Relativity and the A Priori

Reconsidering Friedman's *Dynamics of Reason* in Light of Darrigol's Comprehensibility Principles

Claus Festersen

This is a slightly revised version of my doctoral dissertation, prepared for my oral defence at Roskilde University on October 15, 2025.

Published 2025 by Festersen Academic Press

ISBN 978-87-976379-0-6

In gratitude to my parents, Rita and Hans-Heinrich.

❦

In memory of my academic mentor, Stig Andur Pedersen.

Abstract

During the last three decades a resurgence of interest in the a priori has emerged, with scholars like Michael Friedman arguing for constitutive a priori principles within scientific frameworks, particularly in Newtonian and relativistic physics. This dissertation engages with Friedman's concept of the relativized a priori, offering a critical analysis in relation to the transition from Newtonian physics to special relativity.

The first part explores Michael Friedman's concept of relativized a priori constitutive principles. Chapter 1 serves as an introduction and provides an overview of the content in each chapter. Chapter 2 focuses on Friedman's original formulation in *Dynamics of Reason* (2001), while chapter 3 addresses his article "Synthetic History Reconsidered" (2010).

The second part of the dissertation analyzes the historical transition from Lorentz's ether-based theory of electrodynamics to Ignatowski's derivation of the Lorentz transformations in 1910, with the aim of critiquing Friedman in the third and final part. Chapter 4 provides an overview of the critical conclusions that can be drawn based on the subsequent historical analysis. Chapter 5 analyzes Lorentz's two-tier strategy in his 1895 theory of corresponding states, while chapter 6 explores Henri Poincaré's belief that physical theories aim to fulfill general principles. Chapter 7 analyzes Poincaré's critique of Lorentz's theory, presented at a 1900 conference. Poincaré demonstrated that Lorentz's theory violated the principle of reaction and devised a thought experiment showing the possibility of perpetual motion, which he addressed by extending the principles of reaction and relative motion to the ether. However, Poincaré remained unsatisfied, as ether drift experiments suggested that the principle of relative motion applied only to matter. Chapter 8 explores Lorentz's 1904 electrodynamic theory, which sought to address Poincaré's objections by developing a refined version of the theory of corresponding states, ensuring the invariance of optical phenomena at all orders. However, Lorentz continued to believe that detecting motion relative to the ether might still be possible for other electrodynamic phenomena. Chapter 9 examines Poincaré's response to Lorentz's 1904 theory. Despite Abraham's 1905 demonstration that Lorentz's electron model was logically inconsistent, Poincaré did not discard Lorentz's new theory, as it allowed him to construct electrodynamics consistent with the principle of relativity. This was demonstrated mathemati-

cally by the invariance of Maxwell's equations under Lorentz transformations. Although this invariance ruled out distinguishing ether-based reference frames from other inertial frames empirically, Poincaré still viewed the ether as the medium for electric and magnetic fields. Chapter 10 covers Einstein's 1905 relativity paper, which derived the Lorentz transformations using only the principles of relativity and the constancy of the speed of light, without reference to the ether. Finally, chapter 11 introduces Ignatowski's 1910 derivation of the Lorentz transformations, which did not rely on electrodynamics but on general space-time symmetries. This analysis of the historical shift from Lorentz's 1895 electrodynamics to Ignatowski's derivation of the Lorentz transformations highlights two key inadequacies in Friedman's interpretation of the transition from Newtonian physics to special relativity. First, both Lorentz and Poincaré abandoned Newton's second and third laws of motion to address the null result of ether drift experiments, contrary to Friedman's claim that they upheld Newtonian space-time through these laws. Second, Ignatowski's approach to defining inertial frames using homogeneity and isotropy principles provided an alternative method applicable to both Newtonian and relativistic physics, challenging Friedman's view of the uniqueness of constitutive principles.

In light of the aforementioned shortcomings, the final part of the dissertation aims to provide a clearer understanding of the rationality behind the transition from Newtonian physics to special relativity, drawing on Olivier Darrigol's concepts of comprehensibility principles and interpretive schemes. Chapter 12 begins by introducing these comprehensibility principles, demonstrating how they clarify important aspects of the historical analysis covered in part 2. Rather than defining reference frames empirically, comprehensibility principles guide theory construction by supporting rational inference. Chapter 12 also introduces Darrigol's interpretive schemes, which enable the empirical application of a physical theory. Chapter 13 explores the connection between space-time and mechanics. Landau and Lifshitz conducted the most significant investigation of this connection in 1969 and 1972. They first applied the principle of least action, combined with Ignatowski's definition of inertial frames, to demonstrate the conservation of both linear and angular momentum. Their derivations demonstrate 1) how different principles can define the same inertial frames and 2) how to develop an interpretive scheme for constructing and synchronizing ideal clocks across Newtonian physics, special relativity and ether-based electrodynamics. The latter is the topic of chapter 14. To finally challenge Friedman's concept of constitutive a priori principles in chapter 15, we need only point to the existence and realizability of alternative schemes. As a result, Friedman's claim about the uniqueness of constitutive principles is both logically and empirically incorrect. Instead of supporting Friedman's understanding of constitutive a priori principles, the dissertation therefore advocates for Darrigol's concept of the relativized a priori as presuppositions for what are called comprehensibility principles. For example, Einstein demonstrated that the principle of light and the assumptions of homogeneity and isotropy necessitate a new spatio-temporal framework. However, these principles alone do not empirically define the new framework; that requires interpretive schemes and empirical know-how. Therefore, it is not problematic that different conditions can lead to the same spatio-temporal framework or that various interpretive schemes can empirically realize this structure. Moreover, by highlighting the dialectics of necessity and openness in the historical development of mathematical physics, the interplay between comprehensibility conditions and interpretive schemes promises a more nuanced philosophical understanding of this enterprise.

Acknowledgments

This dissertation marks the culmination of a 20-year-long academic journey. The journey began when, in 2004, Stig Andur Pedersen encouraged me to apply for a PhD scholarship with a project on the role of physics in Ernst Cassirer's *Philosophy of Symbolic Forms* at the Department of Philosophy and Science Studies at Roskilde University. By the end of that year, I received the news that my scholarship application had been granted, with Andur as my supervisor, and I embarked on the project.

However, it quickly became apparent that I struggled to see the potential in Cassirer's philosophical system for providing a deeper understanding of the development of physics. During my studies, I came across Michael Friedman's work *Dynamics of Reason*, whose analysis of the dialectical development of physics through the concept of constitutive a priori principles resonated with me far more profoundly. Initially, Andur was quite skeptical, but I was eventually allowed to change the course of my research. Friedman had written the second part of his work in response to critiques from Felix Mühlhölzer and his students during a stay at the Georg August University of Göttingen. In the autumn semester of 2006, I therefore stayed in Göttingen myself to discuss Friedman's philosophy with Mühlhölzer and a new group of students. Before my departure, it also became clear that I would have the opportunity to discuss *Dynamics of Reason* with Friedman himself, as Andur and Frederik Stjernfelt had succeeded in arranging a conference at the Carlsberg Academy in Copenhagen later that year, focused on Friedman's philosophical work, with Friedman in attendance.

I would like to take this opportunity to thank Friedman for his positive feedback on my presentation. Likewise, I owe my gratitude to Mühlhölzer and his students for the insightful discussions during my stay in Göttingen. I would also like to thank Steen Brock for his active support in seeking PhD scholarships both domestically and internationally during the years 2001-2004, and I am grateful to Jesper Lützen for giving me the opportunity to present my project at the Danish Society for the History of Science in October 2012. In January 2011, I became a high school teacher, and since then I have only been able to work on the dissertation during my summer holidays. I am grateful to Søren Riis, Director of Studies at the Section for Philosophy and Science Studies at Roskilde University, for keeping in touch with me over the past two years and for his encouragement to speed up the completion of the dissertation. The deepest gratitude, however, I owe to Andur, who passed away in 2018. May his memory be honored.

In addition to these academic acknowledgments, I also owe thanks to friends and family. A special thank you goes to Søren Christian Erichsen and Silvia Skov, Per Gammeltoft Jepsen and Annemarie Borregaard, as well as Erik Woodward Knudsen. I would also like to express my gratitude to my brother Carsten Festersen, my sister-in-law Rikke Nim Festersen, and my niece and nephews, Johanna, Kristoffer, and Tobias. Last but certainly not least, I owe an immense debt of gratitude to my parents, Rita and Hans-Heinrich Festersen, for their unwavering support throughout my life.

Contents

CHAPTER 1

Introduction

> "It would be hard to find a more crucial epistemological problem than that of the character of a priori knowledge."
>
> Alberto Coffa[1]

1.1 RELATIVITY

In 1905, Albert Einstein, a twenty-six-year-old patent officer in Bern, published a revolutionary theory of space and time in a paper entitled "Zur Elektrodynamik bewegter Körper." This theory became known as the special theory of relativity and consisted of two parts. In the first part, Einstein argued that space and time are relative rather than absolute terms of measurement. In the second part, he demonstrated how electrodynamics could be reconciled with his new space-time theory.

According to the renowned historian Paul Johnson, Einstein's genius and the sophistication of his reasoning were compared to a form of art and attracted growing interest worldwide. Not even the outbreak of the First World War deterred scientists from pursuing Einstein's search for an all-encompassing general theory of relativity that would also cover gravitational fields. In 1915, news reached the scientific community that he had been successful. Four years later, Sir Arthur Eddington led an expedition to the West African island of Principe, during which he photographed a solar eclipse that provided empirical confirmation of the general theory of relativity.[2]

Johnson explains Einstein's comprehensive scientific and social influence as follows:

> Einstein's theory and Eddington's much-publicized expedition to test it aroused enormous interest throughout the world in 1919. No exercise in scientific verification, before or since, has ever attracted so

[1] Coffa 1995, p. 1.
[2] Johnson 2013, pp. 1–4.

> many headlines or become a topic of universal conversation. [...] The scientific genius impinges on humanity, for good or ill, far more than any statesman or warlord. Galileo's empiricism created the ferment of natural philosophy in the seventeenth century, which adumbrated the scientific and industrial revolutions. Newtonian physics formed the framework of the eighteenth-century Enlightenment [...] So, too, the public response to relativity was one of the principal formative influences on the course of twentieth-century history.[3]

Philosophers in particular were quick to embrace Einstein's theories of relativity.

1.2 LOGICAL EMPIRICISM

The most important philosophical insight that the school of logical empiricism drew from Einstein was, for example, that the Kantian notion of synthetic a priori knowledge (i.e. substantial knowledge about the world based on pure reason) could no longer be upheld. However, if there are indeed no examples of synthetic a priori knowledge, then the logical empiricists also had to abandon Kant's idea of a characteristic transcendental task for philosophy that focuses on identifying and investigating the synthetic a priori elements within scientific knowledge. Although the logical empiricists rejected all non-empirical forms of knowledge, they also hesitated to call logical and mathematical knowledge empirical. Instead of postulating a special source of knowledge, they made a compromise by denying that logical and mathematical propositions have factual content. These kinds of propositions were supposedly known a priori solely in the sense that they were not acquired by empirical means. This should not be a cause for concern, because they are merely analytic propositions that are true by definition. To summarize, the logical empiricists regarded philosophy as a branch of logic and saw the task of philosophy as the systematic investigation of the various scientific domains by means of logical analysis.[4]

1.3 QUINE'S NATURALISM

Another philosophical response to Einstein was to question the possibility of a priori knowledge altogether. In his seminal works "Two Dogmas of Empiricism" (1951) and "Epistemology Naturalized" (1969), Willard Van Orman

[3] Johnson 2013, pp. 3–5.
[4] Shaffer and Veber 2011.

Quine took this step, which laid the foundation for philosophical naturalism, one of the most influential schools of thought of our time. According to Quine, our knowledge system should be conceptualized as a web of interconnected beliefs in which sensory input only comes into contact with the periphery. When we are confronted with an experience that contradicts our established knowledge system, we have the opportunity to decide where we need to make adjustments. These can be made near the periphery of the system, but they can also affect the most fundamental aspects of science, including the truths of logic and mathematics that form the core of the system. Since the latter are firmly entrenched, we are more reluctant to revise them. Quine's crucial point, however, was that all our beliefs are subject to potential revision when we are confronted with new experience:

> The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atom physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges. Or, to change the figure, total science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery causes readjustments in the interior of the field. [...] But the total field is so underdetermined by its boundary conditions and experience that there is much latitude of choice as to what statements to re-evaluate in light of any single contrary experience. [...] If this is right, [...] it becomes folly to seek a boundary between synthetic statements, which hold contingently on experience, and analytic statements, which hold come what may. Any statement can be held true, come what may, if we make drastic enough adjustments elsewhere in the system. [...] Conversely, by the same token, no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics; and what difference is there in principle between such a shift and the shift whereby Kepler superseded Ptolemy, Einstein Newton, or Darwin Aristotle?[5]

In Quine's holistic picture of knowledge, empirical evidence extends across all elements of our entire knowledge system. Therefore, all branches of knowledge, including logic, mathematics and philosophy, are equally em-

[5] Quine 1951, pp. 39–40.

pirical. From this point of view, the logical empiricists misunderstood the robustness of logic and mathematics as absolute immunity, which led them to mistakenly draw a sharp distinction between synthetic a posteriori and analytic a priori statements. As the end of the quote makes clear, Einstein's seminal contributions to theoretical physics played a crucial role in shaping this naturalistic perspective.

1.4 THE RETURN OF THE A PRIORI

According to David Stump, "[o]ne may think that Quine's critique of the analytic/synthetic distinction would have put the final nail in the coffin of a priori knowledge, and that a holism in which all scientific statements are justified empirically has replaced the notion of any special status for what was formerly considered to be a priori."[6] Along the same vein, Michael Veber and Michael Shaffer remark that "[i]n the last fifty years, radical empiricism and philosophical naturalism have gained wide popularity among epistemologists, almost, some would say, to the point of becoming a philosophical orthodoxy."[7] But they concede that in the last decade there has been a renaissance of interest in the a priori among contemporary epistemologists who are "dissatisfied with purely empirical attempts to account for logic, mathematics, and some other examples of apparently a priori knowledge."[8] In particular, what they call the constitutive view of a priori knowledge "is rooted in the Kantian idea that in order for empirical knowledge to be possible, some principles must be accepted a priori."[9] However, "[w]hile Kant was primarily concerned with establishing this sort of view at the level of individual subjects, contemporary constitutivists tend to argue for this view on the basis of scientific practice."[10] The most prominent advocate of this view, we may add, is Michael Friedman, a philosopher and historian of science at Stanford University.

1.5 FRIEDMAN AND THE A PRIORI

In his book *Dynamics of Reason* (2001), Friedman identifies a central flaw in Quine's holistic conception of scientific knowledge, namely its depiction

[6] Stump 2021, p. 203.
[7] Shaffer and Veber 2011, p. 3.
[8] Ibid., p. 3.
[9] Ibid., p. 7.
[10] Ibid., p. 7.

of the interconnectedness of different components within our knowledge system (such as mathematics, physics and geography) as symmetrical in relation to neighboring sensory experiences. As a starting point for his critique of this anti-apriorist conception of knowledge, Friedman strongly disagrees with the idea that this supposedly scientifically oriented form of anti-apriorism derives any plausibility from the historical development of mathematical natural science from Ptolemy via Kepler and Newton to Einstein. In contrast, Friedman argues that revolutionary changes in mathematical physics undermine and refute rather than support Quine's figure. The frameworks of Newtonian physics, special relativity and general relativity each comprise a unique set of a priori constitutive principles that empirically define their respective spatio-temporal structures. Newton's three laws of motion, for example, are the only set of principles that empirically define the inertial frames of Newtonian mechanics, within which all proper empirical laws then apply. In other words, without the a priori presupposition of the laws of motion, proper empirical laws, such as the universal law of gravitation, would have no empirical meaning. To distinguish this notion of constitutive a priori principles more clearly from the logical empiricist picture of an arbitrary coordination between conceptual thought and sensory experience, Friedman expands the core of this account in his article "Synthetic History Reconsidered" (2010) to include a so-called historicized version of transcendental philosophy. What makes this enterprise truly transcendental, we are told, is the fact that the inner logic of successive intellectual situations from Kant to Einstein proceeds in the light of Kant's original theory.

1.6 RECONSIDERING RELATIVITY AND THE A PRIORI

In his review of *Dynamics of Reason* (2004), however, Marc Lange draws attention to critical ambiguities and possible historical inadequacies in Friedman's account. What Lange finds particularly inadequate in Friedman's argument about the existence of a priori constitutive principles is the lack of demonstration for their uniqueness. Unfortunately, Lange does not substantiate his own position that alternative sets of constitutive principles could be possible for each of the three spatio-temporal structures. This dissertation therefore attempts to provide an in-depth analysis and critique of Friedman's concept of the relativized a priori, taking into account the historical development of the special theory of relativity. It consists of three parts.

Part I: Friedman and the Dynamics of Reason

The first part examines Michael Friedman's concept of relativized a priori constitutive principles within Newtonian and relativistic physics: while chapter 2 examines the original 2001 conception, chapter 3 deals with Friedman's 2010 article.

Part II: The Development of Special Relativity Reconsidered

The second part of the dissertation investigates the history of the transition from Hendrik Anton Lorentz's 1895 ether-based theory of electrodynamics to Vladimir Ignatowski's derivation of the Lorentz transformations in 1910. Chapter 4 engages with Lange's review. Chapter 5 analyzes Lorentz's two-tier strategy in his 1895 theory of corresponding states: a general treatment of electrostatics and optics accounting for most first-order ether drift experiments, on the one hand, and a series of specific hypotheses, each introduced to explain a particular second-order experiment, on the other. Chapter 6 examines Henri Poincaré's view that physical theories work towards the fulfillment of certain general principles, while chapter 7 analyzes the Frenchman's criticism of Lorentz's theory in the light of his physics of principles. The latter was expressed in Poincaré's contribution to a conference held in 1900 in honor of the 25th anniversary of Lorentz's doctorate. In the first part of his paper, Poincaré proved that Lorentz's theory did not satisfy the principle of reaction. Secondly, he constructed a thought experiment demonstrating that this result implies the possibility of perpetual motion in conjunction with the principle of relative motion. Finally, he showed how to resolve this apparent contradiction in Lorentz's theory by extending the principles of reaction and relative motion to include the ether. Poincaré was not satisfied with his solution, however, because the null result of ether drift experiments strongly suggested that the principle of relative motion applied only to matter. Chapter 8 examines Lorentz's 1904 electrodynamic theory, which attempted to meet the objections raised by Poincaré by formulating an improved version of the theory of corresponding states in which the invariance of optical phenomena held at every order. For other electrodynamic phenomena, Lorentz still believed that it might be possible to detect motion with respect to the ether. Chapter 9 studies Poincaré's reaction to Lorentz's 1904 theory. In 1905, the German physicist Max Abraham published a paper in which he proved the inconsistency of Lorentz's electron model within the framework of Lagrangian mechanics. Despite this inadequacy, Poincaré did not reject Lorentz's 1904 theory, as it enabled him to construct electrodynamics in such a way that it would rigorously meet what he now called the principle of relativity. Mathematically, this was expressed by the invariance of Maxwell's equations under the Lorentz transformations. Even though Lorentz invariance implied the impossibility of distinguishing ether-based reference frames from other inertial frames, Poincaré retained the ether as the carrier of the electric and magnetic fields. While the Frenchman submitted the main article on his new theory in July 1905, it was first published in 1906. Chapter 10 analyzes Albert Einstein's article "Zur Elektrodynamik bewegter Körper," which was published in June 1905. In this article, Einstein demonstrated how to obtain the Lorentz transformations solely by means of the principle of relativity and the principle of the constancy of the speed of light. In particular, he did not introduce the ether and, accordingly, did not understand length contraction as a dynamical influence of the ether on matter. Instead, the contraction was explained as a kinematical consequence of the spatio-temporal structure of special relativity. Chapter 11 is devoted to another derivation of the Lorentz transformations, which goes back to an article published in 1910 by the Russian physicist Vladimir Ignatowski. Since the special theory of relativity created a common space-time framework for all other theories, Ignatowski's point of departure was that electrodynamic principles, such as the constancy of the speed of light, should not play a fundamental role in the derivation of the Lorentz transformations. He achieved this goal by deriving the Lorentz transformations and their Galilean limit by presupposing the principle of relativity, the homogeneity and isotropy of space, and the homogeneity of time. Which case applied was then a matter of empirical investigation.

This analysis of the historical development from Lorentz's 1895 theory of electrodynamics to Ignatowski's derivation of the Lorentz transformations reveals two major shortcomings in Friedman's interpretation of the transition from the Newtonian to the special relativistic framework. First, both Lorentz and Poincaré abandoned Newton's second and third laws of motion in their effort to reconcile electrodynamics with the null result of ether drift experiments. Contrary to Friedman's claims, they did consequently not constitute Newtonian space-time by utilizing Newton's laws of motion. Instead, they adhered to Newtonian space-time by elevating measurements conducted at rest relative to the alleged ether as real. Second, in contrast to Friedman's view of necessity in terms of uniqueness, Ignatowki encountered an alternative way of defining inertial frames using the principles of homogeneity and isotropy that would work properly for both Newtonian and special relativistic physics.

In addition to these two main points, part 2 contains other shortcomings in Friedman's account. For example, Lorentz's 1904 theory of electrodynamics did not fulfill the principle of relativity. In particular, Lorentz did not have a physical interpretation of local time in terms of dilated clocks. Instead, the concept of local time was a purely mathematical construct in the theory of corresponding states. Moreover, in his 1905 Lagrangian formulation of electrodynamics, Poincaré did not introduce Lorentz's mechanisms, such as electron contraction, as a primary hypothesis to prove the principle of relativity. On the contrary, by presupposing that the Lagrangian had to meet the principle of relativity, Poincaré derived how space and time coordinates had to transform under boosts of a laboratory frame from rest to uniform rectilinear motion relative to the ether. In line with the reverse strategy, Poincaré did not elaborate on the procedures for measuring space and time in his 1905 paper. Nevertheless, his Sorbonne lectures of 1906 document that he retained the ether to distinguish between real and apparent measurements. Although the lectures showed why space and time measurements had to be transformed according to the Lorentz transformations, they did not explain how the dilation of clocks was possible.

