/c) \nabla \times (\mathbf{v} \times \mathbf{E})$. When two vector fields agree in their curls, then, by a standard theorem, they agree up to an additive term in the form of a gradient of an arbitrary scalar field φ . Hence

$$\mathbf{H} = -(1/c)(\mathbf{v} \times \mathbf{E}) + \nabla \varphi \quad (C3)$$

Maxwell's equation (M1) and (M3) cannot fix the induced field \mathbf{H} any more closely, since they are unable to specify the irrotational part of a magnetic field. If we presume that the processes of Maxwell's equation (M3) are unable to generate irrotational magnetic fields, then it is natural (but not essential) to conceive of the component $\nabla \varphi$ of the field as independent of the motion of the charges and set it to zero as a boundary condition.

Invoking the Lorentz force law (L), it now follows that the force on the test charge e is

$$\mathbf{f}/e = \mathbf{E} + (1/c)(\mathbf{v} \times \mathbf{H}) = \mathbf{E} - (1/c)^2(\mathbf{v} \times (\mathbf{v} \times \mathbf{E})) \quad (C4)$$

This force is in general *unequal* to that of (C1), so the resulting observable accelerations would allow us to distinguish the two cases of the charges at rest or in uniform motion in the ether.⁴¹

Prior to the application of the Lorentz force law (L), the analysis conforms to the two charge partial theory of Table 1. The induced magnetic field (C3) can be computed indirectly from Maxwell's

⁴¹ The forces will be equal only in the special cases in which the velocity \mathbf{v} has been chosen to be parallel to \mathbf{E} so that $\mathbf{v} \times \mathbf{E} = 0$. Note that no stipulation for $\nabla \varphi$ can remedy the inequality by eradicating the induced \mathbf{H} field, except perhaps at a single point. The induced field $\mathbf{H} = -(1/c)(\mathbf{v} \times \mathbf{E})$ has non-vanishing curl, whereas the field $\mathbf{H} = \nabla \varphi$ is irrotational.

equations (M1) and (M3) as above; or it may be computed directly from the field transformation law $\mathbf{E} = \mathbf{E}'$, $\mathbf{H} = \mathbf{H}' - (1/c)(\mathbf{u} \times \mathbf{E}')$ (4). Since $\mathbf{H}'=0$ and $\mathbf{v}=\mathbf{u}$, we have $\mathbf{H} = -(1/c)(\mathbf{v} \times \mathbf{E})$ for the induced magnetic field. So, using this field transformation law, the disposition of fields (but not forces) in the two charge thought experiment can be given Galilean covariant treatment.

The weak point of this calculation is the transformation $\mathbf{E} = \mathbf{E}'$ of (C2). The situation is analogous to the assumption $\mathbf{H} = \mathbf{H}'$ in the computation of the magnet and conductor in Appendix A. Both seem entirely natural. Here we merely assume that a moving charge distribution carries with it a clone of the electrostatic field it carried when at rest in the ether. However Maxwell's theory is sufficiently complete for it to specify what field accompanies moving charges. It is a result to be deduced and not postulated independently. A fully relativistic analysis would eradicate the velocity dependence of the result. I will not pursue it here since, among other things, it would require abandoning the Galilean covariant transformation for the force \mathbf{f} in favor of a Lorentz transformation rule.