Part III: The Comprehensibility of Nature

Given the above inadequacies, the final part of the dissertation attempts to pave the way for a better understanding of the rationality associated with the transition from Newtonian to special relativistic physics, using Olivier Darrigol's concepts of comprehensibility principles and interpretive schemes. The first part of chapter 12 introduces Darrigol's comprehensibility principles and shows how they illuminate key parts of the historical analysis of part 2. Rather than purporting to define reference frames empirically, comprehensibility principles are characterized by a theory-generating power that facilitates rational inference. Examples include Poincaré's effort to derive Lorentz's mechanisms by applying the principle of relativity to the Lorentz-Maxwell equations, and Ignatowki's endeavor to free special relativity from its contingent ties to electrodynamics by proving the Lorentz transformations on the basis of the relativity principle and the requirements of homogeneity and isotropy. The second part of chapter 11 presents Darrigol's interpretive schemes. While comprehensibility conditions serve as premises in deriving a theory, they do not by themselves tell us how to apply the constructed theory to experience. The latter depends on the empirical realization of interpretive schemes that describe ideal devices and measurement procedures. Examples include Poincaré and Einstein's methods of synchronizing clocks by light signals or Roberto Torretti's design of ideal clocks based on the conservation of angular momentum. Chapter 13 examines the connections between the space-time structure and the laws of mechanics. Ignatowki gave an alternative definition of inertial frames through the conditions of homogeneity and isotropy. This possibility suggests a connection between these conditions and the laws of mechanics. Lev Landau and Evgeny Lifshitz (1969, 1972) carried out the most prominent investigation of this connection. First, by applying the principle of least action in conjunction with Ignatowki's definition of inertial frames, Landau and Lifshitz demonstrated how to derive the conservation of both linear and angular momentum. Second, by requiring that interactions propagate at either infinite or finite speed, they derived the remainder of Newtonian mechanics for the infinite case and the remainder of special relativistic mechanics for the finite case. While the second proof explains how different sets of principles can define the same group of inertial frames, the first derivation allows us to formulate an interpretive scheme for the construction and synchronization of ideal clocks shared across Newtonian physics, special relativity and ether-based electrodynamics. The latter is the topic of chapter 14. Both Poincaré and Einstein imagined the synchronization of given ideal clocks by light signals. In comparison, Roberto Torretti based the construction and synchronization of Newtonian clocks on the conservation of linear and angular momentum in his book *Relativity and Geometry* (1983). Following the article "The Relativistic Trolley Paradox" (2016) by Vadim Matvejev, Oleg Matvejev and Øystein Grøn, I show how Torretti's interpretive scheme can be extended to the frameworks of special relativity and ether-based electrodynamics.

To finally challenge Friedman's concept of constitutive a priori principles in chapter 15, we need only point to the existence and realizability of alternative schemes. While the arrival of the Global Positing System (GPS) has confirmed the realizability and effectiveness of clock synchronization using electromagnetic waves, a decade ago researchers at NASA developed spherical gyroscopes similar to Torretti's ideal clocks as part of the Gravity Probe B experiment. In other words, there is a definite answer to Lange's question: Friedman's uniqueness claim is both logically and empirically false. The American might reply that my account of the logical and empirical possibility of alternative schemes cannot, because of its retrospective character, call into question his historically oriented view of constitutive a priori principles. In particular, this account arguably tells us little about the historical development from Kant and Newton to Einstein and therefore cannot falsify his historicized version of transcendental philosophy. However, as my historical investigation shows, the latter has its own inadequacies.

Rather than endorsing Friedman's understanding of constitutive a priori principles, the dissertation advocates Darrigol's conception of the relativized a priori as presuppositions of so-called comprehensibility principles on which we base deductions employed in theory construction. For example, Einstein showed how the principle of light and the assumptions of homogeneity and isotropy necessitate a new spatio-temporal framework. Nevertheless, these principles do not define the new framework empirically in and by themselves. The latter requires both interpretive schemes and empirical know-how. Accordingly, it is neither problematic for the status of these conditions that other sets of conditions can also imply the same framework, nor that multiple interpretive schemes can empirically realize this structure. Moreover, by emphasizing the dialectical character of necessity and openness of the historical development of mathematical physics, the dynamic interplay between comprehensibility conditions and interpretive schemes promises a more adequate philosophical understanding of this enterprise. Friedman might counter that the comprehensibility of nature merely displaces the transcendental character of mathematical physics to theory construction and therefore welcome the distinction between comprehensibility conditions and interpretive schemes as a clarification of the dynamics of reason.

Part I

FRIEDMAN AND THE *DYNAMICS OF REASON*

CHAPTER 2

Dynamics of Reason

This chapter is devoted to an analysis of Friedman's book *Dynamics of Reason*. It was published in 2001 and should be considered a classic in the tradition of scientific philosophy. Friedman's aim in this book is to combat what he sees as two major dilemmas in our philosophical situation today: the naturalistic predicament in the figure of Quinean epistemological holism and post-Kuhnian conceptual relativism most prominently attributed to the Edinburgh School of Barnes and Bloor. The first step of Friedman's attack is to meet the naturalistic challenge by formulating a concept of dynamic and relativized a priori constitutive principles within mathematical natural science. His second step is to develop a concept of universal, transhistorical scientific rationality in response to sociological conceptual relativism.

2.1 CHALLENGING NATURALISM

What Friedman finds most lacking in Quine's holistic conception of scientific knowledge as a vast web of interconnected beliefs on which sensory impressions act only at the periphery is that it portrays the totality of the various parts of our system of knowledge – scientific disciplines such as mathematics, physics and geography – as functioning symmetrically in relation to adjacent sensory experiences. For Quine, the web of beliefs is therefore not a stratified system, but rather a flat, holistic structure whose different parts equally face "the tribunal of sensory experience"[1] in such a way that absolutely "no statement is immune to revision:"[2]

> [T]otal science is like a field of force whose boundary conditions are experience. A conflict with experience at the periphery occasions readjustments in the interior of the field [...] But the total field is so undetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to re-evaluate in the light of any single contrary experience. [...] Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. [...] Conversely, by the same token, no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying

[1] Quine 1951, p. 38.
[2] Ibid., p. 40.

> quantum mechanics; and what difference is there in principle between such a shift and the shift whereby Kepler superseded Ptolemy, or Einstein Newton, or Darwin Aristotle?[3]

Friedman's starting point for criticizing this anti-apriorist conception of knowledge is the vehement denial that this supposedly scientifically oriented version of anti-apriorism gains any plausibility from the historical development of mathematical natural science from Ptolemy via Kepler and Newton to Einstein. On the contrary, from Friedman's point of view, the Quinean figure is not supported by revolutionary changes within mathematical physics, but rather undermined and disconfirmed. To reach this conclusion, Friedman draws on resources from a rather surprising direction, namely the logical-empiricist tradition represented by the early writings of Reichenbach and Carnap. The reason why this should come as a surprise is at least twofold. First, this tradition is usually seen as starkly opposed to neo-Kantian conceptions of thought, and second, it forms the criticized foil on which Quine begins to erect his position in the now famous article "Two Dogmas of Empiricism" (1951) (from which I have quoted above) and by which, according to the usual telling of the story, he simultaneously sealed the execution of the logical-empiricist tradition that was finally carried out eleven years later in Kuhn's book *The Structure of Scientific Revolutions* (1962). However, according to Friedman, the simplifications and drawbacks of this account of the history and the decline of logical empiricism, which is firmly anchored in the Quinean naturalistic tradition, have been pointed out in the last two decades by historically oriented philosophers such as Alberto Coffa, Alan Richardson, Robert DiSalle and himself.

2.2 REICHENBACH AND THE RELATIVIZED A PRIORI

In Friedman's view, the important moral of these historical studies is that the philosophical outlook of early Reichenbach and early Carnap was so strongly influenced by Kantian transcendental philosophy that they did not simply reject Kant's conception of the synthetic a priori outright in the light of Sir Arthur Eddington's experimental confirmation of the phenomenon of light bending in 1919 predicted by Einstein's general theory of relativity due to the application of a non-Euclidean space-time structure. In contrast to the reactions from the camp of Kantian orthodoxy to

[3] Quine 1951, pp. 39–40.

Einstein's theory and Eddington's findings, however, the two scientifically educated philosophers did not outright reject the theory of relativity from a supposedly higher philosophical vantage point. Instead, they attempted to develop a new, but still unambiguous Kantian notion of a priori principles, which at the same time carefully took into account the revolutionary developments in mathematics and mathematical physics since Kant. According to Friedman, the clearest version of this new conception of the a priori was formulated by Reichenbach in his first book *Relativitätstheorie und Erkenntnis apriori*, published in 1920. Reichenbach's first insight was to distinguish two different meanings of Kant's conception of the synthetic a priori:[4]

1. necessarily true or true for all times,
2. constituting the concept of object.

His next, and at the same time crucial, step was to argue that the profound lesson we must draw from relativity is that we must consign the former meaning to the flames while retaining the latter. In other words, Reichenbach argued, in agreement with Kant, that every physical theory necessarily contains a priori constitutive principles that ground the empirical meaningfulness of its mathematically formulated laws. The underlying idea is that scientific (i.e. truly objective) knowledge must be formulated or framed in a mathematical formalism. However, for this to be possible in the first place, coordinating principles are necessary, whose function is to establish a mapping between pure mathematics and concrete empirical phenomena, so that the laws acquire a definite empirical meaning. We must therefore distinguish between constitutive coordinating principles ('axioms of coordination') and proper empirical laws ('axioms of connection'). Nevertheless, Reichenbach went beyond Kant by claiming that these principles of coordination change and evolve in the transition from one theory to its successor. Therefore, we must abandon Kant's belief in a fixed set of constitutive a priori principles by moving to a conception of relativized and dynamic constitutive (and in this sense still) a priori principles. Friedman's basic contention, to which we now turn, is that this early logicist-empiricist conception of the a priori provides a much more accurate and vivid account of the historical development of mathematical natural science than Quine's anti-apriorism allows.

[4] Reichenbach 1965, p. 48.

2.3 THE TRIPARTITE STRUCTURE OF NEWTONIAN PHYSICS

Friedman thus sets himself the task of further developing Reichenbach's analysis of the constitutive a priori by explaining in more detail how this conception should be applied in the attempt to elucidate the structure of Newtonian, special relativistic and general relativistic space-time theory. The crucial question in this context is which part of each theory Friedman identifies as fulfilling the constitutive function of coordination. With regard to Newtonian space-time, he states right at the beginning that Newton himself distinguished three parts of his theory:

1. A calculus of infinitesimals for conceptualizing infinite limiting processes,

2. concepts of force and matter determined by the three laws of motion,

3. the law of universal gravitation.

It is a historical fact that Newton developed each of the three components – mathematics, mechanics and gravitational physics – to create a unified theory of terrestrial and celestial motion. But, Friedman argues against Quine, this alone does not mean or imply that the three parts function symmetrically in the overall conjunction. According to Friedman, the claimed asymmetries come to light when we take a closer look at the mutual relationships between the three components. To start with the relationship between the mathematical part and the mechanical part, Newton's second law of motion states that the (impressed) force $\mathbf{F}$ on a body is equal to the (inertial) mass m of the body times its acceleration $\mathbf{a}$, where the acceleration is defined as the double derivative $\ddot{\mathbf{x}} = \mathrm{d}^2\mathbf{x}/\mathrm{d}t^2 (= \mathrm{d}\mathbf{v}/\mathrm{d}t)$ of the body's position $\mathbf{x}$. In other words, this mechanical law cannot even be formulated without the infinitesimal calculus, and for this very reason it is a fundamental misconception to regard the two parts as fulfilling symmetrical functions in relation to the overall theory. Rather, we should understand the mathematical part as providing "elements of the language or conceptual framework, we might say, within which the rest of the theory is then formulated."[5]

For Friedman, there is a similarly asymmetrical relationship between mechanics and gravitational physics. The gravitational part of Newton's theory is encapsulated in the law of universal gravitation, which states that

5 Friedman 2001, p. 36.

between any two bodies in the universe there is a (gravitational) force of attraction whose magnitude is directly proportional to the product of the two masses and inversely proportional to the square of the distance between them. Friedman notices that this law is only empirically well defined if we know in which frame of reference the accelerations in question are defined. In other words, the law presupposes a privileged frame of reference in relation to which its empirical meaning is fixed. Such a frame of reference is called an inertial frame and is itself determined as a frame in which the three laws of motion apply. From this Friedman concludes: "[W]ithout the Newtonian laws of mechanics the law of universal gravitation would not even make empirical sense, let alone give a correct account of the empirical phenomena."[6] Another way to illustrate the asymmetry between the law of gravitation and its mechanical counterparts would be to note that the mechanical laws, unlike the universal law of gravitation, do not assert the existence of a specific type of force. For example, the second law only says that *if* a body is acted upon by an (impressed) force, *then* this force is related to the mass of the body in the particular way prescribed by the law. The three laws together thus rather function as a kind of constitution of concepts like force and acceleration figured in empirical laws such as the law of universal gravitation. For Friedman, therefore, these laws are not proper empirical laws, but rather the kind of a priori constitutive principles he was looking for.

2.4 THE TRIPARTITE STRUCTURE OF GENERAL RELATIVITY

So much for Newtonian physics. Friedman goes on to argue that an analogous account of the role of constitutive a priori principles also applies to general relativity. In particular, this theory also consists of an asymmetric tripartite structure consisting of pure mathematics, constitutive a priori coordination principles and properly empirical laws. The crucial difference between the theories from Friedman's point of view is that the constitutive framework comprising the purely mathematical components and the a priori coordination principles has changed substantially in the historical development from its Newtonian predecessor to its revolutionary Einsteinian successor. For example, while Newtonian physics is encapsulated in Euclidean geometry and infinitesimal calculus, general relativity is framed in the non-Euclidean geometry developed by Bernhard Riemann in his gener-

[6] Ibid., p. 36.

alized theory of manifolds. This revolutionary new mathematics uses the concept of a variably curved space-time geometry, which could not even be formulated at the time of Newton and consequently could not play a constitutive role in Newtonian physics. For this very reason, the Newtonian coordination principles cannot fulfill a similar constitutive function in relation to this revolutionary new kind of pure geometry. The application of Riemann's new ideas within the framework of the general theory of relativity thus presupposes that Newton's laws of motion are replaced by a new form of coordination. What Einstein came up with was the principle of equivalence, which identifies the trajectories of bodies that are only influenced by gravity as geodesics (straightest possible paths) in the variably curved space-time geometry. According to Einstein's theory, the variable curvature itself is determined by the gravitational field equations as a function of the mass-energy distribution in the universe. Friedman further explains the relationship between the principle of equivalence and the field equations by pointing out that the latter would have no empirical meaning without the equivalence principle. For without the principle, we would not have singled out some empirical phenomena as counterparts to the mathematical concepts determined by the equations – in this case, the concept of geodesics: "The principle of equivalence does precisely this, however, and without this principle the intricate space-time geometry described by Einstein's field equations would not even be empirically false, but rather an empty mathematical formalism with no empirical application at all."[7] Again, Friedman concludes that the three parts that make up general relativity theory function asymmetrically, in the sense that the first two parts function as a constitutive framework for the third and thus only proper empirical part. The general conclusion we should draw from the development of mathematical natural science since the heyday of Newton thus fundamentally contradicts Quinean epistemological holism:

> In periods of deep conceptual revolution, it is precisely these constitutively a priori principles which are themselves subject to change – under intense pressure, no doubt, from new empirical findings and especially anomalies. [...] [But t]he idea of necessary presuppositions constituting the conditions of possibility of the properly empirical parts of a scientific theory is not undermined [on the basis of

[7] Friedman 2001, p. 38.

these revolutionary developments] but rather confirmed![8]

2.5 LOGICAL SPACE AND EMPIRICAL SPACE

In order to clarify the concept of constitutivity outlined above, Friedman introduces a distinction between the space of logical possibilities on the one hand and the space of empirical possibilities on the other. Friedman begins the discussion that leads to the introduction of this distinction by pointing out that the role of what he calls a priori constitutive principles is to provide “the necessary framework within which the testing of properly empirical laws is then possible.”[9] For, as we have seen, without the prior acceptance of a constitutive framework, no proper empirical laws have been assigned to concrete empirical phenomena, and for this very reason they still belong to the realm of abstract mathematics. For example, counting Eddington’s experimental test of the phenomena of light bending in the vicinity of heavy bodies such as the sun as empirical evidence for general relativity *presupposes* the a priori acceptance of the constitutive framework of this very theory.[10]

According to Friedman, we can and should therefore distinguish between two very different understandings of empirical theory testing, depending on whether we evaluate the experiment from an internal or external perspective; in our case, from the internal perspective of an Einsteinian physicist or the external perspective of a Newtonian one. Friedman’s point is that both physicists agree that Einstein’s theory provides more accurate predictions. That is, from a purely pragmatic or instrumental perspective of theory testing, they agree. Unlike the Einsteinian, however, the Newtonian can only accept this calculation as a kind of black-box prediction because, from his point of view, “the constitutive framework of general relativity is not even possible or coherent, and there is thus no sense in which Einstein’s field equations can actually be empirically true.”[11] In Friedman’s account, then, the internal perspective facilitates a conception of empirical theory testing that is stronger than a purely instrumental account, in

[8] Ibid., p. 45.

[9] Ibid., p. 83.

[10] In contrast to the classical physicists, true adherents of Einstein should resist speaking of the phenomena of light bending. This is because, according to the general theory of relativity, light travels along geodesics of space-time and thus along the straightest possible paths.

[11] Friedman 2001, p. 84.

the sense that the former (but not the latter) allows for the notion of genuine empirical evidence in the context of an already accepted framework. In other words, a constitutive framework defines what Friedman calls a *space of empirical possibilities* that acts as a background against which experiments can then determine which empirical possibilities are actually the case. Furthermore, Friedman introduces the notion of *space of logical possibilities*, which is given by the purely mathematical part of a constitutive framework. For instance, "[w]ithout the Riemannian theory of manifolds, we might say, the space-time structure of general relativity is not even logically possible, and so, a fortiori, it is empirically impossible as well."[12] The difference between these two notions is that logical possibility is supposed to be a necessary but not sufficient presupposition of empirical possibility, i.e. the latter is defined only when the logical possibilities under consideration are successfully coordinated with some concrete empirical phenomena. According to Friedman, then, it is precisely the constitutive or coordinating principles that ultimately define what he calls a real as opposed to a merely logical possibility. What the distinction brings to the fore is first of all the clarification that the space of logical possibilities of our constitutive framework is independent of its space of empirical possibilities: the former is presupposed by the latter. For example, it is possible to hold that the empirical possibility of general relativity is incomprehensible and at the same time affirm the comprehensibility of Riemann's theory of manifolds, which constitutes the logical space of general relativistic possibilities.

2.6 CONSTITUTIVE PRINCIPLES AS NECESSARY PRESUPPOSITIONS

At the beginning of the second part of his book, Friedman tries to explain in more detail "what exactly it means for (putatively) a priori principles to be necessary conditions of empirical knowledge."[13] The first thing he notices is that the statement that *A* is a constitutive condition of *B* means, in more standard terminology, that *A* is a *presupposition* of *B*. In other words, "*A* is a necessary condition, not simply of the truth of *B*, but of *B*'s meaningfulness or possession of a truth value."[14] Within the framework of Newtonian physics, for example, we *only* know how to apply the universal law of gravitation if we presuppose the truth of the laws of motion: The former has

[12] Friedman 2001, p. 84.
[13] Ibid., p. 73.
[14] Ibid., p. 73.

empirical content only relative to an inertial frame defined as a reference frame in which the laws of motion apply. Friedman then notes, however, that the idea of a presupposition alone is by no means sufficient to delimit the Kantian concept of the necessary condition, which he attempts to revive:

> Yet the mere idea of a presupposition in this sense is of course much too weak to capture the Kantian notion we are after [. . .] We want to reserve this characterization for particularly fundamental presuppositions lying at the basis of mathematical physics – principles which, accordingly, can plausibly be taken as fundamental presuppositions of *all* empirical truth (at least in the natural sciences).[15]

Without imposing such an additional requirement, Friedman notices that any presupposition would automatically count as a constitutive principle in his sense. However, he emphasizes that he must add a qualification, since this is certainly not his intention. Admittedly, his characterization of constitutive principles as "principles which [. . .] can plausibly be taken as fundamental presuppositions of *all* empirical truth" helps to grasp the Kantian notion he is alluding to. The question remains whether this helps to clarify things. For what reasons do we have to adhere to the existence of such fundamental principles? In particular, what does it mean to recognize whether it is *plausible* that a principle is presupposed by *at least all* natural scientific knowledge?

Although Friedman does not take this question at face value, his discussion of the laws of motion undoubtedly indicates that he regards the latter as necessarily presupposed by not only the universal law of gravitation, but by all empirical knowledge that can be obtained within the framework of Newtonian physics. For within the framework of Newtonian physics, the laws of motion define the unambiguous spatio-temporal arena of all empirical laws:

> [A]s we now understand it, these laws of motion define a privileged class of relative spaces or reference frames (what we now call inertial frames) in which the modern concepts of space, time, and motion then unambiguously apply [. . .] The Newtonian laws of motion are

[15] Ibid., p. 74.

> thus presuppositions of the properly empirical laws of Newtonian physics (such as the law of gravitation).[16]

Insofar as other contemporary branches of natural science, such as chemistry and geology, also used standard methods to measure time and place, we can add that they too presupposed the laws of motion. According to Friedman, the Newtonian laws of motion thus empirically define concepts of space, time and motion that every natural science of the time at least implicitly applied. Furthermore, Friedman's account of the necessity of constitutive principles implies, as we shall see later, that only the Newtonian laws of motion enable us to empirically define the spatio-temporal framework of Newtonian physics. If this were not the case, it would be difficult to understand how these laws can function as necessary presuppositions of proper empirical laws. In other words, the universal law of gravitation itself can only presuppose that the Newtonian space-time structure is somehow empirically well defined. According to Friedman's account, however, this can only be the case through Newton's laws of motion. This brings us to Friedman's final characteristic of constitutive principles. Since their only function is to coordinate or mediate between abstract mathematical formalism on the one hand and concrete empirical phenomena on the other, they are themselves assigned to a special class of non-empirical physical principles.

2.7 THE MICHELSON-MORLEY EXPERIMENTS

Friedman begins his argument by stating that the famous experiments of Michelson and Morley appear to provide a direct empirical test of the so-called light principle. This principles states that the speed of light is the same in every inertial frame regardless of the speed of the light source, and in Friedman's view it is a constitutive principle within the framework of special relativity:

> The famous interferometer experiments of Michelson and Morley (1882, 1887), for example, which result in no detectable influence of the motion of the earth on the velocity of light, seem to supply as good an empirical test as can be imagined for the invariance of the velocity of light in different inertial frames and thus for the light principle as it is first introduced in the special theory of relativity.[17]

[16] Friedman 2001, pp. 76–7.
[17] Ibid., p. 86.

But despite appearances, Friedman goes on to argue that the null result of these experiments cannot be taken as a crucial experiment in favor of the relativistic light principle, even though the principle would be empirically untenable if the experiments had not had this result:

> A coordinating principle must always have a counterpart in reality, and, if such a counterpart does not exist, the principle is empirically vacuous and thus useless.[21] The crucial question, however, is whether such a principle can thereby become empirically false.[18]

The empirical counterpart to the constitutive framework of Newtonian physics corresponds, for example, to the existence of an inertial frame in which the laws of motion hold with a high degree of approximation. Friedman himself wants to enforce the conclusion that such principles cannot be empirically tested in the same way as the empirical laws themselves. In the case of the light principle, for example, he argues that the result of the Michelson-Morley experiments only shows that we cannot detect any effects on the behavior of light due to the motion of the earth in relation to the ether. The experiment therefore leaves other alternatives open than the special relativistic interpretation of the null result in terms of the invariance of the speed of light across inertial frames. What Friedman is getting at, of course, is the historical fact that Lorentz integrated the result of the Michelson-Morley experiments into the Newtonian spatio-temporal framework. In Friedman's view, this is the reason why we cannot consider the experiment as a true empirical test of special relativity with respect to its alternatives.

2.8 LORENTZ VS. EINSTEIN

To clarify this matter, Friedman attempts to make clearer what he sees as the fundamental difference between Lorentz's electrodynamical theory and Einstein's special theory of relativity by comparing the choice between them to cases of empirical underdetermination. In contrast to the later cases, he emphasizes, the two alternatives have no constitutive framework in common:

> Whereas Lorentz and Fitzgerald take an essentially classical background structure for space, time, and motion to be already sufficiently well defined and only subsequently locate the new empirical

[18] Ibid., p. 87.

> discovery in question [the null result] as a peculiar (but additional) empirical fact formulated against the background of this classical structure, Einstein calls the whole classical structure into question and uses the very same empirical discovery empirically to define a new fundamental framework for space, time, and motion entirely independent of the classical background.[19]

As we have seen, the empirical counterpart of the classical background structure amounts to the existence of a set of inertial frames defined as reference frames in which the laws of motion hold. In other words, these are not empirical laws, but constitutive principles or empirical criteria that a reference frame must fulfill in order to count as an inertial frame. What Lorentz did in this context, according to Friedman's point of view, was to presuppose that the setup for the Michelson-Morley experiments was at rest relative to an already sufficiently well-defined inertial frame in the Newtonian sense. In contrast, Einstein defined a new notion of an inertial frame in which Newton's second and third laws of motion were replaced by the light principle as a constitutive condition:

> Einstein uses his light principle *empirically to define* a fundamentally new notion of simultaneity and, as a consequence, fundamentally new metrical structures for both space and time (more precisely, for space-time).[20]

From Friedman's point of view, the null result of the Michelson-Morley experiments is thus not seen as an empirical discovery within an already well-defined spatio-temporal framework, but rather shows that the rest frame of the laboratory in question fulfills the new constitutive condition, which means that this rest frame corresponds to an inertial frame in the special relativistic sense. As Friedman notes, the authors influenced by Poincaré express this altered epistemological function of the light principle by saying "that Einstein has 'elevated' an empirical law to the status of a convention – or, as I myself would prefer to put it, to the status of a coordinating or constitutive principle."[21] According to Friedman, this formulation makes sense insofar as "[i]t is precisely here that an essentially non-empirical element

[19] Friedman 2001, p. 88.
[20] Ibid., p. 88.
[21] Ibid., p. 88.

of 'decision' must intervene, for what is at issue, above all, is giving a radically new space-time structure a determinate *empirical meaning* – without which it is not even empirically false but simply undefined."[22] From this point of view, the *decision* to define the concept of an inertial frame by means of the light principle does not at all hinge upon the null result of the Michelson-Morley experiments. As Friedman points out, this is consistent with Einstein's explanation that this experimental result did not play a prominent role in his development of special relativity. To define a particular spatio-temporal structure empirically by selecting a definite set of coordination principles is thus very reasonable, even if we do not know whether the chosen set of principles will have an inertial frame as its empirical counterpart.

But although Friedman here draws attention to what he sees as the *conventional* character of the coordination of a particular mathematical structure with empirical phenomena, he emphasizes at the same time that he certainly does not mean that we have no empirical motives for preferring one coordination to another. He writes about the case we are considering:

> Indeed, the new empirical discovery in question – undetectability of differences in inertial motion in electrodynamics – provides us with strong empirical motivation, not only for entertaining a new coordination, but also (as Einstein was apparently also the first to see) for doubting the adequacy of the classical coordination. For, if there were in fact an empirical counterpart to the classical notion of absolute simultaneity, then there would be (in the context of electrodynamics) an empirical counterpart to absolute velocity as well. But the new empirical discovery strongly suggests that there is no such empirical counterpart [...] The classical spatio-temporal structure, which we had assumed in the context of Newtonian physics to be unproblematically empirically well defined, thereby turns out to be empirically meaningless.[23] Here we certainly have an empirical motivation, and a particularly strong one, for preferring the new empirical coordination effected by Einstein.[23]

Friedman seems to argue that the result of the Michelson-Morley experiments gives us particularly strong reasons to doubt not only that the lab-

[22] Ibid., p. 88.
[23] Ibid., pp. 88–9.

oratory frame corresponds to a Newtonian inertial frame, but the very existence of such frames altogether. For since the laws of Maxwellian electrodynamics are Lorentz-invariant, in the framework of Newtonian physics they should only hold true in a frame at rest relative to the ether, which in turn would make "absolute velocity" empirically well defined. Against this background, the undetectability of differences in inertial motion in electrodynamics should be some kind of empirical motivation "for preferring the new empirical coordination effected by Einstein," even though Friedman immediately reminds us that "this situation is not happily likened to more standard cases of empirical underdetermination, where two empirically equivalent hypotheses face off against the background of a common constitutive framework, and methodological principles such as simplicity or conservativeness are then invoked to settle the question."[24]

In this section I have tried to make clear that Friedman begins to walk a fine line: he wants to fend off naturalism by developing a notion of relativized a priori constitutive principles, but without committing himself to one or another version of conceptual relativism. The first indication of Friedman's balancing act is his introduction of a distinction between first-order empirical evidence, or direct empirical evidence, on the one hand, and second-order empirical evidence, or empirical motivation, on the other. The former are formulated against the background of a constitutive framework in which the space of logical possibilities has been given an empirically well-defined spatio-temporal realization by means of a set of non-empirical constitutive principles. In other words, under conditions that we can control in the laboratory, measurements provide us with first-order empirical evidence to determine which empirical possibilities happen to be the case. As mentioned above, Friedman himself refers to such experiments as genuine empirical tests, or empirical tests that are evaluated from the internal perspective of a particular constitutive framework. On the other hand, an empirical motivation is not meant to express direct empirical evidence. For if this were the case, it would be very difficult to understand how Friedman can claim that constitutive principles are non-empirical in nature. He makes this point by emphasizing that ultimately we still have the freedom to decide whether or not we want to change our set of constitutive principles. Therefore, strictly speaking, it is not irrational to keep the old constitutive framework. But if this is so, we can ask what kind of reason an

[24] Friedman 2001, p. 89.

empirical motivation corresponds to? As we will see below, Friedman introduces the notion of empirical reason by linking it closely to his account of direct empirical evidence. His concept of rationality is therefore an intra-paradigmatic concept from the outset. The question therefore arises as to how Friedman extends his account to include empirical motivations. How does he explain the development of a priori constitutive principles as a rational process?

2.9 THE RELATIVISTIC CHALLENGE

Friedman's point of departure is that this development has a purely logical side on the one hand and an empirical side on the other. The first problem consists in the question of how one can picture the expansion of the space of logical possibilities. Again, Friedman's favorite example is the development from Euclidean geometry to Riemann's theory of manifolds, enclosing the former as a special case of possible manifolds. The latter problem is the question of how to understand the development of a priori coordinating principles that constitute the space of empirical possibilities. Friedman addresses these two questions by clarifying why and how the problem of the change of a priori principles is related to Kuhn's well-known discussion of incommensurability and its sociological flavored offspring of conceptual relativism. In particular, Friedman points out that the distinction between instrumental and communicative rationality formulated by Jürgen Habermas applies to the situation at hand insofar as it is consistent with the distinction between an instrumental conception of empirical theory testing and Friedman's internal conception of testing:

> Standards of communicative rationality are given by what I [. . .] call an empirical space of possibilities or space of reasons, in that agreement on the constitutive principles definitive of such an empirical space of possibilities (mathematical principles and coordinating principles) yields agreement on what can count as an empirical reason or justification for any given empirical possibility. A shared constitutive framework thereby facilitates shared mutually comprehensible rational argumentation.[25]

The first thing Friedman points out in this passage is that any empirical space of possibilities involves 1) a purely mathematical structure that con-

[25] Ibid., p. 93.

stitute the internal or formal relationships between any two points in that space, and 2) a set of coordinating principles that determine how that structure is coordinated to empirical phenomena that humans can observe. For example, we can model the surface of the earth as a 2-dimensional mathematical sphere using the laws of astronomy as coordination principles. If we accept this coordination, any sailor well-trained in the art of navigation will not only be able to determine the position of his ship, i.e. the empirical possibility it occupies, but he will also be able to rationally justify why he is certain that the ship is where he takes it to be. But of course he cannot convince a person who believes that the earth is flat, nor a person who has no idea about navigation (the latter may still believe the sailor because he happens to know that the sailor has passed navigation school). In particular, the sailor's identification of the ship's position through the correct use of the sextant and the nautical table is only valid as an empirical reason or justification for people who already agree with the constitutive principles of our earth model. In other words, only what I have called the internal perspective enables rational argumentation. From an external perspective, the sailor's ability to maneuver the ship from one destination to another can only be evaluated as an expression of some kind of instrumental rationality or pragmatic skill. To take another example: The Newtonian's stubbornness in not accepting Einstein's constitutive framework leads him to be unable to regard Einstein's prediction of the phenomenon of light bending as anything more than a black-box prediction. This means that the Newtonian evaluates every single application of Einstein's theory as a purely instrumental rational act. In other words, the Newtonian is unable to understand such an application as a communicative rational activity in the sense that it is based on a shared constitutive framework that enables "constitutive mutually comprehensible rational argumentation." What Friedman takes all of this to show is that the shift from one constitutive framework or paradigm to its successor raises the challenge of conceptual relativism, for "what is rationally acceptable within one paradigm may not be so according to the standards [...] of the other."[26]

2.10 THE IDEA OF PROSPECTIVE COMMUNICATIVE RATIONALITY

For Friedman, then, there is a fairly obvious sense in which Kuhn's idea of incommensurability or non-intertranslatability between successive consti-

[26] Friedman 2001, p. 95.

tutive frameworks in a scientific revolution has real significance. Moreover, Kuhn's defense of scientific rationality based on the increase of solved scientific puzzles in (truly revolutionary) transitions from one paradigm to its successor is inadequate. In Friedman's view, this is because Kuhn's attempt to meet the relativistic challenge turns out to reflect a purely instrumental or pragmatic conception of rationality. And as Friedman has argued, in an important sense such an account does not answer the real problem, but rather points to a quite uncontroversial fact, namely that even the stubborn Newtonian must chalk up Einstein's predictions as a pragmatic success. The deeper problem of incommensurability, which was addressed above, thus remains untouched:

> Our problem [...] is to explain how a revolutionary transition from one scientific paradigm or constitutive framework to another can be communicatively rational, despite the fact that we are in this case faced with two essentially different and even incommensurable 'logical spaces.' [...] How [...] can it ever be (communicatively) rational to accept the later constitutive framework? How, in particular, can there ever be empirical evidence that counts as an empirical reason, in our sense, in support of the later framework?[27]

To overcome the threat of conceptual relativism, for example, one must articulate why Newtonians have good (communicative) reasons to give up their stubbornness and abandon their previously preferred constitutive framework in favor of Einstein's. Above all, this means that the usual gambit of pointing to the phenomenon of convergence between frameworks (prominently exemplified by the transitions from Newtonian physics via special relativity to general relativity) as a supposed bulwark against relativism is insufficient, because the idea of such convergence can only be communicated to someone who is already looking from the *retrospective* and thus from the wrong point of view: Convergence is achieved by showing how the old space of possibilities can be acquired as an approximate special case of the new space. What we need is not a retrospective reformulation of the old constitutive framework out of the new, but rather what Friedman calls an explanation from the *prospective* stance: we need to articulate in what sense it can be prospectively rational to shift from the old paradigm to the new. For example, we need to explain why even the stubborn

[27] Ibid., pp. 95–6.

Newtonian has good (communicative) reasons to become an Einsteinian.

So let us look at Friedman's explanation of how it can become prospectively rational to give up one space of empirical possibilities for another. The first thing Friedman notes is that, since we first have a new standard procedure of empirical testing after the transition, the problem is not to explain why it is rational to accept the truth of a properly empirical theory such as Einstein's new theory of gravitation. Rather, the problem is to give an account that explains "how Einstein's new theory of gravitation becomes a rational or reasonable possibility in the first place – to explain, as it were, how it first became a live option."[28] He then goes on to identify the comparison of scientific groups working within different constitutive frameworks to linguistic communities of radically separate languages as the problematic core of Kuhn's account of incommensurability. Although Friedman himself adheres to the idea of incommensurability, he argues that the new constitutive framework is always formulated in response to problems and concerns raised within the old framework. And it is for this reason that Friedman feels compelled to conclude that "succeeding conceptual frameworks in a scientific revolution are more akin to different stages of development within a common linguistic or cultural tradition."[29] From this observation it is then only a single step to the articulation of the idea of prospective inter-paradigmatic rationality at which he aims: the concepts and principles of the new constitutive framework not only converge retrospectively to the concepts and principles of the old framework under certain boundary conditions, "but they also develop out of, and as a natural continuation of, the old concepts and principles."[30] Thus, although the practitioners of the new framework apply concepts and principles that have no counterpart in the old framework, they are nevertheless able to rationally invoke the practitioners of the old framework "using empirical and conceptual resources that are already available at precisely this earlier stage."[31]

With this in mind, Friedman returns to Einstein's introduction of the light principle into the special theory of relativity. First, he notes that Einstein's work was not only formulated against the background of Lorentz's approach to electrodynamics, which attempted to accommodate the null result of the Michelson-Morley experiments in a classical spatio-temporal

[28] Friedman 2001, p. 100.
[29] Ibid., p. 100.
[30] Ibid., p. 101.
[31] Ibid., p. 101.

structure, but that it was also related to the development of the concept of inertial frame in the late nineteenth century, which clarified the issue of absolute versus relative motion in Newtonian mechanics. Thus, when Einstein "took an already well-established empirical fact (the empirical indistinguishability of different inertial frames by optical and electrodynamical means) and 'elevated' it to the status of a convention or coordinating principle,"[32] he was appealing to pre-existing conceptual resources and problems. Physicists working within the old framework were therefore not only able to understand the new coordination, but also prepared to recognize it as a reasonable live option.

The next point Friedman emphasizes is that these resources and problems do not belong to the scientific realm, but are clearly philosophical in nature. For example, to show that the long tradition of thinking about the question of absolute versus relative motion is largely philosophical in nature, we just have "to keep in mind the large amount of unresolved intellectual disagreement surrounding it up to the present day."[33] Thus, in addition to the levels of empirical laws and constitutive principles, we need to distinguish a third, meta-scientific level with clearly philosophical considerations:

> Here we are concerned with what I want to call meta-paradigms or meta-frameworks, which play an indispensable role in mediating the transmission of (communicative) rationality across revolutionary paradigm shifts, despite the fact that they are incapable, by their very nature, of the same degree of (communicatively) rational consensus as first-level or scientific paradigms.[34]

But, as Friedman immediately notes, distinguishing between philosophical and scientific reflection in this way leads directly to a new difficulty. If philosophical reflection is distinguished from scientific reflection because the former is incapable of achieving the same degree of rational consensus as the latter, how can philosophical reflection ever help to achieve a new rational consensus during scientific revolutions? How is it possible that the philosophical realm itself, characterized by unresolved questions and concerns, can facilitate rational agreement that the time has come to abandon the old constitutive framework in favor of a radically new one?

---

[32] Ibid., p. 102.
[33] Ibid., p. 105.
[34] Ibid., p. 105.

Friedman's answer to this question involves three steps. First, he notes that the only thing the required rational consensus has to accomplish is to make the new framework recognizable as a possible live option. Second, even though distinctively philosophical debates by their very nature do not lead to a stable consensus on the outcomes reached, Friedman argues that the scientific parties arguing against each other nevertheless manage to reach a relatively stable agreement on the important contributions to the debate. This means that each of them must be taken very seriously by a responsible attitude in the ongoing debate. Finally, he points out that distinctive philosophical reflection can shake up scientific reflection in such a way that "controversial and conceptually problematic philosophical themes become productively intertwined with relatively uncontroversial and unproblematic scientific accomplishments" so as "to facilitate the introduction and communication of a new scientific paradigm."[35]

Einstein's advent of the special theory of relativity again serves as one of Friedman's main examples. According to the latter, the first of the three points is evident in the case of special relativity, since Einstein self-consciously appealed to the traditional debate about absolute versus relative motion in order for his introduction of a new coordination to be taken seriously as an alternative to the classical coordination. Second, although the debate about the reality of absolute motion was characterized by disagreement at the philosophical level, there was a relatively stable consensus that thinkers such as Newton, Leibniz, Kant and Mach had articulated the main arguments, which any new contribution to the debate therefore had to treat with the utmost seriousness. Third, Friedman claims that the contemporary clarification of the status of absolute versus relative motion within Newtonian physics through the development of the concept of an inertial frame by physicists such as Carl Gottfried Neumann, Ludwig Lange, Lord William Thomson Kelvin and Peter Guthrie Tait corresponded to a "relatively uncontroversial and unproblematic scientific contribution to this debate."[36] Moreover, in formulating the new coordination, Einstein used the newly discovered empirical fact regarding the speed of light in a very surprising way. Friedman therefore concludes that the classical physicists had to take Einstein's proposal seriously on their own terms.[37]

[35] Friedman 2001, p. 107.
[36] Ibid., p. 107.
[37] Ibid., p. 108.

2.11 CONCLUSION

This chapter examined Michael Friedman's neo-Kantian conception of relativized a priori constitutive principles within Newtonian and relativistic physics, which he expounded in his book *Dynamics of Reason* (2001). According to this view, the frameworks of Newtonian physics, special relativity and general relativity each contain unique constitutive principles that empirically define their spatio-temporal structure. For example, Newton's three laws of motion define the inertial frames of Newtonian mechanics, which are necessary for proper empirical laws such as the universal law of gravitation to have any empirical meaning at all. Since the constitutive principles serve as presuppositions for the meaningfulness of the empirical laws, they are themselves non-empirical in nature. In order to explain the rationality behind the abandonment of one constitutive framework for another, Friedman develops a concept of prospective inter-paradigmatic rationality in view of their incommensurability. Einstein's introduction of the principle of light into the special theory of relativity serves as Friedman's main example. He notes that it is related to the development of the concept of inertial frame in the late nineteenth century and appealed to conceptual resources and problems that were already in place. Physicists working within the Newtonian framework were therefore able to understand and recognize the new coordination as a reasonable live option.

CHAPTER 3

Synthetic History Reconsidered

Since its publication in 2001, Friedman's *Dynamics of Reason* has been at the center of much debate. The most comprehensive discussion of Friedman's conception of the relativized constitutive a priori can be found in the collection *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science* (2010), which concludes with a long article by Friedman himself entitled "Synthetic History Reconsidered." As the title and the length of the article suggest (230 pages!), Friedman has used the publication of the compilation as a welcomed opportunity to reconsider his neo-Kantian interpretation of the foundations of the mathematical sciences from Newton to Einstein. The following review will focus primarily on the section entitled "Synthetic History and the Dynamics of Reason," where we find Friedman's detailed analysis of this history from Helmholtz's contribution to the foundations of geometry via Poincaré's hierarchical conception of the mathematical sciences to Einstein's creation of the special theory of relativity.

3.1 SYNTHETIC HISTORY AS TRANSCENDENTAL PHILOSOPHY

Friedman opens the section with the admission that the concept of coordination between mathematical formalism and empirical phenomena (which he adopted from the early Reichenbach) conveys a completely unsatisfactory idea of the application of mathematics in modern physics. For "it assumes an overly simplified 'formalistic' account of modern abstract mathematics, and, even worse, it portrays such abstract mathematics as being directly attached to intuitive perceptible experience at one fell swoop."[1] Friedman then goes on to approve Thomas Ryckman's main line of argument in *Reign of Relativity* (2005) by conceding that Edmund Husserl's phenomenological version of transcendental philosophy seems much better suited than early logical empiricism to do justice to the application of abstract mathematics to concrete empirical phenomena. At the beginning of his phenomenological approach, Husserl rejected a sharp distinction between conceptual thought and intuitive experience, without renouncing the characteristic Kantian project of a transcendental explanation of the possibility of objective knowledge. In contrast to Kant's appeal to the transcendental

[1] Friedman 2010, p. 698.

schematism of the understanding, Husserl found "his ultimate grounding in the transcendental phenomenological structure of the ordinary and always present life-world itself, as it is directly and immediately given to all of us."[2] The positive sciences then necessarily arise from inductive regularities within the life-world through a process of continual correction that attempts "continually to correct and refine such regularities by reference to more accurate and comprehensive scientific laws of nature."[3]

To better understand how the application of abstract mathematics to concrete experience is possible, Friedman argues, we therefore need a transcendental explanation of how mathematical theories of space, time and motion have been successively related to empirically given frames of reference embedded in the phenomenological life-world itself. In other words, such an explanation must capture the radical shifts that characterize the actual historical development of space-time physics from Newton to Einstein. However, since this development appears to be the result of a sequence of contingent historical circumstances, we are forced not only to abandon Kant's version of the synthetic a priori, but also to question the very idea of developing "a philosophical understanding of the evolution of modern science that is at once genuinely historical and properly transcendental."[4] In Friedman's reading, Husserl attempted to respond to this challenge by pointing to the necessary origin of the historical development of science in transcendental features of the life-world.

However, Friedman argues that even if we endorse Husserl's transcendental account of the life-world, this does not solve the problem. For this account tells "us nothing specific about any given stage in the development of the mathematical-physical science," and consequently "it would seem that it cannot transcendentally illuminate the historical process by which the empirical meaning of any given stage is actually constituted."[5] Instead, he points to his own approach to the dynamics of reason as a historical form of transcendental philosophy that attempts to meet precisely this challenge. Thus, while Friedman has no intention of abandoning the core of the dynamics of reason, i.e. the account of the role of relativized constitutive principles in the development of mathematical physics from Newton to Einstein, he concedes that he needs to strengthen the transcen-

[2] Ibid., p. 690.
[3] Ibid., p. 688.
[4] Ibid., p. 696.
[5] Ibid., p. 696.

dental significance of these principles in particular historical contexts.