References

- Beck, Anna (trans.) (1983) [English Translation of] *The Collected Papers of Albert Einstein: Volume 1: The Early Years: 1879-1902*. P. Havas, consultant. Princeton: Princeton University Press.
- Beck, Anna (trans.) (1996) [English Translation of] *The Collected Papers of Albert Einstein: Volume 4: The Swiss Years: 1912-1914*. D. Howard, consultant. Princeton: Princeton University Press.
- Beck, Anna (trans.) (1995) [English Translation of] *The Collected Papers of Albert Einstein: Volume 5: The Swiss Years: Correspondence, 1902-1914*. D. Howard, consultant. Princeton: Princeton University Press.
- Becker, Richard (1964) *Electromagnetic Fields and Interactions*. 2 Vols. New York: Blaisdell; republished as one volume, New York: Dover, 1982.
- Born, Max (1962) *Einstein's Theory of Relativity*. New York: Dover.
- Darrigol, Olivier (1996) "The Electrodynamical Origins of Relativity Theory," *Historical Studies in the Physical Sciences*, 26, pp. 241-312.
- Drude, Paul (1900) *Lehrbuch der Optik*. Leipzig: S. Hirzel.
- Earman, John, Glymour, Clark and Rynasiewicz, Robert (1983) "On Writing the History of Special Relativity," *PSA 1982*, Volume 2. Peter D. Asquith and Thomas Nickles, eds., East Lansing: Philosophy of Science Association, pp. 403-416.
- Ehrenfest, Paul (1912) "Zur Frage der Entbehrlichkeit des Lichtäthers," *Physikalische Zeitschrift*, 13, pp. 317-19.
- Einstein, Albert (1905) "Zur Elektrodynamik bewegter Körper," *Annalen der Physik*, 17, pp. 891-921; *Papers*, Vol. 2, Doc. 23.
- Einstein, Albert (1907) "Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen," *Jahrbuch der Radioaktivität und Elektronik*, 4, pp. 411-62; *Papers*, Vol. 2, Doc. 47.
- Einstein, Albert (1909) "Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung," *Deutsche Physikalische Gesellschaft, Verhandlungen*, 7(1909), pp. 482-500; also in *Physikalische Zeitschrift*, 10, pp. 817-825; *Papers*, Vol. 2, Doc. 60.
- Einstein, Albert (1909a) "Zum gegenwärtigen Stand des Strahlungsproblems," *Physikalische Zeitschrift*, 10, pp. 185-93; *Papers*, Vol. 2, Doc. 56.
- Einstein, Albert (1910) "Le Principe de Relativité et ses Conséquences dans la Physique Moderne," Trans. E. Guillaume. *Archives des Sciences Physiques et Naturelles*, 29, pp. 5-28, 125-44; Doc. 2 in *Papers*, Vol. 3.
- Einstein, Albert (1911) "Die Relativitätstheorie," *Naturforschende Gesellschaft in Zürich, Vierteljahrsschrift*, 56, pp. 1-14; *Papers*, Vol. 3, Doc. 17.
- Einstein, Albert (1912-14) Manuscript on the Special Theory of Relativity. *Papers*, Vol. 4, Doc. 1.
- Einstein, Albert (1914-15) Lecture notes for course on relativity at the University of Berlin, Winter Semester, 1914-1915. *Papers*, Vol. 6, Doc. 7.

- Einstein, Albert (1915) "Die Relativitätstheorie," in P. Hinnenburg, ed., *Die Kultur der Gegenwart. Ihre Entwicklung und ihre Ziele*. As reproduced in *Papers*, Vol. 4, Doc. 21.
- Einstein, Albert (1917) *Über die spezielle and die allgemeine Relativitätstheorie (Gemeinverständlich)*. Braunschweig: Friedr. Vieweg & Sohn; *Papers*, Vol. 6, Doc. 42. 15th expanded edition, trans. R. W. Lawson, *Relativity: the Special and the General Theory*. London: Methuen, 1954.
- Einstein, Albert (1918) "Bemerkung zu E. Gehrckes Notiz 'Über den Äther'," *Deutsche Physikalische Gesellschaft. Verhandlungen* 20, p.261.
- Einstein, Albert (1918-1919) Lecture notes for course on special relativity at the University of Berlin and the University of Zurich, Winter Semester, 1918-1919. *Papers*, Vol. 7, Doc. 12.
- Einstein, Albert (1920) "Fundamental Ideas and Methods of the theory of Relativity, Presented in Their Development," *Papers*, Vol. 7, Doc. 31.
- Einstein, Albert (1920a) *Äther und Relativität*. Berlin: Julius Springer. *Papers*, Vol. 7, Doc. 38.
- Einstein, Albert (1922) *Vier Vorlesungen über Relativitätstheorie*. Braunschweig: Friedr. Vieweg & Sohn (*Papers*, Vol. 7, Doc. 71); translated as *The Meaning of Relativity*. Princeton: Princeton Univ. Press, 1922.
- Einstein, Albert (1949) *Autobiographical Notes*. P. A. Schilpp, trans. and ed., La Salle and Chicago: Open court, 1979.
- Einstein, Albert (1954) "Relativity and the Problem of Space" Appendix V added to 5th edition of Einstein (1917).
- Einstein, Albert (1956) "Autobiographische Skizze," pp. 9-17 in Seelig, Carl (ed.) *Helle Zeit—Dunkle Zeit*. Zurich: Europa Verlag.
- Fölsing, Albrecht (1997) *Albert Einstein: A Biography*. Trans E. Osers. New York: Viking.
- Föppl, August (1894) *Einführung in die Maxwell'sche Theorie der Elektrizität*. Leipzig: B. G. Tuebner.
- Forman, Paul (1975) "Ritz, Walter," pp. 475-81 in C. C. Gillispie (ed.), *Dictionary of Scientific Biography*, Vol. XI, New York: Scribners, 1975.
- Hoffmann, Banesh (1982) "Some Einstein Anomalies," pp. 91-105 in G. Holton and Y. Elkana (eds.) *Albert Einstein" Historical and Cultural Perspectives*. Princeton: Princeton University Press.
- Holton, Gerald (1973) "Influences of Einstein's Early Work," Ch.7 in *Thematic Origins of Scientific Thought: Kepler to Einstein*. Cambridge, MA: Harvard University Press.
- Hovgaard, William (1931) "Ritz's Electrodynamical Theory," *Journal of Mathematics and Physics*, 11(1931-32), pp.218-254.
- Howard, Don and Stachel, John(eds.) (2000) *Einstein: the Formative Years, 1879-1909: Einstein Studies, Volume 8*. Boston: Birkhäuser.
- Jammer, Max and Stachel, John (1980) "If Maxwell had Worked between Ampère and Faraday: An Historical Fable with a Pedagogical Moral," *American Journal of Physics*, 48, pp. 5-7.
- Janssen, Michel (1995) *A Comparison Between Lorentz's Ether Theory and Special Relativity in the Light of the Experiments of Trouton and Noble*. PhD Dissertation. Department of History and Philosophy of Science, University of Pittsburgh.