Friedman begins this supplementary account with Kant's answer to the question of how mathematical physics is possible, noting that although Kant was operating against the contingent background of the intellectual resources of his time, the solution he found was by no means arbitrary:

> Geometry, for Kant, is limited to the classical system of Euclid; the pure understanding or pure intellect is delimited by the logical forms of Aristotle; the available conceptions of space and time are exhausted by the Leibnizian and Newtonian alternatives; and so on. Kant's construction of a synthetic description of our faculties of sensibility and understanding can only be understood against the background of precisely these resources – mathematical, logical, metaphysical, and theological – as Kant delicately navigates within them and eventually radically transforms them. The revolutionary and completely unexpected result, that space and time are pure forms of our (human) faculty of sensibility and that, considered independently of sensibility, our faculty of understanding yields no (theoretical) cognition at all, then emerges as the practically unique solution to the problem set by the existing intellectual resources: It is the only available conception of our rational faculties that does simultaneous justice to both Newtonian mathematical physics and Leibnizian (as opposed to Newtonian) natural theology and metaphysics. [...] [T]he intellectual situation in which he found himself had a definite "inner logic" – mathematical, logical, metaphysical, and theological – which allowed him to triangulate, as it were, on a practically unique (and in this sense necessary) solution.[6]

Friedman's next aim is to extend this understanding of Kant's transcendental philosophy to post-Kantian developments by tracing how the "inner logic" of successive intellectual situations evolved from Kant's original theory in response to new developments in the mathematical sciences and post-Kantian scientific philosophy. The existence of a certain "inner logic" for each successive intellectual situation is thus meant to imply that this so-called "synthetic history" can still count as transcendental philosophy, insofar as this "enterprise does not collapse into total contingency."[7] In the

6 Friedman 2010, pp. 701–2.
7 Ibid., p. 702.

next section we will see how Friedman applies his synthetic method to the new intellectual situation in the 18th century created by the development of non-Euclidean geometry.

3.2 HELMHOLTZ'S SOLUTION TO THE "SPACE PROBLEM"

At Kant's time, non-Euclidean geometries were not yet properly developed. In practice, therefore, there was no other possibility than to attribute Euclidean geometry to the structure of intuitive space. It was for this reason that this attribution was necessarily part of Kant's solution. However, the exceptional standing of Euclidean geometry changed with the development of hyperbolic and elliptic geometry in the course of the 18th century. To account for this, according to Friedman, Helmholtz developed "a conception of space representing 'the minimal such generalization that is consistent with the idea that our 'subjective' and '*necessary* form of external intuition' may be described by a non-Euclidean geometry."[8] In particular, Helmholtz's conception retained the transcendental trait of Kant's original within an empiricist theory of spatial perception. The fundamental idea was that our ability to locate objects in space is not an inherent capacity, but a gradually learned skill that is acquired "by a process of 'unconscious inductive inferences' based on regularities or associations among our sensations."[9] For example, the ability to shave oneself in front of a mirror is not innate, because in order to move the razor properly, we must have learned that mirror reflections are left-right reversed. The same regularities between our sensations and bodily motions on which our ability to locate objects in space depends, Friedman argues on Helmholtz's behalf, then imply that our representation of space itself must conform to a condition of free mobility: Our ability to localize rigid objects at different positions in space is based on the condition that these objects retain their spatial shape when displaced. On the basis of this principle, Helmholtz finally deduced that space must correspond to one of the geometries with constant curvature: hyperbolic, elliptic or Euclidean. In contrast to Kant's theory of space, the axioms of Euclidean geometry are therefore not built into the necessary conditions that underlie our spatial intuition, but "the specifically Euclidean character of physical space is [rather] a *merely* empirical fact about the actual behavior of our measuring instruments as we move them around in accordance

[8] Ibid., fn. 289.
[9] Ibid., p. 631.

with the condition of free mobility."[10] Friedman concludes, however, that while Helmholtz's conception is empiricist insofar as the determination of the curvature of space is a matter of empirical investigation, it is also distinctively Kantian insofar as "space has a '*necessary* form' expressed in the condition of free mobility."[11]

3.3 POINCARÉ ON SPACE AND RELATIVITY

Friedman continues his transcendental history of the foundations of the mathematical sciences from Newton to Einstein by evaluating Poincaré's contribution to the latter:

> Poincaré's conception is the minimal extension of Helmholtz's consistent with the more sophisticated group-theoretic version of the principle of free mobility due to Lie [. . .], the new perspective on the relativity of motion due to the modern concept of an inertial frame [. . .], and, most importantly, the apparently paradoxical new situation in electrodynamics arising in connection with precisely this relativity of motion.[12]

Helmholtz's theorem on the principle of free mobility was later proved more rigorously by the Swedish mathematician Sophus Lie in the context of group theory.[13] From Friedman's point of view, it was Poincaré who incorporated this new development into the representation of physical space as part of his overall solution to the problem of the relativity of motion within a hierarchical framework of the mathematical sciences.

The fundamental feature of Poincaré's hierarchy is that each science in it presupposes the sciences at all higher levels. Beginning with arithmetic at the top of the hierarchy, the continuum of real numbers follows at the next lower level, which in turn is succeeded by the concept of space. As already indicated, Poincaré interpreted the latter as a three-dimensional mathematical continuum arising from our sensory experience of bodily motion, refining Helmholtz's solution to the so-called "space problem" by utilizing Lie's group-theoretic account of free mobility:

---

[10] Friedman 2010, pp. 644–5.
[11] Ibid., p. 705.
[12] Ibid., fn. 298.
[13] Ibid., fn. 141.

> In order to apply mathematics to this intuitive experience [of bodily motion] we then need to form a corresponding rigorous concept, and this, for Poincaré is the concept of a group – in accordance with which, as an idealization, we assume that our bodily displacements are closed under the operation of composition, that this operation always has a unique well-defined inverse, and so on. Moreover, we also assume, as an additional idealization, that we thereby arrive at a continuous group in the sense of Lie [...] It now follows, by the Helmholtz-Lie theorem, that space, in this sense, has a (metrical) geometry [of constant curvature], and so, for Poincaré, we now have the beginnings of an explanation for how we are able to apply the exact science of geometry to our sensory experience.[14]

The original conclusion of Poincaré's answer to the space problem thus contradicted both Kant's a priorism and Helmholtz's empiricism. The geometric structure of physical space is neither a necessary nor an empirically determined form, but rather a free convention of our mind that we impose on our sensory experience of bodily displacements. First, the group-theoretic idealization of "our rough and approximate sensory experience"[15] is not imposed on us, but is freely chosen. Second, the measurements facilitated by such an idealization cannot be made with an accuracy sufficient to determine the curvature of space. Instead, we choose to utilize Euclidean geometry for our measurements because it is convenient and simple compared to all other alternatives. Therefore, Friedman emphasizes that Poincaré, by recognizing geometries with variable curvature as possible but inconvenient alternatives, did not agree with Helmholtz's understanding of free mobility as a necessary condition for spatial measurements. However, our experience with bodily motion has guided and influenced our decision to adopt Euclidean geometry as the most convenient way for spatial measurement, so that Poincaré's conventionalism does not turn into geometric nominalism either.

At the next level down in the hierarchy, Poincaré placed classical or Newtonian mechanics. In particular, he argued that this science presupposes that spatial geometry is already in place as a standard for the force-free behavior of moving bodies. More generally, he argued that by "elevating" the merely empirical fact that spatial measurements roughly meet

[14] Ibid., p. 644.
[15] Ibid., p. 645.

the rules of Euclidean geometry to the status of a convention, we are able to impose "a precise mathematical framework" on our sensory experience, "within which alone our properly physical theories can subsequently be formulated."[16] According to Friedman, the essential question discussed in Poincaré's treatment of mechanics concerns the problem of absolute space and the relativity of motion. The French mathematician initiated this discussion with the so-called "law of relativity," which states that "[t]here is no absolute space and we can conceive nothing but relative motions."[17] From Friedman's point of view, this law expresses the conclusion of a complicated argument based on "one of the most striking and important applications of Poincaré's hierarchical conception [of science]."[18] The main argument is as follows:

> Poincaré's key idea is that what he calls the (physical) "law of relativity" rests squarely on the "relativity and passivity of space" and therefore reflects the circumstance, essential to free mobility and the Helmholtz-Lie theorem, that the space constructed from our experience of bodily displacements is both homogenous and isotropic: all points in space, and all directions through any given point, are, necessarily, geometrical equivalent.[19]

Friedman begins his careful presentation of this argument by setting out the law of relativity in its definitive version:

> [A]bsolute position or orientation of a system of bodies in space can have no physical effect whatsoever, and neither can any *change* (velocity) of such absolute position or orientation.[20]

Next, he interprets the principle of "relativity and passivity of space" as the condition that space must be both homogeneous and isotropic. As we have seen, this condition is already a direct consequence of Poincaré's group-theoretic analysis of free mobility insofar as it restricts spatial geometry to a constant curvature. Since the homogeneity and isotropy of space imply (by their very definition) that neither absolute position nor absolute rotation can have any physical effect, the first part of the law of relativity is now

[16] Friedman 2010, p. 708.
[17] Poincaré 1952, p. 90.
[18] Friedman 2010, p. 705.
[19] Ibid., p. 705.
[20] Ibid., p. 705.

satisfied. The second part, Friedman argues, then easily follows from the first:

> But it then follows, similarly, that the velocity with which such a supposed absolute position or orientation can have no physical effect either, and, accordingly, the laws of the phenomena must be entirely independent of all such absolute velocities (of translation or rotation) as well.[21]

He therefore comes to the conclusion that Poincaré built the law of relativity on his philosophy of geometry.

But precisely for this reason, Friedman continues, Poincaré's discussion of the relativity of motion now encounters serious problems, since "velocities of absolute *rotation* do appear to have physical effects."[22] Poincaré himself was of course painstakingly aware of this dilemma. In chapter VII of *La Science et l'Hypothèse*, he introduced the principle of relative motion stating that "[t]he movement of any system ought to obey the same laws whatever it is referred to fixed axes or to the movable axes which are implied in uniform motion in a straight line."[23] In contrast to the law of relativity, the principle of relative motion does not apply to rotational motion and thus, as Friedman emphasizes, corresponds to what we today call Galilean relativity. Poincaré then noted that experiments, although they should satisfy the law of relativity, seem to be limited to the principle of relative motion. In particular, he mentioned the flattening at the poles and Foucault's pendulum experiment as apparent evidence for the true absolute rotation of the earth. To resolve this difficulty, Poincaré devoted the following discussion to the case of the uniformly rotating earth. He conceded that the attribution of a true absolute rotation is the simplest explanation. However, there are also more complicated explanations involving Coriolis and centrifugal forces. From Poincaré's point of view, the statement that "the earth really rotates" is therefore not a fact, but a free decision of our mind, which we base on its simplicity and convenience.

At the lowest level of the hierarchy, we finally come to the proper empirical sciences such as the theories of gravitation and electrodynamics. According to Friedman, the main problem with the state of the latter was the

[21] Ibid., p. 649.
[22] Ibid., p. 649.
[23] Poincaré 1952, p. 111.

lasting impossibility to detect the absolute velocity of the earth with respect to the ether by drift experiments such as the Michelson-Morley interference experiments of 1881 and 1887. Lorentz had continuously contributed to the resolution of this problem in the years 1886-1904 by first developing a theory that explained why such experiments were unsuccessful up to first order in v/c, where v denoted the expected absolute velocity of the earth. To account for the null result of the Michelson-Morley experiments, he extended his original explanation to the second order by appealing to length contraction. According to Friedman's reading, in 1904 he finally "erected a theory in which no experiments of any order would succeed, and he showed that the further hypothesis of local time (what we now call time-dilation) accomplishes this."[24]

Although he made his most important contributions to electrodynamics in general and to the null result of the ether drift experiments in particular after 1902, Friedman argues, already in *La Science et l'Hypothèse* "Poincaré had formulated his 'law of relativity,' and accordingly he had already expressed considerable dissatisfaction with Lorentz's ongoing piece-meal attempts to explain the negative results of the aether experiments."[25] In particular, he demanded in 1902 that an adequate theory of electrodynamics should strictly realize what Einstein would shortly thereafter call the "principle of relativity." Nonetheless, Poincaré's implementation of this principle in the electrodynamics that he formulated in his 1905-6 papers "Sur la Dynamique de l'Électron," Friedman continues, was profoundly influenced by the "Lorentzian tradition," insofar as the French mathematician still used "all the special Lorentzian mechanisms (local time, the contraction hypothesis – including the deformable electron, and so on)."[26] It follows that Poincare's understanding of the principle of relativity was incompatible with Einstein's:

> This principle, for Poincaré, was never a mere empirical law, even one which, in accordance with his general methodology, could later be "elevated" to the status of a convention [...] Rather, the idea of the relativity of motion is equally based, for Poincaré, on an a priori "law of relativity" which itself [...] is not so much a physical as a purely geometrical principle. [...] Poincaré's understanding

[24] Friedman 2010, p. 651.
[25] Ibid., p. 652.
[26] Ibid., p. 655.

> of the principle of relativity is therefore inextricably connected with his hierarchical conception of how the mathematical and empirical sciences are related. In particular, "the physical sciences properly so-called" presuppose that all higher levels of the hierarchy are already in place, and the most important such (empirical) science, for Poincaré, is electrodynamics. [...] This science [...] presupposes that a theory of mechanics is already in place – which, for Poincaré, is of course Newtonian.[27]

In other words, due to his hierarchical concept of science, Poincaré was forced to implement the principle of relativity at every level of his hierarchy. The homogeneity and isotropy of space guaranteed its implementation at the level of geometry. Next, Newton's three laws of motion defined the kinematical structure of an inertial frame at the level of mechanics and accordingly implemented the principle of relativity in its Galilean version. Finally, Poincaré had no choice but to implement it in electrodynamics, using the "Lorentzian mechanisms" mentioned above.

3.4 EINSTEIN'S SOLUTION TO THE RELATIVITY OF MOTION

In his 1905 paper "Zur Elektrodynamik bewegter Körper," Friedman argues that Einstein "totally ignores the hierarchical conception of the mathematical and physical sciences Poincaré had carefully constructed"[28] by implementing the principle of relativity directly at the level of kinematics. In particular, Einstein based the relativity of motion not on a priori motivations arising from the philosophy of geometry, but on the mounting empirical evidence suggesting that electrodynamic phenomena do not delimit a distinguished (ether) frame. This strategy enabled him to view Lorentz's mechanism as a consequence and not a prerequisite of his theory:

> [F]rom what we now call the relativity of simultaneity it follows that the length of a rigid rod varies in different inertial frames and so does the temporal duration measured by an ideal clock. Length-contraction and time-dilation are thus direct, purely kinematical consequences of Einstein's new definition of simultaneity, and no special hypotheses involving the assumed electronic constitution of matter

[27] Ibid., pp. 655–6.
[28] Ibid., p. 657.

> (including the hypothesis of the deformable electron) are needed at all.[29]

Einstein's definition of simultaneity stipulates that light needs the same amount of time to travel from A to B as it does to return from B to A. He justified this with the invariance of the speed of light in all inertial frames, which in turn was due to the application of the principle of relativity to the light postulate (which states that light has a constant speed c, regardless of the speed of the source). As already indicated, these two principles arose from the null result of all ether drift experiments through a process of "elevation" by which it was assumed that the experimental validity up to the second order in v/c would apply rigorously to all orders. The elevation of the two principles then allowed him to define the revolutionary kinematical structure of an inertial frame in special relativity "that suffices to solve the current problems of electrodynamics independently of all more specific hypotheses."[30] Thus, while Einstein disregarded Poincaré's hierarchy of science, he fundamentally revised the kinematical framework by adopting the Frenchman's methodology of elevation quite rigorously. In Friedman's reading, then, there is a direct line of development from Kant's original version of transcendental philosophy to Einstein's solution to the "apparently paradoxical situation in electrodynamics" at the beginning of the twentieth century:

> A central contention of Kant's original version of transcendental philosophy, as we know, is that the three Newtonian Laws of Motion are not mere empirical laws but a priori constitutive principles on the basis of which alone the Newtonian concepts of space, time, and motion can then have empirical application and meaning. What we have just seen is that Einstein's two fundamental "presuppositions" or "postulates" play a precisely parallel role in the context of special relativity. [...] But we have also seen significantly more. [...] Whereas Helmholtz's principle of free mobility generalized and extended Kant's original theory of geometrical construction within our "subjective" and "*necessary* form of external intuition," Poincaré's idea that specifically Euclidean geometry is then imposed on this form by a "convention or definition in disguise" represents an extension or continuation of Helmholtz's conception. For Poincaré, specifically Euclidean

29 Friedman 2010, p. 655.
30 Ibid., p. 657.

> geometry is applied to our experience by [...] a process of "elevation" [...] This same process of "elevation," in Einstein's hands, then makes it clear how an extension or continuation of Kant's original conception can also accommodate new and surprising facts – in this case, the very surprising empirical discovery [...] that light has the same constant velocity in every inertial frame.[31]

Although Helmholtz and Poincaré extended the scope of possible spatial geometries to non-Euclidean spaces with constant curvature, both ultimately favored Euclidean geometry for empirical and conventional reasons, respectively. As we have seen, Poincaré went on to presuppose the classical Newtonian framework at the level of mechanics. In his effort to explain "why specifically electrodynamic phenomena do not reveal the true absolute velocity of the earth" at the level of proper empirical sciences, he was therefore unable to implement the principle of relativity without invoking "additional special mechanisms."[32] It was thus left to Einstein to impose a non-Euclidean framework on "our rough and approximate perceptual experience" by utilizing the two elevated laws as constitutive principles. Accordingly, from Friedman's point of view, we should regard the constitution of the kinematical structure of special relativity as the very first implementation of the relativized a priori:

> Einstein's creation of special relativity, from this point of view, represents the very first instantiation of a relativized and dynamical conception of the a priori – which, in virtue of precisely its historical origins, has a legitimate claim to be considered as genuinely constitutive in the transcendental sense.[33]

Friedman's historicized version of transcendental philosophy is said to have shown not only how successive intellectual situations developed against the background of Kant's original theory in general and the constitutive role of Newton's laws of motion in particular, but also that each such situation consequently had its own "inner logic" to which its solution necessarily had to conform. For Friedman, it follows in particular that we should regard the principle of relativity and the light principle as "genuinely constitutive in the transcendental sense."

[31] Ibid., p. 708.
[32] Ibid., p. 655.
[33] Ibid., p. 708.

3.5 THE TRANSCENDENTAL NECESSITY OF CONSTITUTIVE PRINCIPLES

Discussing Robert DiSalle's book *Understanding Space-Time* (2006) and contribution to the collection entitled "Synthesis, the Synthetic A Priori, and the Origins of Modern Space-Time Theory," Friedman returns to the question of what he understands by the "transcendental necessity of the principles."[34] According to DiSalle's interpretation, Friedman explains, "critical conceptual analysis [. . .] results in new empirical definitions of fundamental physical concepts (true rotational motion, simultaneity, inertia) which, in an important sense, are *uniquely* determined in the context in question."[35] This insight then leads DiSalle both to a devastating critique of the logical empiricists' notion of arbitrary coordination and to a conception of transcendental necessity:

> If a certain way of defining a concept is shown to be a condition of the possibility of employing that concept at all, [. . .] then it can hardly be seen as an arbitrary coordination. Nor, therefore, can the argument for it be seen as an a priori appeal to convenience or simplicity of the framework that the definition constitutes. The argument, rather, reveals the new conception in its transcendental role, as uniquely making possible the synthesis of experience under formal concepts.[36]

In his commentary on this passage, Friedman agrees with DiSalle insofar as he now also rejects the logical empiricists' notion of arbitrary coordination in favor of a "truly transcendental constitutive a priori," and also argues for "the transcendental necessity of the principles in question – *relative* to a given intellectual situation."[37] Friedman also concedes, however, that DiSalle's notion of transcendental necessity perhaps delimits the relevant kind of uniqueness somewhat too strictly. Particularly in the period 1905-12, before the advent of general relativity, Friedman acknowledges that Poincaré had every right to reject the new geometrical-mechanical framework of special relativity in favor of the classical Newtonian framework. From the perspective of Minkowski's relativistic space-time, "the central problem of Poincaré's hierarchy is that it makes the three-dimensional geometry of space prior to the four-dimensional geometry of space-time."[38]

---

[34] Friedman 2010, p. 727.
[35] Ibid., p. 727.
[36] DiSalle 2010, p. 529.
[37] Friedman 2010, p. 727.
[38] Ibid., fn. 296.

But despite this fact, Poincaré did not only resolve the null result of ether drift experiments within the classical geometrical-mechanical framework. His ether-based theory of electrodynamics was in fact "both empirically and mathematically equivalent to Einstein's."[39] Hence, Friedman emphasizes, "I would claim only that Einstein's new geometrical-mechanical framework is the unique *non-classical* alternative (unique minimal *extension* of the classical framework) appropriate to the intellectual situation in which he found himself."[40] Poincaré died in 1912 and accordingly did not live to comment upon the creation of the general theory of relativity and its implementation of spatial geometries with variable curvature in the years 1912-5. In contrast to the kinematical structure of special relativity, which survived the transition to general relativity as local frames, for Friedman, "[w]hat Poincaré called the 'relativity and passivity of space' is now definitely a thing of the past, together with Poincaré's own philosophical and mathematical motivations for what he called the principle of relativity."[41]

3.6 CONCLUSION

To summarize, in his 2010 recapitulation of the dynamics of reason, Friedman adheres to the very core of his 2001 conception of the relativized constitutive a priori. In particular, he still interprets relativized constitutive a priori principles as the unique conditions that make the empirical application of a particular mathematical-physical theory possible. However, in order to distinguish his conception more clearly from the logical positivist picture of arbitrary coordination between conceptual thought and sensory experience, Friedman feels compelled to extend the core of his account to include a historicized version of transcendental philosophy. "[W]hat makes this enterprise properly 'transcendental,'" he insists, is "that the 'inner logic' of the successive intellectual situations in question proceeds against the background of, and explicitly in light of, Kant's original theory."[42]

[39] Ibid., p. 658.
[40] Ibid., fn. 353.
[41] Ibid., p. 660.
[42] Ibid., fn. 300.

Part II

THE DEVELOPMENT OF SPECIAL RELATIVITY RECONSIDERED

CHAPTER 4

Reconsidering Friedman and the Dynamics of Reason

In this part, I will prepare my subsequent evaluation of Friedman's concept of the relativized a priori by reconsidering the transition from the classical framework of Newtonian mechanics to special relativity from a historical perspective. Since Marc Lange's review of *Dynamics of Reason* (2004) has shaped my understanding of the need for such a reconsideration, we begin this chapter with a brief examination of Lange's analysis. In my view, Friedman's work sheds considerable light on the transition from Newtonian to relativistic physics. However, as we shall see, Lange points out crucial ambiguities and possible historical inadequacies in Friedman's account. In particular, Friedman does not argue for his claim that Newton's laws of motion comprise the only possible set of constitutive principles for Newtonian physics. Unfortunately, Lange himself does not formulate a convincing argument for his position that alternative constitutions are possible. This is more than understandable, considering that his contribution is a review of Friedman's work. I have therefore continued where Lange left off by tracing in detail the historical development from Lorentz's ether-based theory of electrodynamics to Einstein's 1905 relativity paper. To facilitate the reading of the latter, I conclude this chapter with a brief overview of the conclusions that can be drawn from it.

4.1 LANGE'S CRITIQUE OF FRIEDMAN

One of the first things Lange notes is that "Friedman's point [. . .] that *every* empirical prediction made by Newton's law of gravitation requires Newton's laws of motion"[1] does not seem to be conclusive, for Friedman "has not argued that the *particular* 'coordinating principles' of Newtonian mechanics are necessary for Newton's law of gravitation to have any bearing on physical reality."[2] Rather, the latter has shown "that without *something* to play the role of Newton's laws of motion, Newton's law of gravitation is irrelevant to physical reality."[3] I addressed this issue above in my discussion of Friedman's account of the necessity of constitutive principles. I argued that Friedman's position seems to imply that only the Newtonian laws of

[1] Lange 2004, p. 704.
[2] Ibid., p. 704.
[3] Ibid., p. 704.

motion allow us to empirically define the spatio-temporal domain of Newtonian physics. From this perspective, Lange's critique can be framed as saying that it may well be that some principles are necessary to empirically define the classical space-time structure, but Friedman has not given us any reasons why this is only possible using the Newtonian laws of motion. One way to respond to Lange's critique would be to argue that Friedman was really only concerned with drawing attention to the need to empirically define a particular spatio-temporal framework by means of one or another set of constitutive principles. Lange concedes that such an interpretation may not be entirely wrong, as Friedman is quite ambiguous on this issue. For example, he remarks that the standard procedure for testing a particular empirical law "could not even be set up in the first place without some or another coordinating principle already in place."[4] But even if we grant Friedman this much, according to Lange he is still unable to give a satisfactory explanation for the transition of Maxwellian electrodynamics from Newtonian to relativistic physics. This is because Friedman must interpret Maxwell's equations in such a way that they change their meaning during such a transition:

> Unlike Newtonian gravitational theory, however, Maxwellian electromagnetic theory survived intact the shift from Newtonian to relativistic physics [...] Maxwellian electromagnetic theory could be interpreted as describing the electric and magnetic fields as they appear in a reference frame at absolute rest (or at least relative to the ether). Alternatively, it could be interpreted as describing the electric and magnetic fields as they appear in any inertial frame. [...] However, since Friedman regards the 'coordinating principles' as constitutively *a priori* relative to Maxwellian theory, he could not say that here we have one and the same theory coupled with either classical or relativistic mechanics. Rather, in Friedman's view, there has been a change in *meanings* of various terms in Maxwell's [...] equations (cf. pp. 98-99).[5]

First of all, Lange points out that, with respect to the classical spatio-temporal framework, the Maxwellian theory can only be true in one reference frame.[6] This is related to the fact that the Maxwellian equations are

[4] Friedman 2001, p. 81.
[5] Lange 2004, p. 705.
[6] This frame of reference is then interpreted as being at rest relative to the ether.

invariant under Lorentz transformations instead of Galilean transformations. However, in the context of special relativity, the Maxwellian equations can be interpreted as being valid in any inertial frame, which is due to the fact that any change of coordinates from one inertial frame to another is now determined by a Lorentz transformation.[7] Lange finally argues that since Friedman claims that proper empirical laws only have empirical content or truth value relative to a spatio-temporal framework, a change in the framework of an empirical law must imply that at least some of its terms change meaning.

The reason Lange disagrees with Friedman on this point is that the conceptual foundations of Maxwellian electromagnetic theory "remained somewhat fluid and unsettled until the advent of relativity" and it therefore may "make less sense to speak of Maxwellian theory as undergoing 'meaning change' than to speak of the theory as itself highly ambiguous (albeit ultimately in a fruitful way)."[8] This ambiguity implies in particular that the Maxwellian theory does not presuppose any particular set of principles that empirically define frames of reference in which the theory then applies, but only some set of such principles. For Lange, this means that the claim that the theory did not survive the transition from Newtonian to relativistic physics unscathed is highly dubious – especially because Maxwellian electromagnetic theory played a pivotal role in the introduction of the new relativistic framework. Lange attempts to bolster his argument by noting that a similar point can be made with respect to Friedman's explanation of the constitutive function of Newton's laws of motion relative to Newton's law of gravitation. In contrast to our modern understanding of Newton's laws of motion as implicitly defining "inertial frames," Lange notes that even Friedman himself repeatedly states that Newton understood them differently. We can add that Newton understood his laws of motion, to which Lange alludes, as straightforward empirical laws defined relative to absolute space. This shows, according to Lange, that it is implausible to claim that Newton's law of gravitation has no empirical content if no frame of reference is specified:

> Even without any privileged reference frames being picked out, Newton's law of gravitation is not entirely empty of content regarding

[7] This is of course not surprising since Einstein self-consciously developed the special theory of relativity to work that way.

[8] Lange 2004, p. 705.

> physical reality; it might even be taken as asserting that there is *some* set of principles for picking out reference frames under which it applies with empirical adequacy. Friedman might accept all this and insist merely that without some set of principles, the law of gravitation is empirically useless. But again, without some such set of force laws [...], Newton's laws of motion are empirically useless. Insofar as all this goes, the situation seems roughly symmetrical.[9]

Thus Lange points out that while Friedman may agree that Newton's law of gravitation presupposes only one or another set of constitutive principles, he nevertheless maintains the distinction between principles and empirical laws by claiming that properly empirical laws are empirically useless until a particular set of such principles has been specified. Against this, Lange argues that we can also regard the principles themselves as empirically useless if there are no force laws. Lange does not elaborate on his criticism, but he probably has something like the following in mind: without knowledge of empirical force laws, we will never be able to recognize which frames of reference are actually inertial frames. For if we do not know what kind of forces exist, then we cannot be sure whether we have correctly recognized all the forces in a reference frame. And as a result, we also cannot know whether all forces in a reference frame occur in pairs. In other words, we are not able to recognize whether a reference frame satisfies all three Newtonian laws of motion.

After laying out his arguments for the symmetry between Friedman's constitutive principles and empirical laws, Lange concludes his review with an interpretation of the philosophical implications of the Michelson-Morley experiments. He begins his explanation by questioning Friedman's claim that although "the undetectability of differences in inertial motion [...] provides us with strong empirical motivation, not only for entertaining a new coordination , but also [...] for doubting the adequacy of the classical coordination,"[10] the choice between the two constitutive frameworks "is not happily viewed, in any sense, as a case of empirical testing."[11] In particular, Lange poses the rhetorical question of why Einstein's proposal cannot be tested by the usual empirical means, even though it can be empirically motivated.

---

9 Lange 2004, p. 707.
10 Friedman 2001, p. 88.
11 Ibid., p. 89.

As we have seen, Friedman answers this question by arguing that the Michelson-Morley experiments cannot be regarded as a genuine empirical test of Einstein's new coordination simply because Lorentz successfully formulated an electromagnetic theory of moving bodies that embeds the null result in a classical spatio-temporal framework. Although Lange concedes that the experiments do not constitute evidence for special relativity over Lorentz's theory, he claims that the fact that "a classical spatio-temporal structure can accommodate these results only by way of some maneuver along the Lorentz-Fitzgerald lines may surely qualify as some evidence against that structure."[12] According to Lange, a very good explanation of why the Lorentz-Fitzgerald maneuver is so unsatisfactory was formulated by Einstein himself, who questioned whether nature had placed us in an ether gale only to design the laws of nature in such a way that we could not detect the gale. In other words, a classical spatio-temporal structure can only accommodate the null result as "an unlikely coincidence in the laws of nature, a remarkable case of 'fine-tuning'" that "cries out to be eliminated – as Einstein does."[13] So even without the existence of an alternative, we have ample empirical reasons to be dissatisfied with Lorentz's theory. Lange is of course aware that Friedman will quickly retort that this critique does not distinguish between first-level and second-level scientific theories, to which a distinction between first-level and second-level empirical reasons corresponds. As we have seen, only the former are genuine empirical evidence, while the latter can nevertheless help to motivate a transition between constitutive principles. Friedman introduces these distinctions, as Lange goes on to note, "on the grounds that [...] a first-level theory's empirical motivation goes straightforwardly through second-level theory, whereas the second-level theory's empirical motivation proceeds only by the grace of a third-level philosophical meta-framework and therefore is not happily viewed as a case of empirical testing."[14] Friedman's sharp distinction between constitutive principles and empirical laws thus presupposes a very sharp distinction between the scientific-constitutive level and the philosophical meta-level. But, Lange now argues, it is precisely this latter distinction that simply cannot be drawn sharply. To support his argument, he points out that Einstein's argument against Lorentz is based on

[12] Lange 2004, p. 709.
[13] Ibid., p. 709.
[14] Ibid., p. 710.

considerations of unity and parsimony, which, while not self-explanatory, are precisely the sort of considerations that – even according to Friedman – also determine the choice of theory within a single constitutive framework. Lange therefore concludes that Einstein's argument seems to provide only standard empirical grounds for the new framework, which means that a sharp distinction between genuine empirical evidence on the one hand and distinct philosophical empirical motivations on the other does not seem to be warranted.[15]

4.2 THE DEVELOPMENT OF SPECIAL RELATIVITY

The remaining chapters of part 2 aim to answer some critical questions left open by Lange's review by analyzing in detail the historical development from Lorentz's ether-based theory of electrodynamics to Einstein's 1905 relativity paper. More specifically, this analysis reveals two key points: First, in order to reconcile electrodynamics with the null result of ether drift experiments, both Lorentz and Poincaré eventually abandoned Newton's second and third laws of motion. Consequently, from 1904 onwards, they did no longer utilize Newton's laws of motion as constitutive principles in Friedman's sense. For this reason, neither theory can be regarded as a classical interpretation of electrodynamics in Friedman's sense. Secondly, in 1910 Vladimir Ignatowki published an alternative derivation of the special theory of relativity which, in contrast to Friedman's concept of necessity in the sense of uniqueness, used a different definition of inertial frames than Einstein's 1905 paper. In addition to these two main points, part 2 contains the following list of historical shortcomings in Friedman's interpretation: Lorentz's 1904 theory of corresponding states did not fulfill the principle of relativity. Consequently, it was not empirically equivalent to Einstein's. In particular, Lorentz had no physical interpretation of local time (in terms of dilated clocks or otherwise). The "inner logic" of Lorentz's solution to the null results therefore differs radically from Friedman's account. Before 1905, for example, Lorentz did not interpret the concept of local time as indicating the reading of clocks moving relative to the ether. Instead, local time was a purely mathematical construct in the so-called theory of corresponding states, which allowed Lorentz to compare the results of an experiment performed both at rest and in motion relative to the ether.

This would explain why Friedman shifted his attention from Lorentz

[15] Lange 2004, p. 711.

to Poincaré from 2001 to 2010. In agreement with Friedman's interpretation, the Frenchman provided a Lagrangian formulation of electrodynamics in 1905 that was empirically equivalent to that of Einstein. However, as already indicated, the Lagrangian did not imply Newtonian mechanics, but special relativistic mechanics. In contrast to Friedman's interpretation, Poincaré was thus prepared to reject Newtonian mechanics in 1905 despite the hierarchy of science in *La Science et l'Hypothèse* (1902). To understand this development, we must consider Poincaré's 1900 lecture in honor of Lorentz's doctoral anniversary. In this lecture, Poincaré succeeded in reconciling Newtonian mechanics with the null result of ether drift experiments to a certain degree of accuracy. This resolution involved both 1) the extension of the principle of relativity to ordinary matter and the ether and 2) the first physical interpretation of Lorentz's concept of local time in terms of the apparent readings of clocks in uniform rectilinear motion relative to the ether. The latter resulted from the application of a method for synchronizing clocks that Poincaré had published in 1898.[16] However, although there was no experimental evidence against this reconciliation, Poincaré emphasized his conviction that the principle of relativity must apply to matter alone. He therefore predicted that Newtonian mechanics would have to be replaced, regardless of the revolutionary nature of such a step. Poincaré's 1905 Lagrangian formulation of electrodynamics was the result of his efforts.[17] The strategy of his theory, as the Frenchman emphasized in his book *Science et Méthode* (1908), was to reverse the logic of Lorentz's approach. Stated otherwise, contrary to Friedman's assertion, Poincaré did not introduce Lorentz's mechanisms, such as the contraction of the electron, as primary hypotheses to prove the principle of relativity. On the contrary, by assuming that (a Lagrangian formulation of) Lorentz's 1904 ether-based version of Maxwell's equations has to satisfy the principle of relativity (for matter only), Poincaré deduced how space and time coordinates must transform when a laboratory frame is boosted from rest to uniform rectilinear motion relative to the ether. Instead of assuming Lorentz's controversial mechanisms, as Poincaré emphasized in 1908, this strategy allowed him to derive them as consequences of a hypothesis that nobody had been able to challenge: the principle of relativity. Although Poincaré himself was the originator of these transformations, he called them Lorentz

[16] Poincaré 1898.
[17] Poincaré 1905b; Poincaré 1906.

transformations. In line with the reverse strategy, Poincaré did not elaborate on the procedures of space and time measurements in his 1905 and 1906 papers. Nevertheless, his Sorbonne lectures of 1906, among others, show that he still adhered to the distinction between real and apparent measurements, using his 1898 procedure for the synchronization of clocks. While this analysis showed why space and time measurements had to be transformed according to the Lorentz transformations, it did not explain how the dilation of clocks was possible. As far as I know, this situation remained unchanged until Poincaré's death in 1912. As the above illustrates, Friedman's interpretation of Poincaré's work up to 1902 would fit well if one disregards the fact that Poincaré's 1900 theory of electrodynamics did not yet correspond to the principle of relativity (for matter only). The problem is that Poincaré's 1905 theory of electrodynamics (published in 1905 and 1906), while finally satisfying the principle of relativity, implied that Newtonian mechanics had to be sacrificed along the way.

4.3 CONCLUSION

This chapter begins with an examination of Lange's review of *Dynamics of Reason* (2005), in which he points out that Friedman has not provided a convincing argument for the uniqueness of the constitutive principles of a particular geometrical-mechanical framework. For example, certain principles might be necessary to define the Newtonian framework empirically. However, Friedman has not given any reasons to suggest that this is only possible by utilizing Newton's laws of motion. In addition, the chapter prepares my later assessment of Friedman's concept of the relativized a priori. I do this by pointing out historical inadequacies in Friedman's account, which are documented in the remaining chapters of part 2. First, both Lorentz and Poincaré abandoned Newton's second and third laws of motion before the end of 1905 in order to reconcile electrodynamics with the null result of ether drift experiments. Secondly, Ignatowki published a derivation of the Lorentz transformations in 1910, which defined inertial frames in a different way to Einstein's 1905 paper.

Lorentz's 1895 Theory of Corresponding States

According to the wave theory of light founded by Augustin Fresnel in the early 19th century, light waves corresponded to transversal waves in a medium referred to as the *ether*. In his treatise of 1873, James Clerk Maxwell incorporated the assumption of the ether into his theory of electricity by identifying it as the carrier of the electromagnetic action. In particular, he argued that light corresponded to electromagnetic disturbances which, according to his theory, propagated in the ether (understood as the medium filling the "vacuum") at a constant speed c.[1] But unlike Fresnel, who assumed that the ether did not participate in the motion of the earth (to explain the aberration of stars), Maxwell followed the competing theory of George Gabriel Stokes, according to which the ether was dragged along by the earth in such a way that everywhere on the earth's surface the velocity of the ether was equal to that of the earth.[2] More precisely, Stokes understood the motion of the ether around the earth as an irrotational motion of an incompressible fluid around a spherical object.

Lorentz himself intervened in this debate by rejecting Stokes' theory because it "requires the existence of a velocity potential for the motion of the ether, which is incompatible with the equality of the velocities of the earth and the adjacent ether."[3] In other words, the irrotational motion of the ether around the earth had to imply a finite slip on its surface, as opposed to the alleged drag.[4] Lorentz went on to say that Fresnel's assumption of a stationary ether required that "the motion of the earth must influence the time required by light to travel to and fro between two points rigidly fixed to the earth."[5] He noted, however, that under this assumption a "serious difficulty [. . .] had arisen in an interference experiment made by Michelson."[6] Let us now turn to the description of this experiment in order to analyze the problem that its result posed for the reconciliation of physics at the time.

1 Torretti 1983, p. 37.
2 Darrigol 2006, p. 7.
3 Lorentz 1937b, p. 219.
4 Darrigol 2006, p. 8.
5 Lorentz 1937b, pp. 219–20.
6 Ibid., p. 219.

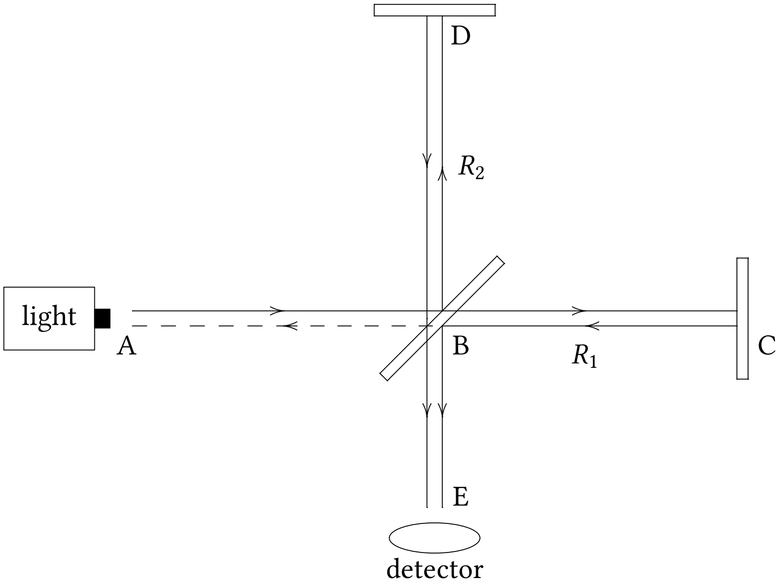


Fig. 5.1: The Michelson experiment of 1881.

5.1 THE MICHELSON-MORLEY EXPERIMENTS

The setup for Michelson's experiment consisted of an interferometer made up of two perpendicular arms *BC* and *BD*, both of length $l = l_1 = l_2$. At *A* there was a light source that emitted a beam that ran along *AC*. A half-silvered mirror, which formed an angle of 45° with *AC*, was placed at *B*, splitting the light beam into the rays R_1 and R_2. Beam R_1 was sent along *BC*, reflected in *C* and returned to *B* where it was (partially) reflected and continued along *BE*. The beam R_2 traveled along *BD*, was reflected in *D* and returned to *B*. At *B* it was (partially) transmitted and continued along *BE*, interfering with (the reflected part of) R_1. The resulting beam was then detected at *E*.

Let us now assume that the interferometer was placed so that *AC* lied in the direction of the earth's motion through the ether. Applying the Galilean rule for the addition of velocities and Fresnel's assumption of the independence of the speed of light from the relative motion of its emitter through the ether, the speed of R_1 should be $c - u$ between *B* and *C* and $c + u$ between *C* and *B* in the rest frame of the interferometer, where $\mathbf{u}$ was the relative velocity between the earth and the ether.[7] The time $\bar{t}_1$ measured in the laboratory frame, which R_1 should require for the journey from B to C and back again, was therefore given according to classical mechanics by

$$\bar{t}_1 = \frac{l}{c-u} + \frac{l}{c+u} = \frac{2l/c}{1-\frac{u^2}{c^2}}. \tag{5.1}$$

If we introduce $\gamma = \left(1-\frac{u^2}{c^2}\right)^{-\frac{1}{2}}$, we have

$$\bar{t}_1 = \frac{2l\gamma^2}{c}. \tag{5.2}$$

We now use $\bar{\mathbf{v}}$ to denote the velocity of R_2 in the laboratory frame between B and D. If we apply the Galilean rule for the addition of velocities and remember that the speed of light in the ether frame is equal to c, we get

$$\mathbf{c} = \bar{\mathbf{v}} + \mathbf{u} \tag{5.3}$$

from which it follows that

$$\bar{v} = \sqrt{c^2 - u^2}. \tag{5.4}$$

Similarly, the speed of R_2 between D and B was equal to $\sqrt{c^2-u^2}$. The time $\bar{t}_2$ that R_2 should take to travel from B to D and back was therefore

$$\bar{t}_2 = \frac{2l}{\sqrt{c^2-u^2}} = \frac{2l/c}{\sqrt{1-\frac{u^2}{c^2}}} = \frac{2l\gamma}{c}. \tag{5.5}$$

The time difference $\Delta\bar{t}$ between $\bar{t}_1$ and $\bar{t}_2$ was therefore

$$\Delta\bar{t} = \bar{t}_1 - \bar{t}_2 = \frac{2l}{c}\gamma(\gamma-1) > 0. \tag{5.6}$$

Finally, let us assume that we have rotated the interferometer clockwise by 90° around B. Then $\bar{t}_1$ has increased by $2lc^{-1}\gamma(\gamma-1)$, while $\bar{t}_2$ has decreased by $2lc^{-1}\gamma(\gamma-1)$, which means that the time difference $\Delta\bar{t}$ between them has changed by

$$\frac{4l}{c}\gamma(\gamma-1) \approx \frac{2lu^2}{c^3}. \tag{5.7}$$

[7] For details see Rindler 2006, p. 5, or Torretti 1983, p. 29.

When Michelson carried out this experiment in 1881, the change in the time difference caused by the rotation of the interferometer should have been large enough to cause a detectable shift in the interference pattern. But no such shift was observed. Not even when Michelson and Morley improved the accuracy of the first experiment in 1887 by sending the beams R_1 and R_2 back and forth several times along BC and BD respectively.