- Janssen, Michel *et al.* (2002) *The Collected Papers of Albert Einstein: Volume 7: The Berlin Years: Writings, 1918-1921*. Princeton: Princeton University Press. ("Papers, Vol. 7.")
- Janssen, Michel and Stachel, John (forthcoming) "The Optics and Electrodynamics of Moving Bodies," in J. Stachel, *Going Critical*.
- Klein, Martin J. *et al.* (1993) *The Collected Papers of Albert Einstein: Volume 3: The Swiss Years: Writings, 1909-1911*. Princeton: Princeton University Press. ("Papers, Vol. 3.")
- Klein, Martin J. *et al.* (1995) *The Collected Papers of Albert Einstein: Volume 4: The Swiss Years: Writings, 1912-1914*. Princeton: Princeton University Press. ("Papers, Vol. 4.")
- Kox, A. J. *et al.* (1996) *The Collected Papers of Albert Einstein: Volume 6: The Berlin Years: Writings, 1914-1917*. Princeton: Princeton University Press. ("Papers, Vol. 6.")
- Lorentz, Hendrik A. (1895) *Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpern*. Leiden: E. J. Brill.
- Lorentz, Hendrik A. (1904) "Electromagnetic Phenomena in a System Moving with any Velocity less than that of Light," *Proceedings of the Academy of Sciences, Amsterdam*, 6; reprinted pp. 11-34 H. A. Lorentz *et al.*, *The Principle of Relativity*. Trans W. Perrett and G. B. Jeffrey. Methuen, 1923; Dover, 1952.
- Miller, Arthur (1981) *Albert Einstein's Special Theory of Relativity*. Reading, MA: Addison-Wesley.
- Norton, John D. (1984) "How Einstein Found His Field Equations: 1912-1915," *Historical Studies in the Physical Sciences*, 14 (1984), pp. 253-315. Reprinted in D. Howard and J. Stachel (eds.), *Einstein and the History of General Relativity: Einstein Studies Vol. I*, Boston: Birkhauser, pp101-159.
- Norton, John D. (1993) "General Covariance and the Foundations of General Relativity: Eight Decades of Dispute," *Reports on Progress in Physics*, 56 , pp.791-858.
- Norton, John D. (manuscript), "Einstein's Special Theory of Relativity and the Problems in the Electrodynamics of Moving Bodies that Led him to it," prepared for *Cambridge Companion to Einstein*, M. Janssen and C. Lehner, eds., Cambridge University Press.
- O'Rahilly, Alfred (1938) *Electromagnetics*. Longmans, Green & Co.; republished as *Electromagnetic Theory: A Critical Examination of Fundamentals*. 2 Vols., New York: Dover, 1965.
- Pauli, Wolfgang (1921) "Relativitätstheorie," in *Encyklopädie der mathematischen Wissenschaften, mit Einschluss an ihrer Anwendung*. Vol.5, *Physik*, Part 2. Arnold Sommerfeld, ed., Leipzig: B.G. Teubner, 1904-1922, pp.539-775. [Issued November 15, 1921] English translation, *Theory of Relativity*. With supplementary notes by the author. G. Field, trans. London: Pergamon, 1958; reprint New York: Dover.
- Ritz, Walter (1908) "Recherches Critique sur l'Électrodynamique Générale," *Annales des Chimie et de Physique*, 13, pp. 145-275; reprinted in Ritz (1911), pp. 317-426.
- Ritz, Walter (1908a) "Recherches Critiques sur les Théories Électrodynamiques de Cl. Maxwell and de H. -A. Lorentz," *Archives des Sciences physique et naturelles*, 26, pp. 209-36; reprinted Ritz (1911), pp. 427-92.
- Ritz, Walter (1909) "Über die Grundlagen der Elektrodynamik und die Theorie der Schwarzen Strahlung," *Physikalische Zeitschrift*, 9, pp. 903-907; reprinted as Ritz (1911), pp. 493-502.

- Ritz, Walter (1909a) "Zum gegenwärtigen Stand des Strahlungsproblems: Erwiderung auf den Aufsatz des Herrn A. Einstein," *Physikalische Zeitschrift*, 10, pp. 224-25.
- Ritz, Walter (1911) *Gesammelte Werke. Oeuvres*. Paris: Gauthier-Villars.
- Ritz, Walter and Einstein, Albert (1909), "Zum gegenwärtigen Stand des Strahlungsproblems," *Physikalische Zeitschrift*, 10, pp. 323-34; reprinted as *Papers*, Vol. 2, Doc. 57 and Ritz (1911), pp. 507-508.
- Rynasiewicz, Robert (1988) "Lorentz's Local Time and the Theorem of Corresponding States," PSA 1988, Vol. 1, A. Fine and J. Leplin, eds., East Lansing: Philosophy of Science Association, pp. 67-74.
- Rynasiewicz, Robert (2000) "The Construction of the Special Theory: Some Queries and Considerations," in Howard and Stachel (2000), pp. 159-201.
- Shankland, R. S. (1963/73) "Conversations with Einstein," *American Journal of Physics*, 31(1963), pp. 47-57; 41(1973), pp. 895-901.
- Sitter, W. de (1913) "Ein astronomischer Beweis für die Konstanz der Lichtgeschwindigkeit," *Physikalische Zeitschrift*, 14, p. 429.
- Stachel, John (1982) "Einstein and Michelson: the Context of Discovery and Context of Justification," *Astronomische Nachrichten*, 303(1982), pp. 47-53; reprinted Stachel (2002), pp. 177-90.
- Stachel, John (1984) "The Generally Covariant Form of Maxwell's Equations," pp. 23-37 in M. Berger, ed., *J. C. Maxwell: The Sesquicentennial Symposium*. North-Holland.
- Stachel, John (1987) "Einstein and Ether Drift Experiments," *Physics Today*, 40(1987), pp. 45-47; reprinted Stachel (2002), pp. 171-76.
- Stachel, John (1989) "Quale canzone cantarono le sirene: come scopro Einstein la teoria speciale della relatività?" *L'Opera di Einstein*. Umberto Curi (ed.). Ferrara, Gabriele Corbino & Co. (1989), pp. 21-37; translated and reprinted as "'What Song the Sirens Sang:' How Did Einstein discover Special Relativity?" in Stachel (2002), pp. 157-69.
- Stachel, John (2002) *Einstein from 'B' to 'Z': Einstein Studies, Volume 9*. Boston: Birkhäuser.
- Stachel, John et al. (eds.) (1987) *The Collected Papers of Albert Einstein: Volume 1: The Early Years: 1879-1902*. Princeton: Princeton University Press. ("Papers, Vol. 1")
- Stachel, John et al. (eds.) (1989) *The Collected Papers of Albert Einstein: Volume 2: The Swiss Years: Writing, 1900-1902*. Princeton: Princeton University Press. ("Papers, Vol. 2.")
- Stachel, John et al. (1989a) "Einstein on the Theory of Relativity," Headnote in *Papers*, Vol. 2., pp. 253-74.
- Tolman, Richard C. (1912) "Some Emission Theories of Light," *Physical Review*, XXXV, pp. 136-43.
- Wertheimer, Max (1959) *Productive Thinking*. New York: Harper & Bros.