5.2 LORENTZ'S ETHER PROGRAM

Let us now turn to Lorentz and his program, which aimed to formulate a theory of electrodynamics that could explain, among other things, the null result of the Michelson-Morley experiments. From 1892 and until the introduction of the special theory of relativity by Einstein in 1905, Lorentz adhered to some fundamental ontological assumptions. Like his predecessors, he tacitly accepted Newton's spatio-temporal framework, including the Galilean transformations. He also assumed that matter consisted of both ponderable matter and the ether itself. Following Fresnel, matter was completely permeable to the ether, so that it could move while the ether remained at rest.[8] Furthermore, Lorentz assumed that all ponderable bodies contained small, electrically charged particles – so-called "ions" or "electrons" – and that all electromagnetical phenomena were based on the position and motion of these particles.[9] More precisely, the state of the stationary ether at any point corresponded to the values of two vector fields, the dielectric displacement **d** and the magnetic force **h**, which were determined by the position and motion of the ions according to Lorentz's microscopic version of Maxwell's equations:

$$\begin{aligned} \nabla \cdot \mathbf{d} &= \rho, & \nabla \times \mathbf{d} &= -\frac{1}{4\pi c^2}\frac{\partial \mathbf{h}}{\partial t}, \\ \nabla \cdot \mathbf{h} &= 0, & \nabla \times \mathbf{h} &= 4\pi\left(\rho \mathbf{v} + \frac{\partial \mathbf{d}}{\partial t}\right), \end{aligned} \tag{5.8}$$

where ρ denoted the microscopic charge density of the ions, **v** the velocity of the ions and c the speed of light in the ether. The force **f** that the ether exerted on the ions per unit charge at each point was then given by

$$\mathbf{f} = 4\pi c^2 \mathbf{d} + \mathbf{v} \times \mathbf{h}.^{10} \tag{5.9}$$

[8] Lorentz 1937c, p. 3.
[9] Ibid., p. 5.

As already indicated, the equations referred to a frame S at rest relative to the ether.

Before proceeding, it is instructive to relate Lorentz's equations to their Maxwellian counterparts. In their most complete form, Maxwell's phenomenological equations for ponderable bodies, which govern the relationship between the macroscopic fields of the electric force **E**, the electric displacement **D**, the magnetic force **H** and the magnetic induction **B**, can be expressed as

$$\nabla \cdot \mathbf{D} = \rho_M, \qquad \nabla \times \mathbf{E} = -\frac{D\mathbf{B}}{Dt},$$
$$\nabla \cdot \mathbf{B} = 0, \qquad \nabla \times \mathbf{H} = \mathbf{j} + \frac{D\mathbf{D}}{Dt}, \tag{5.10}$$

where ρ_M denoted the macroscopic charge density and **j** the macroscopic conduction current.[11] $\frac{D}{Dt}$ was called the convective derivative and was defined by

$$\frac{D}{Dt} = \frac{\partial}{\partial t} - \nabla \times (\mathbf{v} \times \ ) + \mathbf{v}(\nabla \cdot \ ). \tag{5.11}$$

Maxwell assumed that the ether was completely dragged by ponderable matter. Therefore, he could regard **D** and **B** as the states of a single medium, consisting of ether and matter, moving at any point in space with a well-defined velocity **v**. Moreover, as Darrigol has emphasized,

> [f]or Maxwell and his followers, the charge density and the conduction current **j** were not primitive concepts: the former corresponded to the longitudinal gradient of the polarization or "displacement" **D**, and the latter to the dissipative relaxation of this polarization in a conducting medium.[12]

Lorentz, on the other hand,

> learned Maxwell's theory in a reinterpretation by Hermann Helmholtz that accommodated the continental interpretation of charge,

[10] Lorentz 1936, §§ 12, 96–97 as well as Lorentz 1937c, §§ 5-8, 12. Lorentz used semi-rationalized electromagnetic units (ursprünglich elektromagnetische Einheiten) in the 1890s. For a detailed comparison between different unit systems, see Lorentz 1904, pp. 82–87.

[11] Darrigol 2006, p. 1.

[12] Ibid., p. 3.

current, and polarization in terms of the accumulation, flow, and displacement of electric particles.[13]

In particular, he assumed that 1) the parts of the ether that are not penetrating ions behave like the ether in a vacuum, 2) the permittivity ϵ and the permeability μ have the same value everywhere as in a vacuum and 3) the total current is the sum of the microscopic convection current (the flow of ions) and the displacement current (the variation of the displacement of ions in dielectrics). This allowed him to reinterpret Maxwell's equations for the stationary ether ($\mathbf{v} = \mathbf{0}$ everywhere) at the molecular level to obtain equations (5.8). Having set out the basic ontological core of Lorentz's theory, let us now examine the development of his thoughts on this subject that led him to reconcile the core with the null result.

5.2.1 The Origin of the Lorentz Transformations

Lorentz's starting point proved to be a crucial problem, to which he devoted much effort in his work "La Théorie électromagnetique de Maxwell et son Application aux Corps mouvants" published in 1892: how to solve the field equations for a material system moving through the ether at constant velocity $\mathbf{u}$.[14] The central importance of this problem arose from the fact that such a solution would allow comparisons with empirical measurements on earth: For any short time interval, a slowly accelerating body like the earth could be considered to move in a straight line and uniformly. If we choose the coordinates (x, y, z, t) of the ether frame S such that $\mathbf{u}$ has the components $u_x = u, u_y = 0, u_z = 0$ relative to S and let $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ denote the coordinate system adapted to $\overline{S}$ in standard configuration relative to (x, y, z, t), Lorentz assumed that

$$\begin{aligned} \overline{x} &= x - ut, & \overline{y} &= y, \\ \overline{z} &= z, & \overline{t} &= t. \end{aligned} \tag{5.12}$$

That is, he naturally assumed the Galilean transformations, which correspond to the classical space-time structure. Lorentz's next step was to express his set of partial equations (5.8) in the new coordinates $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$.

[13] Darrigol 2006, p. 10.

[14] Lorentz 1936, §§ 132-8. Lorentz wrote equations in component form in 1892. I rewrite them in vector form.

Applying the chain rule, he calculated that

$$\nabla = \overline{\nabla}, \qquad \frac{\partial}{\partial t} = \frac{\partial}{\partial \bar{t}} - u\frac{\partial}{\partial \bar{x}}. \tag{5.13}$$

Denoting the velocity of the microscopic charge density $\overline{\rho}$ in the moving frame $\overline{S}$ by $\overline{\mathbf{v}}$, it followed that the absolute velocity of the density becomes $\overline{\mathbf{v}} + \mathbf{u}$. Although Lorentz did not state it explicitly, he assumed that forces are invariant under Galilean transformations. Since spatial dimensions remain unchanged under Galilean transformations, the conservation of charge meant that the microscopic charge density $\overline{\rho}$ with respect to the moving system was equal to the microscopic charge density ρ with respect to the system adapted to the ether:

$$\overline{\rho}(\overline{\mathbf{r}}) = \rho(\mathbf{r}), \tag{5.14}$$

where

$$\overline{\mathbf{r}} = \mathbf{r} - t\mathbf{u}.$$

Substituting these relationships into (5.8) then yields the following equations for the moving frame $\overline{S}$:

$$\begin{aligned} \overline{\nabla}\cdot\overline{\mathbf{d}} &= \overline{\rho}, & \overline{\nabla}\times\overline{\mathbf{d}} &= -\frac{1}{4\pi c^2}\diamond\overline{\mathbf{h}}, \\ \overline{\nabla}\cdot\overline{\mathbf{h}} &= 0, & \overline{\nabla}\times\overline{\mathbf{h}} &= 4\pi\diamond\overline{\mathbf{d}} + 4\pi\overline{\rho}(\overline{\mathbf{v}} + \mathbf{u}), \end{aligned} \tag{5.15}$$

$$\overline{\mathbf{f}} = 4\pi c^2\overline{\mathbf{d}} + (\overline{\mathbf{v}} + \mathbf{u}) \times \overline{\mathbf{h}}, \tag{5.16}$$

where $\diamond$ denotes the scalar operator $\frac{\partial}{\partial \bar{t}} - u\frac{\partial}{\partial \bar{x}}$. Lorentz himself did not change the variables for the field quantities $\mathbf{d}$ and $\mathbf{h}$ as he continued to evaluate them in the ether frame. This was a rather confusing strategy, considering that he was investigating experiments performed in the moving frame $\overline{S}$. I have only introduced barred vectors to emphasize that Lorentz considered these fields as functions of the moving coordinates $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$. In other words, $\overline{\mathbf{d}}(\overline{\mathbf{r}}) = \mathbf{d}(\mathbf{r})$ and $\overline{\mathbf{h}}(\overline{\mathbf{r}}) = \mathbf{h}(\mathbf{r})$, where $\overline{\mathbf{r}} = \mathbf{r} - t\mathbf{u}$.

Lorentz went on to demonstrate that the four field equations (5.15) correspond to a partial differential equation of the form

$$(c^2\overline{\nabla}^2 - \diamond^2)f = G(\bar{t}, \bar{x}, \bar{y}, \bar{z}). \tag{5.17}$$

In view of the availability of a general solution for the three-dimensional wave equation

$$\left(c^2\nabla^2 - \frac{\partial^2}{\partial t^2}\right) f = 0, \tag{5.18}$$

he then tried to find a transformation that would reduce the special case of the four field equations corresponding to $(c^2\overline{\nabla}^2 - \diamond^2)f = 0$ to an instance of the familiar wave equation (5.18). The changes of variables that finally solved the problem for him were

$$x' = \gamma\overline{x}, \qquad y' = \overline{y},$$
$$z' = \overline{z}, \qquad t' = \overline{t} - \frac{u\gamma^2}{c^2}\overline{x}. \tag{5.19}$$

In terms of (x, y, z, t) he thus obtained:

$$x' = \gamma(x - ut), \qquad y' = y,$$
$$z' = z, \qquad t' = \gamma^2\left(t - \frac{u}{c^2}x\right). \tag{5.20}$$

These equations corresponded to the Lorentz transformations except for an additional factor of γ in the expression of t'. As Zahar has remarked,

> [w]e may wonder why Lorentz did not put $t' = \gamma t''$ and thus obtain the full invariance of the operator $[c^2\nabla^2 - \frac{\partial^2}{\partial t^2}]$. [...] The answer is simply that Lorentz was not interested at this stage in invariance for its own sake, but only as a means of solving a mathematical problem, and a very particular one at that.[15]

Stated otherwise, the Lorentz transformations were originally introduced as a purely mathematical tool, without any physical meaning being attached to them.

5.2.2 Electrostatics

Lorentz realized shortly after the publication of "La Théorie électromagnetique de Maxwell et son Application aux Corps mouvants" that the transformation formulas lent themselves to a physical interpretation, which he first outlined in his paper "De Relatieve Beweging van de Aarde en den Aether"

[15] Zahar 1989, p. 57.

from the same year and elaborated in more detail in his treatise *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern* published in 1895.

Let us assume that we have an electrostatic system resting in a frame $\overline{S}$ and moving through the ether at a constant velocity $\mathbf{u}$. Under this assumption, $\overline{\mathbf{v}}$ was equal to $\mathbf{0}$ and all quantities in the frame $\overline{S}$ were independent of the time coordinate $\overline{t}$. It followed that the microscopic field equations (5.15), which were adapted to $\overline{S}$, took the form

$$\overline{\nabla}\cdot\overline{\mathbf{d}} = \overline{\rho}, \qquad \overline{\nabla}\times\overline{\mathbf{d}} = \frac{1}{4\pi c^2}u\frac{\partial\overline{\mathbf{h}}}{\partial\overline{x}},$$
$$\overline{\nabla}\cdot\overline{\mathbf{h}} = 0, \qquad \overline{\nabla}\times\overline{\mathbf{h}} = -4\pi u\frac{\partial\overline{\mathbf{d}}}{\partial\overline{x}} + 4\pi\overline{\rho}\mathbf{u}, \tag{5.21}$$

while the force field $\overline{\mathbf{f}}$, which acted on a unit charge, reduced to

$$\overline{\mathbf{f}} = 4\pi c^2\overline{\mathbf{d}} + \mathbf{u}\times\overline{\mathbf{h}}. \tag{5.22}$$

Lorentz then derived the equations

$$\left(\overline{\nabla}^2 - \frac{u^2}{c^2}\frac{\partial^2}{\partial\overline{x}^2}\right)\overline{\mathbf{d}} = \left(\overline{\nabla} - \frac{u}{c^2}\mathbf{u}\frac{\partial}{\partial\overline{x}}\right)\overline{\rho},$$
$$\left(\overline{\nabla}^2 - \frac{u^2}{c^2}\frac{\partial^2}{\partial\overline{x}^2}\right)\overline{\mathbf{h}} = 4\pi\mathbf{u}\times\overline{\nabla}\,\overline{\rho}. \tag{5.23}$$

For each solution $\overline{\varphi}$ of the equation

$$\left(\overline{\nabla}^2 - \frac{u^2}{c^2}\frac{\partial^2}{\partial\overline{x}^2}\right)\overline{\varphi} = \overline{\rho}, \tag{5.24}$$

there was a corresponding solution to (5.23) given by

$$\overline{\mathbf{d}} = \overline{\nabla}\,\overline{\varphi} - \frac{u}{c^2}\frac{\partial\overline{\varphi}}{\partial\overline{x}}\mathbf{u}, \qquad \overline{\mathbf{h}} = 4\pi\mathbf{u}\times\overline{\nabla}\,\overline{\varphi}. \tag{5.25}$$

Substituting (5.25) into expression (5.22), Lorentz finally obtained $\overline{\mathbf{f}}$ as a function of $\overline{\varphi}$:

$$\overline{\mathbf{f}} = 4\pi(c^2 - u^2)\overline{\nabla}\,\overline{\varphi} = \frac{4\pi c^2}{\gamma^2}\overline{\nabla}\,\overline{\varphi}, \tag{5.26}$$

which demonstrated that a solution of (5.24) solved the problem.[16]

Lorentz continued by imagining a second, fictitious electrostatic system S' that was equal to $\overline{S}$, except for two conditions: 1) S' was at rest in the ether and 2) the dimensions in S' were obtained by dilating the dimensions of $\overline{S}$ along the $\overline{x}$-axis by the factor γ. In other words, the spatial position $(\overline{x}, \overline{y}, \overline{z})$ of $\overline{S}$ corresponded to the following spatial position (x', y', z') of S':

$$x' = \gamma\overline{x}, \qquad y' = \overline{y}, \qquad z' = \overline{z}. \tag{5.27}$$

Since corresponding volume elements carried the same charges, the charge density ρ' in the frame S' was given by

$$\rho' = \frac{\overline{\rho}}{\gamma}. \tag{5.28}$$

Furthermore, in the electrostatic system S' the displacement $\mathbf{d}'$ was the gradient of a scalar function φ' satisfying

$$\nabla'^2\varphi' = \rho' = \frac{\overline{\rho}}{\gamma} \tag{5.29}$$

and the force field $\mathbf{f}'$ was

$$\mathbf{f}' = 4\pi c^2\mathbf{d}' = 4\pi c^2\nabla'\varphi'. \tag{5.30}$$

If we rewrite the scalar operator in equation (5.24) using the transformations (5.27), we get

$$\begin{aligned}\left(\overline{\nabla}^2 - \frac{u^2}{c^2}\frac{\partial^2}{\partial\overline{x}^2}\right) &= \left(1 - \frac{u^2}{c^2}\right)\frac{\partial^2}{\partial\overline{x}^2} + \frac{\partial^2}{\partial\overline{y}^2} + \frac{\partial^2}{\partial\overline{z}^2}\\ &= \frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2}\\ &= \nabla'^2.\end{aligned} \tag{5.31}$$

By applying this result to (5.29) and (5.24), we obtain

$$\overline{\varphi} = \gamma\varphi'. \tag{5.32}$$

[16] Lorentz 1937c, §§ 19-22. Lorentz shifted between writing equations in component and vector form in 1895. I have unified his presentation by expressing as many equations as possible in vector form. I have also simplified it by introducing the assumption $\mathbf{u} = (u, 0, 0)$ at an earlier stage.

Therefore,

$$\mathbf{f}' = 4\pi c^2 \nabla' \varphi' = \frac{4\pi c^2}{\gamma} \nabla' \overline{\varphi} = \frac{4\pi c^2}{\gamma} \left(\frac{1}{\gamma} \frac{\partial \overline{\varphi}}{\partial \overline{x}}, \frac{\partial \overline{\varphi}}{\partial \overline{y}}, \frac{\partial \overline{\varphi}}{\partial \overline{z}} \right). \tag{5.33}$$

Comparing equations (5.26) and (5.33), the force field $\overline{\mathbf{f}}$ was thus related to $\mathbf{f}'$ by

$$\overline{f}_x = f'_x, \qquad \overline{f}_y = \gamma^{-1} f'_y, \qquad \overline{f}_z = \gamma^{-1} f'_z,$$

which Lorentz wrote as

$$\overline{\mathbf{f}} = (1, \gamma^{-1}, \gamma^{-1}) \mathbf{f}', \tag{5.34}$$

where $(1, \gamma^{-1}, \gamma^{-1})$ denoted a 3×3 diagonal matrix. This equation between the force fields generated in $\overline{S}$ and S' showed that the force $\overline{\mathbf{f}}$ acting at a point of $\overline{S}$ vanished if and only if the force $\mathbf{f}'$ acting at the corresponding point of S' vanished. Secondly, it demonstrated that the motion of the earth with respect to the ether could have no first-order influence in $\frac{u}{c}$ on electrostatic experiments carried out on the surface of the earth, since $\gamma \doteq 1$ in this approximation.[17]

5.2.3 Theory of Corresponding States

Using a similar technique, Lorentz also demonstrated that the result of optical experiments conducted on earth was independent of the relative motion between the earth and the ether to first-order in $\frac{u}{c}$, where u denoted the relative speed. By averaging the equations (5.15) over volume elements, which contained a large number of ions but could still be considered infinitesimal compared to tangible objects, and by neglecting second-order terms in $\frac{u}{c}$, he was able to derive equations for the macroscopic fields in a homogeneous and isotropic dielectric, which was adapted to a frame $\overline{S}$ moving with velocity $\mathbf{u}$ with respect to the ether.[18] In this effort, he defined the material polarization $\overline{\mathbf{P}}$ as the average electric moment over all ions located in the same volume element:

$$\overline{\mathbf{P}} = \langle e \overline{\mathbf{r}} \rangle, \tag{5.35}$$

[17] Ibid., §§ 23–24.

[18] Ibid., §§ 40-2, 45, 51–2.

where e denoted the charge of an ion and $\overline{\mathbf{r}}$ the displacement of the ion from its (unforced) equilibrium position. This vector field made it possible to introduce the macroscopic polarization density $\overline{\rho}_P$ for a moving polarized body:

$$\overline{\rho}_P = -\overline{\nabla}\cdot\overline{\mathbf{P}}. \tag{5.36}$$

To rephrase, $\overline{\rho}_P$ was the part of the total macroscopic charge density $\overline{\rho}_M$ ($\langle\,\overline{\rho}\,\rangle$) that was generated by polarization. To see this, we note that the total charge generated by the polarization in a given volume $\overline{\mathcal{T}}$ was opposite to the amount of charge squeezed out of the surface $\partial\overline{\mathcal{T}}$:

$$\int_{\overline{\mathcal{T}}} \overline{\rho}_P \; \mathrm{d}\overline{\tau} = -\int_{\overline{\mathcal{T}}} \overline{\nabla}\cdot\overline{\mathbf{P}} \; \mathrm{d}\overline{\tau} = -\int_{\partial\overline{\mathcal{T}}} \overline{\mathbf{P}}\cdot\mathrm{d}\overline{\boldsymbol{\sigma}}. \tag{5.37}$$

In the last step, the divergence theorem has been applied to convert the volume integral into a surface integral. Lorentz then introduced the macroscopic fields

$$\overline{\mathbf{H}} = \langle\,\overline{\mathbf{h}}\,\rangle \tag{5.38}$$

and

$$\overline{\mathbf{D}} = \langle\,\overline{\mathbf{d}}\,\rangle + \;\overline{\mathbf{P}}. \tag{5.39}$$

$\overline{\mathbf{D}}$ was formally equivalent to Maxwell's dielectric displacement because of the identities

$$\overline{\nabla}\cdot\overline{\mathbf{D}} = \langle\,\overline{\nabla}\cdot\overline{\mathbf{d}}\,\rangle + \overline{\nabla}\cdot\overline{\mathbf{P}} = \overline{\rho}_M - \overline{\rho}_P = \overline{\rho}_F, \tag{5.40}$$

where $\overline{\rho}_F$ was the macroscopic density of *free* charge (i.e. it was the part of the total macroscopic charge density $\overline{\rho}_M$ that was not the result of polarization).[19]

Assuming that all ions were bound in the dielectric, Lorentz was able to set $\overline{\rho}_F$ equal to 0. By carefully studying the microscopic polarization processes, he then showed that if monochromatic waves propagated in the dielectric, $\overline{\mathbf{P}}$ would be proportional to the macroscopic force field

$$\overline{\mathbf{F}} = \langle 4\pi c^2\overline{\mathbf{d}} + \mathbf{u}\times\overline{\mathbf{h}}\rangle \tag{5.41}$$

to a reasonable approximation. Put simply,

$$\overline{\mathbf{F}} = \chi\overline{\mathbf{P}}, \tag{5.42}$$

[19] Lorentz 1937c, §§ 40-2.

where χ depended on the frequency of the waves. Defining

$$\epsilon = 1 + \frac{4\pi c^2}{\chi}, \tag{5.43}$$

and by averaging his microscopic field equations, he arrived at a system of first-order equations:

$$\overline{\nabla}\cdot\overline{\mathbf{D}} = 0, \qquad \overline{\nabla}\times\overline{\mathbf{F}} = -\frac{\partial\overline{\mathbf{H}}}{\partial\bar{t}},$$
$$\overline{\nabla}\cdot\overline{\mathbf{H}} = 0, \qquad \overline{\nabla}\times\left(\overline{\mathbf{H}} - \frac{1}{c^2}\mathbf{u}\times\overline{\mathbf{F}}\right) = 4\pi\frac{\partial\overline{\mathbf{D}}}{\partial\bar{t}},$$

$$\epsilon\overline{\mathbf{F}} = 4\pi c^2\overline{\mathbf{D}} + \mathbf{u}\times\overline{\mathbf{H}}, \tag{5.44}$$

for the optics of moving dielectrics.[20] For a system that was at rest with respect to the ether, this set of equations reduced to

$$\nabla\cdot\mathbf{D} = 0, \qquad \nabla\times\mathbf{F} = -\frac{\partial\mathbf{H}}{\partial t},$$
$$\nabla\cdot\mathbf{H} = 0, \qquad \nabla\times\mathbf{H} = 4\pi\frac{\partial\mathbf{D}}{\partial t},$$

$$\epsilon\mathbf{F} = 4\pi c^2\mathbf{D}. \tag{5.45}$$

As already indicated, Lorentz used this result to prove that the relative motion of the earth and the ether could not have a first-order influence in $\frac{u}{c}$ on optical experiments on earth. To support his argument, he approximated the previous change of variables (5.19) to first order in $\frac{u}{c}$:

$$x' = \bar{x}, \qquad y' = \bar{y},$$
$$z' = \bar{z}, \qquad t' = \bar{t} - \frac{u\bar{x}}{c^2}. \tag{5.46}$$

[20] Ibid., §§ 45, 51-2. Lorentz did not limit his derivations to (homogeneous) isotropic dielectrics. To obtain his general equations for (homogeneous) anisotropic dielectrics, replace ϵ with a 3×3 diagonal matrix $(\epsilon_1, \epsilon_2, \epsilon_3)$.

Because t' varied as a function of spatial position, Lorentz called it *local time*.[21] By means of the chain rule, the differentials became

$$\nabla' = \overline{\nabla} + \frac{\mathbf{u}}{c^2}\frac{\partial}{\partial \bar{t}}, \qquad \frac{\partial}{\partial t'} = \frac{\partial}{\partial \bar{t}}. \tag{5.47}$$

Introducing

$$\mathbf{D}' = \overline{\mathbf{D}} + \frac{1}{4\pi c^2}\mathbf{u} \times \overline{\mathbf{H}}, \qquad \mathbf{H}' = \overline{\mathbf{H}} - \frac{1}{c^2}\mathbf{u} \times \overline{\mathbf{F}}, \tag{5.48}$$

and neglecting second-order terms in $\frac{u}{c}$, he arrived at the following equations by rewriting (5.44) using (5.47) and (5.48):

$$\nabla' \cdot \mathbf{D}' = 0, \qquad \nabla' \times \overline{\mathbf{F}} = -\frac{\partial \mathbf{H}'}{\partial t'},$$

$$\nabla' \cdot \mathbf{H}' = 0, \qquad \nabla' \times \mathbf{H}' = 4\pi \frac{\partial \mathbf{D}'}{\partial t'},$$

$$\epsilon \overline{\mathbf{F}} = 4\pi c^2 \mathbf{D}'. \tag{5.49}$$[22]

By pointing out that the form of these equations corresponded exactly to the form of the equations (5.45) for the system of bodies at rest, Lorentz obtained what he called the theorem of corresponding states.[23] It can be paraphrased as follows:

Suppose there is a system of bodies fixed to a frame S which is initially at rest in the ether and a solution to the field equations, such that **D**, **F**, and **H** are certain functions of the coordinates (x, y, z, t) adapted to S. It follows that there is a solution for the same system of bodies moving across the ether with constant velocity **u**, such that $\mathbf{D}'$, $\overline{\mathbf{F}}$, and $\mathbf{H}'$ are the same functions of $(\bar{x}, \bar{y}, \bar{z}, t')$.

According to Lorentz, the significance of this theorem arose from the fact that it implied the absence of effects of the earth's motion on well-known first-order optical experiments.

Let me explain the theorem of corresponding states by distinguishing more clearly between the different systems involved. The theorem in this version is as follows:

---

[21] Lorentz 1937c, p. 81.
[22] ibid., §§ 56-57.
[23] Ibid., § 60.

Suppose there is a system of bodies fixed to a stationary frame S and a solution of the field equations such that the fields **D**, **F** and **H** are functions of the coordinates (x, y, z, t) adapted to S satisfying (5.45). Under the assumption that $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ adapted to a moving frame $\overline{S}$ are related to (x, y, z, t) by the Galilean transformation (5.12), it follows that for the same system of bodies fixed to the moving frame $\overline{S}$, there is a solution such that $\overline{\mathbf{D}}$, $\overline{\mathbf{F}}$ and $\overline{\mathbf{H}}$ – i.e. the fields **D**, **F** and **H** as functions of the moving coordinates $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ – satisfy (5.44). In addition, by a purely mathematical transformation of $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ and $\overline{\mathbf{D}}$, $\overline{\mathbf{F}}$ and $\overline{\mathbf{H}}$ by means of (5.46) and (5.48) it is possible to obtain a coordinate system (x', y', z', t') and vectors $\mathbf{D}'$, $\mathbf{F}'$ and $\mathbf{H}'$ that satisfy the equations (5.45) for the system of bodies at rest.[24] Thus, the same vector functions that give the real fields **D**, **F** and **H** as functions of (x, y, z, t) in the system at rest, will give the fictitious fields $\mathbf{D}'$, $\mathbf{F}'$ and $\mathbf{H}'$ as functions of (x', y', z', t') in the corresponding state defined by the moving system.

This rendition of the theorem makes it clear that Lorentz distinguished between three systems S, $\overline{S}$ and S': a system of bodies at rest, the same system of bodies in uniform rectilinear motion and a fictitious system that was defined by the system in motion but behaved like the system at rest (i.e. "a fictitious system brought to rest"[25] as Darrigol puts it). The reason why the fictitious system brought to rest had a physical meaning for the moving system ($\mathbf{u} \neq \mathbf{0}$) was due to the fact that it allowed a comparison between the moving system in $\overline{S}$ and the corresponding system at rest in S. For example, let us assume that it is dark at a certain point $\mathbf{r} = \mathbf{p}$ in S. This is equivalent to the vanishing of the real fields **D** and **H** at this point. According to the theory of corresponding states, it follows that the fictitious fields $\mathbf{D}'$ and $\mathbf{H}'$ vanish at $\mathbf{r}' = \mathbf{p}$ in S'. However, according to the relations (5.48), $\mathbf{D}'$ and $\mathbf{H}'$ vanish at $\mathbf{r}' = \mathbf{p}$ if and only if $\overline{\mathbf{D}}$ and $\overline{\mathbf{H}}$ vanish at $\overline{\mathbf{r}} = \mathbf{p}$. Therefore, together with the tacit assumption that the observed positions of dark fringes in the moving frame $\overline{S}$ coincided with those generated by $\overline{\mathbf{D}}$ and $\overline{\mathbf{H}}$,[26] the biconditional implied the invariance of patterns of light and darkness under uniform rectilinear motion to first order in $\frac{u}{c}$. In short, the theory of corresponding states explained the negative result of any first-order optical

[24] For the sake of simplicity, I have introduced $\mathbf{F}'(\mathbf{r}', t') = \overline{\mathbf{F}}(\overline{\mathbf{r}}, t')$.
[25] Darrigol 2006, p. 15.
[26] Remember that Lorentz did not transform the field quantities.

Fig. 5.2: Patterns of light and darkness.

ether drift experiment: the motion of the earth had no first-order influence on experiments involving interference, diffraction or polarization.[27]

Finally, the theory made it possible to derive Fresnel's dragging coefficient. Let us assume that we have a plane-wave solution for the fields **D** and **H** with phase

$$\omega t - \mathbf{k}\cdot\mathbf{r} + \delta \tag{5.50}$$

for the rest system S, where **k** denotes the propagation vector and ω the angular frequency. Furthermore, we assume that $\overline{S}$ moves in an arbitrary direction with respect to S given by the velocity **u** and generalize the transformations (5.46) between $\overline{S}$ and S' accordingly:

$$\mathbf{r}' = \bar{\mathbf{r}}, \quad t' = \bar{t} - \frac{\mathbf{u}\cdot\bar{\mathbf{r}}}{c^2}, \tag{5.51}$$

where $\bar{\mathbf{r}} = (\bar{x}, \bar{y}, \bar{z})$ and $\mathbf{r}' = (x', y', z')$. For transversal waves, which propagate in a stationary dielectric with a phase velocity that is characterized

[27] Lorentz 1937c, § 60.

Fig. 5.3: Electromagnetic wave.

by

$$v = \frac{\omega}{k}, \tag{5.52}$$

the theory of corresponding states then attached plane waves in the moving system with phase

$$\begin{aligned}\omega t' - \mathbf{k}\cdot\mathbf{r}' + \delta &= \omega\bar{t} - \frac{\omega}{c^2}\mathbf{u}\cdot\bar{\mathbf{r}} - \mathbf{k}\cdot\bar{\mathbf{r}} + \delta \\ &= \omega\bar{t} - \left(\mathbf{k} + \frac{\omega}{c^2}\mathbf{u}\right)\cdot\bar{\mathbf{r}} + \delta = \omega\bar{t} - \bar{\mathbf{k}}\cdot\bar{\mathbf{r}} + \delta\end{aligned}$$

and phase velocity

$$\bar{v} = \frac{\omega}{\bar{k}} = \frac{\omega}{\left|\mathbf{k} + \frac{\omega}{c^2}\mathbf{u}\right|},$$

where $\bar{\mathbf{k}} = \mathbf{k} + \frac{\omega}{c^2}\mathbf{u}$. To first order in $\frac{u}{c}$, Lorentz thus obtained

$$\bar{v} = v - \frac{v^2}{c^2}u_{\parallel}, \tag{5.53}$$

where $u_{\parallel}$ denoted the component of the velocity $\mathbf{u}$ in the direction of the propagation vector $\bar{\mathbf{k}}$:

$$u_{\parallel} = \frac{\mathbf{u}\cdot\bar{\mathbf{k}}}{\bar{k}} = \frac{\mathbf{u}\cdot\left(\mathbf{k} + \frac{\omega}{c^2}\mathbf{u}\right)}{\left|\mathbf{k} + \frac{\omega}{c^2}\mathbf{u}\right|}.$$

Noting that the phase velocity v in a transparent body at rest was related to the refractive index n by $n = \frac{c}{v}$, formula (5.53) finally turned into

$$\bar{v} = \frac{c}{n} - \frac{u_{\|}}{n^2} \tag{5.54}$$

in accordance with Fresnel's result.[28]

5.2.4 *Lorentz Contraction*

Lorentz was still faced with the problem of accounting for the lack of a second-order effect in the interferometer experiments carried out by Michelson and Morley. To this end, he introduced the hypothesis that the molecular forces that held solid bodies like the arms of the interferometer together were transformed like electrostatic forces during the transition from the ether frame to the moving frame. Thus, if a body moved through the ether at a constant speed u, the equilibrium of molecular forces between the particles could only be maintained if the dimensions of the body were contracted by a factor $\frac{1}{\gamma}$ in the direction of motion.[29] Applying this contraction hypothesis to the Michelson-Morley experiments, it followed that if $|BD|$ was equal to l (before the rotation of the apparatus), $|BC|$ would be equal to $\frac{l}{\gamma}$. For the time difference $\Delta\bar{t}$ we therefore have

$$\Delta\bar{t} = \bar{t}_1 - \bar{t}_2 = \frac{2l/\gamma}{c}\gamma^2 - \frac{2l}{c}\gamma = 0. \tag{5.55}$$

The hypothesis thus enabled Lorentz to explain the invariance of the interference pattern under the rotation of the interferometer by 90° around B as a second-order effect in $\frac{u}{c}$.

5.3 CONCLUSION

Lorentz's 1895 theory demonstrated, among other things, the invariance of electrostatics and optics to first order in $\frac{u}{c}$ and provided an explanation for the null result of the Michelson-Morley experiments to second order in $\frac{u}{c}$. As we have seen, Lorentz obtained these results by introducing a system at rest with respect to the ether next to the moving system under investigation. While the first two results assumed only an imaginary system at

[28] Lorentz 1937c, §§ 68-70.

[29] For details see Lorentz 1937b, p. 222, Lorentz 1937c, §§ 89–92, or Zahar 1989, pp. 62–3.

rest, his treatment of the Michelson-Morley experiments by means of the molecular force hypothesis required a realistic interpretation of the system at rest: the material configuration of the system at rest corresponded to the material configuration of the moving system when it was brought to rest. However, although the material parts determined the state of the electromagnetic field, Lorentz only applied the hypothesis to predict a change in the dimensions of the material parts (such as the arms of the interferometer). The reason why this strategy was permissible in the special case of the Michelson-Morley experiments was that the alleged validity of the Maxwell equations with respect to the ether implied that light would propagate at a constant speed c in the ether, regardless of the relative motion between the light source and the ether. However, in general, Lorentz did not attempt to explain how the hypothesis of contractile bodies could influence optical phenomena. Instead of attempting to achieve a homogenous account of electrical and optical phenomena, Lorentz's 1895 theory consisted of two quite heterogeneous parts: a general treatment of electrostatics and optics accounting for most first-order ether drift experiments, on the one hand, and a number of special hypotheses each introduced to explain a specific second-order experiment, on the other.

CHAPTER 6

Poincaré's Physics of Principles

Let us now take a look at how Lorentz's ideas were received and interpreted by Henri Poincaré. Reviewing various electromagnetic theories in his book *La Science et l'Hypothèse* (1902), he concluded that Lorentz's theory offered the most adequate explanation for a variety of electromagnetic and optical phenomena, including "the aberration of light, the partial impulse of luminous waves, magnetic polarisation, and Zeeman's experiment."[1] Poincaré was not entirely satisfied, however, because Lorentz's 1895 theory did not satisfy certain general principles, which he regarded as the most fundamental achievement in physics to date. For him, the decisive development towards this physics of general principles was the development from the Newtonian to the Lagrangian ideal of mechanics:

> [A] day arrived when the conception of central forces no longer appeared sufficient [...] What was done then? The attempt to penetrate into the detail of the structure of the universe, to isolate the pieces of this vast mechanism, to analyze one by one the forces which put them in motion, was abandoned, and we were content to take as guides certain general principles the express object of which is to spare us this minute study.[2]

Poincaré's list of principles included the principle of the conservation of energy, the principle of energy dissipation (Carnot's principle), the principle of equality of action and reaction (Newton's third law), the principle of relative motion, the principle of conservation of mass and the principle of least action.[3] The reason why he perceived this physics of general principles to be extremely important was, according to the quote, that it enabled physicists to consider the dynamical problem of a whole system without having to decompose it into parts. In contrast to the Newtonian ideal, the application of the Lagrangian method made it possible, for example, to develop the dynamical equations of a rigid body by introducing kinematical constraints without knowing the internal forces responsible for the rigidity of the body.

[1] Poincaré 1952, p. 243.
[2] Poincaré 1929, p. 299.
[3] Ibid., p. 300.

6.1 MECHANICAL EXPLANATION

Poincaré spoke of demonstrating the possibility of a mechanical explanation for a series of natural phenomena without having to find the explanation itself, i.e. without having to reduce the phenomena to motion of ordinary matter or fictitious fluids. He found the paradigmatic example of this superior approach in the work of James Clerk Maxwell on electromagnetic phenomena:

> It is easy now to understand Maxwell's fundamental ideal. To demonstrate the possibility of a mechanical explanation of electricity we need not trouble to find the explanation itself; we need only know the expression of the two functions T and U, which are the two parts of energy, and to form with these two functions Lagrange's equations, and then to compare these equations with the experimental laws.[4]

In his general account of a mechanical explanation of a phenomena, Poincaré first made it clear that such phenomena are accessible to "a certain number of parameters $[q_1, \ldots, q_n]$ which are reached directly by experiment, and which can be measured."[5] Furthermore, the experimentally ascertainable laws of this set of generalized coordinates $q_1, \ldots, q_n$ could be given in the form of differential equations with respect to time t. Presupposing the principle of conservation of energy and the principle of least action, the French mathematician then demonstrated that the following set of constraints were necessary and sufficient conditions for the existence of a complete mechanical explanation:

1. There existed a potential energy function U, which depended solely on the generalized coordinates q_k, and a kinetic energy function T, which depended on the coordinates q_k and their time derivatives $\dot{q}_k$.

2. U and T fulfilled the principle of conservation of energy, which implied that their sum $T + U$ was constant.

3. The theoretically specifiable Lagrangian equations of motion, which could be derived with the aid of the principle of least action, agreed with the experimentally ascertainable laws.[6]

[4] Poincaré 1952, p. 223.
[5] Ibid., p. 217.
[6] Poincaré 1891, pp. 3–6.

Poincaré was the first to prove that these conditions were necessary; i.e. that any mechanical explanation of a series of natural phenomena had to satisfy them. For the sake of argument, he assumed that we are given a mechanical explanation for a series of natural phenomena. Such an explanation would account for the phenomena either by the motion of ordinary matter or by hypothetical fluids. Without loss of generality, Poincaré opted for the latter strategy and assumed that the fluids were formed from a large number p of isolated molecules with masses m_i and coordinates x_i, y_i, z_i. A mechanical explanation of the phenomena would then firstly require that the equations of motion of these hypothetical molecules satisfy both the principle of conservation of energy and the principle of least action. The former presupposed that the total energy was constant and could be divided into two parts:

1. A potential energy function U of the $3p$ coordinates x_i, y_i, z_i, which made it possible to write down the $3p$ equations of motion in the form

$$\begin{aligned} m_i \frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} &= -\frac{\partial U}{\partial x_i}, \\ m_i \frac{\mathrm{d}^2 y_i}{\mathrm{d}t^2} &= -\frac{\partial U}{\partial y_i}, \\ m_i \frac{\mathrm{d}^2 z_i}{\mathrm{d}t^2} &= -\frac{\partial U}{\partial z_i}. \end{aligned} \tag{6.1}$$

2. A kinetic energy function T of the $3p$ time derivatives $\dot{x}_i$, $\dot{y}_i$, $\dot{z}_i$:

$$T = \frac{1}{2} \sum_{i=1}^{p} m_i \left(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2 \right).$$

Secondly, to enable a comparison between the equations of motion (6.1) and the experimentally discernable laws, it was necessary to express the $3p$ coordinates x_i, y_i, z_i as functions of the n measurable quantities q_k. As a result, U became a function of q_k and T of q_k and their time derivatives $\dot{q}_k$, while the equations of motion assumed their Lagrangian form, derivable by means of the principle of least action:

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial T}{\partial \dot{q}_k} - \frac{\partial T}{\partial q_k} + \frac{\partial U}{\partial q_k} = 0. \tag{6.2}$$

This interpretation corresponded to a complete mechanical explanation if and insofar as the experimentally ascertainable laws could be cast in Lagrangian form (6.2). This concluded the necessity part of the proof: The possibility of a mechanical explanation of a series of natural phenomena required the existence of two functions $U(q_1, \ldots, q_n)$ and $T(q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n)$, so that their sum was constant and the equations (6.2) in the form of these functions corresponded to the experimentally ascertainable laws.

Let us take a look at the sufficiency part. To prove it, Poincaré had to show how to specify a mechanical system of hypothetical molecules underlying the series of natural phenomena under consideration by means of functions $U(q_1, \ldots, q_n)$ and $T(q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n)$. This condition, he argued, could be satisfied by the existence of p constants $m_1, \ldots, m_p$ and $3p$ functions of $q_1, \ldots, q_n$,

$$\phi_i(q_1, \ldots, q_n), \quad \psi_i(q_1, \ldots, q_n), \quad \theta_i(q_1, \ldots, q_n) \qquad (i = 1, \ldots, p),$$

which fulfilled the following requirement:

$$T(q_1, \ldots, q_n; \dot{q}_1, \ldots, \dot{q}_n) = \frac{1}{2} \sum_{i=1}^{p} m_i \left(\dot{\phi}_i^2 + \dot{\psi}_i^2 + \dot{\theta}_i^2 \right), \tag{6.3}$$

where, for instance,

$$\dot{\phi}_i = \sum_{k=1}^{n} \dot{q}_k \frac{\partial \phi_i}{\partial q_k}. \tag{6.4}$$

For Poincaré could then consider m_i as the masses and $x_i = \phi_i$, $y_i = \psi_i$, $z_i = \theta_i$ $(i = 1, \ldots, p)$ as the coordinates of the hypothetical molecules. However, he found that by choosing a sufficiently large number p, it was possible to find not just one, but an infinite number of sets of constants and functions that fulfill this requirement. It was thus possible to show that a mechanical explanation can be given without actually finding a particular mechanical configuration underlying them. All a theoretical physicist had to do was to specify appropriate functions of potential and kinetic energy in terms of measurable quantities. He therefore came to the conclusion that applying general principles in this way "to different physical phenomena is sufficient for our learning of them all that we could reasonably hope to know of them."[7]

[7] Poincaré 1929, p. 301.

6.2 EPISTEMOLOGY OF GENERAL PRINCIPLES

With regard to the epistemological status of his general principles, Poincaré remarked that none of them was a priori true because contingent facts about the world had played an important role in their formulation. The truth of the principle of inertia, for example, could not be proved a priori because it was possible to imagine worlds in which a body that was not acted upon by external forces maintained its position or acceleration rather than its velocity.[8] In such worlds, Newton's principle would not apply. Therefore, it could not be true a priori.[9] Similarly, the principle of relative motion could not be demonstrated a priori. In one of Poincaré's formulations, the principle stated that "[t]he movement of any system whatever ought to obey the same laws, whether it is referred to fixed axes or to movable axes which are implied in uniform motion in a straight line."[10] One could also imagine a world in which laws were independent of absolute acceleration and not of absolute velocity. The latter would imply that absolute velocity would be empirically verifiable, so that the principle of relative motion would not apply in this world.

As for the empirical origin of the general principles, Poincaré pointed out that

> [t]he principles are results of experiments boldly generalized; but they seem to derive from their very generality a high degree of certainty. In fact, the more general they are, the more frequent are the opportunities to check them, and the verifications multiplying, taking the most varied, the most unexpected forms, end by no longer leaving place for doubt.[11]

However, the principles should not be taken as empirical truths. For although they were based on bold inductive generalizations from experiments, they had attained a level of generality that made them immune to empirical falsification. Poincaré exemplified how inductive procedures can lead to such irrefutable principles using the example of Newton's three laws of motion. The latter were only defined as strictly true if they were applied

---

[8] The principle of inertia was not on Poincaré's list of general principles. According to the latter, however, Newton's three laws of motion had the same epistemological status as the more general principles.

[9] Poincaré 1952, pp. 91–2.

[10] Ibid., p. 111.

[11] Poincaré 1929, p. 301.

to a system of bodies that was completely abstracted from all external influences. Thus, they would only be strictly true if they were applied to the entire universe. Applied to all other cases, such as the system of celestial bodies in our solar system, they would merely be approximately true. Since the strict case, by its very nature, cannot be verified, he concluded that these principles cannot really be verified experimentally.

Newton's laws of motion could neither be falsified by experiments. Poincaré gave two reasons for his assertion. Firstly, since they applied to completely isolated systems not accessible to experiment, they could "never be submitted to a decisive test."[12] Secondly, even if we were to suppose that a mechanical principle could be verified experimentally, it could always be safeguarded by the introduction of hidden masses whose motion would explain the deviation.[13] To summarize, the truth of general principles could not be established by a priori reasoning or on an empirical basis. But what then was their status? They were conventions, Poincaré claimed, and therefore belonged in his hierarchy of sciences on the level of mechanics.

6.3 CONVENTIONALISM

Conventions were principles that were held to be true, but whose truth was not merely a matter of definition. For even if a convention can neither be confirmed nor falsified by an experiment, its choice "is not absolutely arbitrary; it is not the child of our caprice. We admit it because certain experiments have shown that it will be convenient."[14] Poincaré spoke in this context of elevating an empirical law to the status of a convention. Thus, a principle could play the role of an (approximately true) empirical law in one field and the role of a (strictly true) convention in another field. The principle of relative motion, for example, was both a consequence of Poincaré's conventionalist conception of space and an empirically well-established law of Newtonian physics:

> [I]t [the principle of relative motion] is imposed on us for two reasons: the most common experiment confirms it; the consideration of the contrary hypothesis is singularly repugnant to the mind.[15]

[12] Poincaré 1952, p. 96.
[13] Ibid., pp. 96–7.
[14] Ibid., p. 136.
[15] Ibid., p. 111.

However, its empirical status in the field of electrodynamics remained unclear at the turn of the 20th century (as we shall see later). It would therefore be wrong to interpret Poincaré's distinction as clear cut.

As Stathis Psillos has pointed out, Poincaré did not formulate a theory of what it takes for a convention to be convenient, but he distilled at least three elements.[16] According to the latter, a principle was a convention when

1. it could be applied to the study of nature in a way that gave approximately correct results,
2. it could be extended to cover new facts and predict new phenomena,
3. it could form a common theoretical framework for explaining seemingly dissimilar phenomena.

Put simply, principles were conventions insofar as they were able to play an indispensable regulative role in the development of physics. If a principle no longer played this role, it should be abandoned rather than safeguarded. The latter was a very critical step that should only be considered after careful deliberation.[17]

6.4 THE TRUE RELATIONS OF THINGS

Having clarified Poincaré's view on conventions, let us see how he combined it with an important corollary of his demonstration of the possibility of a mechanical explanation: If one could find a mechanical explanation for one series of phenomena, it was possible to find an infinite number of others. Formally, this was a direct consequence of the sufficiency part of his proof. According to Poincaré, it was also confirmed by the actual history of physics.[18] This history made it abundantly clear that some elements had remained unaltered between transitions: The conceptual wrapping could differ immensely between one physical theory and one of its immediate successors, while its system of differential equations remained unchanged. With respect to these equations, he further remarked:

> [T]hese equations express relations, and if the equations remain true, it is because the relations preserve their reality. They teach us now,

[16] Psillos 1996, pp. 182–3.
[17] Poincaré 1952, pp. 166–7.
[18] Poincaré 1891, pp. 5–6.

> as they did then, that there is such and such a relation between this thing and that; only, the something which we then called *motion*, we now call *electric current*. But these are merely names of the images we substituted for the real objects which Nature will hide for ever from our eyes. The true relations between these real objects are the only reality we can attain.[19]

Thus, while the purported images might fluctuate rapidly between different mechanical explanations for the same phenomena, the empirical discernible laws underlying both explanations could nevertheless remain the same. As we have already mentioned, this circumstance was a consequence of the corollary of the possibility of a mechanical explanation, which in turn presupposed the principle of energy conservation and the principle of least action. Seen in this way, Poincaré explained, the general principles "thus represent the quintessence of innumerable observations:"[20] they expressed what was common to the formulation of numerous physical laws.

6.5 CONCLUSION

In his assessment of Lorentz's 1895 electromagnetic theory, Henri Poincaré acknowledged its suitability for explaining various electromagnetic and optical phenomena, but criticized its failure to adhere to certain general principles, which he regarded as fundamental achievements of physics. He understood this physics of general principles as a development from the Newtonian to the Lagrangian ideal of mechanics. The latter allowed physicists to study entire systems without having to break them down into smaller parts. In particular, this approach enabled mechanical explanations of phenomena without having to know the underlying forces responsible for the observed behavior. For Poincaré, this meant that the true reality in physics lies in the relations expressed by the system of differential equations, and not in the specific images or interpretations used to describe physical phenomena. In line with the above, he regarded these principles as conventions that did not have to be a priori or empirically true. Rather, they were the result of bold inductive generalizations from experiments. Although they were not empirically falsifiable, they played an indispensable role in the progress of physics. However, when they were no longer useful, they should be abandoned and replaced.

[19] Poincaré 1952, p. 161.
[20] Ibid., p. 166.

Poincaré's Contribution to the Lorentz Jubilee

Let us now return to Lorentz's theory of 1895 to find out how Poincaré evaluated it in the context of his physics of principles. In his contribution "La Théorie de Lorentz et le Principe de la Réaction" (1900) to the Lorentz jubilee volume, he was quick to point out a fundamental problem with the theory, namely that it did not fulfill the principle of reaction when applied to ordinary matter. Second, he demonstrated how this fundamental problem was compounded by a thought experiment which showed that the principle of reaction had to apply to ordinary matter, since its negation in conjunction with the validity of the principle of relative motion for ordinary matter implied the possibility of perpetual motion. The rest of the jubilee article was then devoted to his attempt to resolve the problem. As we will examine in detail below, his solution involved both the safeguarding of the two principles via the ether and a crucial physical reinterpretation of Lorentz's theory.

7.1 DERIVATION OF THE PRINCIPLE OF REACTION

Let us begin by explaining Poincaré's thought experiment: Consider an isolated system consisting of two bodies rigidly connected by a rod. If action and reaction did not cancel each other out, the net force acting on the system would not be zero. However, assuming that the principle of relative motion applied, the net force was independent of the position and velocity of the system. Therefore, the isolated system continued to accelerate.[1]

More generally, Poincaré observed that in a conservative system where U was the potential energy, m_i were the point masses of the system, and $\mathbf{v}_i = (\dot{x}_i, \dot{y}_i, \dot{z}_i)$ were their velocities relative to a coordinate system S at rest,[2] we would have the energy equation

$$\sum_i \frac{m_i}{2} \left(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2 \right) + U = C, \tag{7.1}$$

where C was a real constant. Assuming the Galilean rule for the addition of velocities between S and a second coordinate system $\overline{S}$ moving with constant speed u parallel to the x-axis of S,

$$\dot{x}_i = \dot{\overline{x}}_i + u, \qquad \dot{y}_i = \dot{\overline{y}}_i, \qquad \dot{z}_i = \dot{\overline{z}},$$

[1] See Poincaré 2008, p. 481, Poincaré 1895a or Poincaré 1895b.

[2] Poincaré assigned to absolute space the status of a convention within mechanics.

he was able to rewrite the energy equation for absolute motion in terms of the coordinates of the moving system:

$$\sum_i \frac{m_i}{2}\left(\left(\dot{x}_i + u\right)^2 + \dot{y}_i^{\,2} + \dot{z}_i^{\,2}\right) + U = C. \tag{7.2}$$

"By virtue of the *principle of relative motion*," he argued, "U depends only on the *relative* motion being no different from the laws of absolute motion,"[3] from which he derived the energy equation for the conservative system in relative motion:

$$\sum_i \frac{m_i}{2}\left(\dot{x}_i^{\,2} + \dot{y}_i^{\,2} + \dot{z}_i^{\,2}\right) + U = \overline{C}. \tag{7.3}$$

Subtracting equation (7.3) from (7.2), Poincaré obtained

$$\sum_i m_i \dot{x}_i = \overline{K}, \tag{7.4}$$

where $\overline{K}$ was a real constant. Identifying this equation as the "analytic expression of the principle of reaction,"[4] he concluded:

> The principle of reaction appears to us, therefore, as a consequence of the principle of energy and the principle of relative motion. The latter weighs heavily on our thoughts when we consider an isolated system.[5]

By applying this consequence to Lorentz's theory of 1895 (for which, as we will demonstrate in section 7.3, the reaction principle was not satisfied for matter alone), the French mathematician came to the conclusion that the principle of relative motion should not apply to matter alone either:

> But in the case we're considering [electrodynamics], we're not dealing with an isolated system, since we're only considering the ordinary matter, and in addition to that there is still an ether. If all material objects are carried along by a common translation, as, for example, the motion of the Earth, phenomena could be different from

---

[3] Poincaré 2008, p. 19. In my notation, $U = \overline{U}$.

[4] Ibid., p. 19.

[5] Ibid., p. 19. I have corrected the translation of "principe du movement relatif" to "principle of relative motion" (in accordance with the translator's own guideline), and have applied this correction consistently in subsequent quotations.

> those which we would observe in the absence of that translation since the ether could not be carried along by the translation. It seems like the principle of relative motion should not just apply to ordinary matter.[6]

As the quotation makes clear, in line with his conviction of a physics of principles, Poincaré was quick to observe that the above derivation of an "intimate and necessary connection"[7] between the principle of reaction and the principle of relative motion only held in the case of a conservative system. Hence, as he put it in *La Science et l'Hypothèse*, "[i]f we did not wish to change the whole of the science of mechanics," rather than abandoning its famous pillars such as the reaction principle, "we should have to introduce the ether, in order that the action which matter apparently undergoes should be counterbalanced by the re-action of matter on something."[8] The first steps in this direction were taken in the jubilee article itself. We shall see in a moment how Poincaré applied the principle of reaction to matter, including the ether, in such a way that Lorentz's theory of 1895 did justice to it.

7.2 POINCARÉ'S FORMULATION OF LORENTZ'S 1895 EQUATIONS

Let us first look at the way in which Poincaré wrote Lorentz's equations of 1895. In contrast to the Dutch physicist, he used component notation instead of vector notation. The first equations to be considered took the following form in component notation:

$$\rho\eta = -\frac{dg}{dt} + \frac{1}{4\pi}\left(\frac{d\alpha}{dz} - \frac{d\gamma}{dx}\right), \quad (7.5)$$

$$\rho\zeta = -\frac{dh}{dt} + \frac{1}{4\pi}\left(\frac{d\beta}{dx} - \frac{d\alpha}{dy}\right).^{9} \quad (7.6)$$

Observing that Poincaré expressed the components of the velocity of an electron by ξ, η and ζ,[10] these two equations can be identified as the y- and

[6] Poincaré 2008, p. 19.
[7] Ibid., p. 25.
[8] Poincaré 1952, p. 170.
[9] Poincaré 2008, p. 2. Note that Poincaré employed the same units as Lorentz.
[10] Ibid., p. 2.

z-components of the vector equation

$$\nabla \times \mathrm{h} = 4\pi \left(\rho \mathbf{v} + \frac{\partial \mathbf{d}}{\partial t} \right),$$

which corresponded to the fourth of Lorentz's microscopic version of Maxwell's equations (5.8). From this follows,

$$\begin{aligned} \mathbf{v} &= (\xi, \eta, \zeta), \\ \mathbf{h} &= (\alpha, \beta, \gamma), \\ \mathbf{d} &= (f, g, h), \end{aligned}$$

where the terms on the right-hand side of the equation represent Poincaré's notation and the single term on the left-hand side represents Lorentz's notation. It follows that

$$\rho = \sum \frac{df}{dx}$$

corresponded to the first equation of (5.8)

$$\nabla \cdot \mathbf{d} = \rho,$$

whereas

$$\frac{d\alpha}{dx} + \frac{d\beta}{dy} + \frac{d\gamma}{dz} = 0$$

was equivalent to the third equation of (5.8)

$$\nabla \cdot \mathbf{h} = 0.$$[11]

Noticing that

$$K_0 = \frac{1}{c^2},$$

we are finally in a position to reproduce Poincaré's equations

$$\begin{aligned} \frac{df}{dz} &= \frac{dh}{dx} - \frac{K_0}{4\pi} \frac{d\beta}{dt}, \\ \frac{df}{dy} &= \frac{dg}{dx} + \frac{K_0}{4\pi} \frac{d\gamma}{dt}, \end{aligned}$$

[11] Ibid., pp. 2–3.

as the y- and z-components of Lorentz's second equation of (5.8)

$$\nabla \times \mathrm{d} = -\frac{1}{4\pi c^2}\frac{\partial \mathbf{h}}{\partial t}.^{12}$$

Having shown in detail how to translate back and forth between Poincaré's component notation and Lorentz's vector notation, I would like to elaborate on the jubilee contribution by using the latter for the sake of clarity.

7.3 DERIVATION OF THE INVALIDITY OF THE PRINCIPLE OF REACTION

Poincaré began his jubilee contribution with the proof that Lorentz's theory of 1895 did not satisfy the reaction principle applied to matter:

> To start, let's briefly review the calculation by which one shows that, in the theory Lorentz, the principle of the equality of action and reaction is not correct, at least if one wishes to apply it solely to material objects.
>
> Let us find the sum of all the ponderable forces applied to all the electrons situated in the interior of a certain volume. That result, or rather its projection on the X axis, is represented by the integral
>
> $$X = \int \rho d\tau \left(\eta\gamma - \zeta\beta + \frac{4\pi f}{K_0}\right).^{13}$$

To rephrase, Poincaré wrote down the resulting force exerted on the charge within a given volume $\mathcal{T}$ as the integral of the Lorentz force exerted on the microscopic charge density ρ over the volume in question:

$$\int_{\mathcal{T}} \rho \mathbf{f}\, \mathrm{d}\tau = \int_{\mathcal{T}} \rho \mathbf{v} \times \mathbf{h} + 4\pi c^2 \mathbf{d}\, \mathrm{d}\tau.$$

By calculating this integral by parts over the entire space, he showed that the resulting force $\mathbf{F}$ that a charge distribution would experience from its self-field was given by

$$\mathbf{F} = -\frac{\mathrm{d}}{\mathrm{d}t}\int \mathbf{d} \times \mathbf{h}\, \mathrm{d}\tau. \tag{7.7}$$

By observing that this integral did not necessarily vanish, he had proven the invalidity of the reaction principle in the case of Lorentz's theory.

---

[12] Poincaré 2008, p. 3.
[13] Ibid., pp. 1–2.

7.4 THE PRINCIPLE OF REACTION APPLIED TO MATTER AND ETHER

Poincaré continued with the following remark:

> If, therefore, we call one of the material masses being considered M, and the components of its velocity V_x, V_y, and V_z, and if the principle of reaction were applicable, then we should have:
>
> $$\sum MV_x = \text{const.}, \quad \sum MV_y = \text{const.}, \quad \sum MV_z = \text{const.}$$
>
> On the contrary, we would have:
>
> $$\begin{aligned} \sum MV_x + \int d\tau(\gamma g - \beta h) &= \text{const.}, \\ \sum MV_y + \int d\tau(\alpha h - \gamma f) &= \text{const.}, \\ \sum MV_z + \int d\tau(\beta f - \alpha g) &= \text{const.}^{14} \end{aligned}$$

Let us rewrite the second equation in vector notation by introducing the mass density ϱ and its velocity field $\mathbf{V}$:

$$\int \varrho \mathbf{V}\, \mathrm{d}\tau + \int \mathbf{d} \times \mathbf{h}\, \mathrm{d}\tau = \mathbf{N}, \tag{7.8}$$

where $\mathbf{N}$ is a constant vector. Poincaré himself did not bother to mention this, but equation (7.8) followed by applying Newton's second law of motion in conjunction with equation (7.7):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \int \varrho \mathbf{V}\, \mathrm{d}\tau &= -\frac{\mathrm{d}}{\mathrm{d}t} \int \mathbf{d} \times \mathbf{h}\, \mathrm{d}\tau. \\ \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\left(\int \varrho \mathbf{V}\, \mathrm{d}\tau + \int \mathbf{d} \times \mathbf{h}\, \mathrm{d}\tau\right) &= 0 \\ \Rightarrow \int \varrho \mathbf{V}\, \mathrm{d}\tau + \int \mathbf{d} \times \mathbf{h}\, \mathrm{d}\tau &= \mathbf{N}. \end{aligned}$$

Next, Poincaré identified $c^2(\mathbf{d} \times \mathbf{h})$ as the energy flux vector that appears in the Poynting equation:

$$4\pi c^2 \int_{\mathcal{T}} \rho \mathbf{d} \cdot \mathbf{v}\, \mathrm{d}\tau + \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathcal{T}} \frac{1}{8\pi} \mathbf{h}^2 + 2\pi c^2 \mathbf{d}^2\, \mathrm{d}\tau + \int_{\partial \mathcal{T}} c^2 (\mathbf{d} \times \mathbf{h}) \cdot \mathrm{d}\boldsymbol{\sigma} = 0 \tag{7.9}$$

---

[14] Ibid., pp. 3–4.

According to Poincaré, the first term represented the work done by the ether on the charges within the volume $\mathcal{T}$ during a unit of time dt. To see this, note that the work done by the fields of the ether on a charge q was given by

$$(q\mathbf{f})\cdot\mathrm{d}\mathbf{l} = q\left(\mathbf{v}\times\mathbf{h} + 4\pi c^2\mathbf{d}\right)\cdot\mathbf{v}\mathrm{d}t = 4\pi c^2 q\mathbf{d}\cdot\mathbf{v}\,\mathrm{d}t.$$

Due to the fact that the charge in a volume element dτ was ρdτ, the rate at which work was done on the charges in a volume $\mathcal{T}$ was therefore

$$\frac{\mathrm{d}W}{\mathrm{d}t} = 4\pi c^2\int_{\mathcal{T}}\rho\mathbf{d}\cdot\mathbf{v}\,\mathrm{d}\tau.$$

The second term in (7.9) represented the time derivative of the total energy stored in the ether within a volume $\mathcal{T}$, while the third term comprised the rate at which energy is dissipated through the surface $\partial\mathcal{T}$. Against this background, the Poynting equation stated that the decrease of energy stored in the ether was equal to the work performed on the charges plus the energy dissipated through the surface.

Defining the electromagnetic energy density

$$J = \frac{1}{8\pi}\mathbf{h}^2 + 2\pi c^2\mathbf{d}^2$$

and substituting this expression into the Poynting equation, Poincaré obtained

$$\int_{\mathcal{T}}\frac{\partial J}{\partial t}\,\mathrm{d}\tau = -\int_{\partial\mathcal{T}}c^2(\mathbf{d}\times\mathbf{h})\cdot\mathrm{d}\boldsymbol{\sigma} - 4\pi c^2\int_{\mathcal{T}}\rho\mathbf{d}\cdot\mathbf{v}\,\mathrm{d}\tau. \tag{7.10}$$

Based on this equation, he interpreted electromagnetic energy as a fictitious fluid:

> We can consider the electromagnetic energy as a fictional fluid of which the density is $K_oJ\,[\frac{J}{c^2}]$ and which travels through space in conformance with Poynting's law. We just need to realize that the fluid is not indestructible, and in the element of volume $d\tau$, during one unit of time, a quantity $\frac{4\pi}{K_0}\rho d\tau\sum f\xi$ $[4\pi c^2\rho\mathbf{d}\cdot\mathbf{v}\,\mathrm{d}\tau]$ is destroyed [...] It is that which prevents us from considering our fictional fluid as a sort of "real" fluid.[15]

[15] Poincaré 2008, p. 5.

To achieve the described conformity, he further noted that the velocity of the fluid **U** had to be defined in such a way that the product of its density $\frac{J}{c^2}$ and velocity satisfied

$$\frac{J}{c^2}\mathbf{U} = \mathbf{d} \times \mathbf{h}. \tag{7.11}$$

Following Poincaré's idea of electromagnetic energy as a fictitious fluid, we write equation (7.10) in differential form, divide it by c^2, shift the term of the energy flux to the left-hand side and insert (7.11) into the result:

$$\frac{\partial}{\partial t}\left(\frac{J}{c^2}\right) + \nabla\cdot\left(\frac{J}{c^2}\mathbf{U}\right) = -4\pi\rho\mathbf{d}\cdot\mathbf{v}. \tag{7.12}$$

The left-hand side of the equation now corresponds to the continuity equation for a fluid with density $\frac{J}{c^2}$ and velocity **U**. In contrast to the latter, however, the right-hand side of (7.12) is equal to $-4\pi\rho\mathbf{d}\cdot\mathbf{v}$ instead of 0. Assuming that Poincaré inadvertently forgot to divide the right-hand side by c^2, this means that the fluid would decrease by a quantity of $4\pi\rho\mathbf{d}\cdot\mathbf{v}\mathrm{d}\tau$ in a volume element $\mathrm{d}\tau$ during a time interval $\mathrm{d}t$.

Substituting equation (7.11) into (7.8), Poincaré arrived at

$$\int \varrho\mathbf{V}\,\mathrm{d}\tau + \int \frac{J}{c^2}\mathbf{U}\,\mathrm{d}\tau = \mathbf{N},$$

which he interpreted as the conservation of total momentum of ordinary matter and fictitious fluid. That is to say, he had shown how to reformulate the reaction principle so that it was applicable to matter including the ether (in the sense that the state of the ether simply corresponded to the values of the electromagnetic fields **d** and **h** at each point in space).

7.5 THE RECOIL OF A HERTZIAN OSCILLATOR

The first consequence that Poincaré drew from this view was that, "since the electromagnetic energy behaves as a fluid which has inertia, [...] if any sort of device produces electromagnetic energy and radiates it in a particular direction, that device must *recoil* just as a cannon does when it fires a projectile."[16] This is the case of a Hertzian oscillator fixed at the focus of a parabolic reflector and emitting a constant rate of radiation. For example, if the device had a mass m_H of 1 kg and emitted 3,000,000 J in a single

[16] Ibid., p. 8.

direction at the speed of light, Poincaré predicted that the recoil speed v_H would be 1 cm/s.[17] Let us reconstruct the necessary calculations based on the above. We begin by assuming that the beam emitted by the oscillator moves in the x-direction. Applying the principle of conservation of total momentum of matter and ethereal fluid, we obtain

$$m_H \mathbf{v}_H + \frac{E}{c^2}\mathbf{c} = \mathbf{0},$$

where $\mathbf{v}_H$ denotes the recoil velocity of the emitter and E the radiated energy. Solving the equation for $\mathbf{v}_H$, we get:

$$\mathbf{v}_H = -\frac{E}{m_H c}\,\hat{\mathbf{x}} = -\frac{3{,}000{,}000\,\mathrm{J}}{1\,\mathrm{kg}\cdot 300{,}000\,\frac{\mathrm{km}}{\mathrm{s}}}\,\hat{\mathbf{x}} = -1\,\frac{\mathrm{cm}}{\mathrm{s}}\,\hat{\mathbf{x}},$$

where $\hat{\mathbf{x}}$ denotes a unit vector in the x-direction. Due to the magnitude of the recoil, Poincaré ended this preliminary discussion of the Hertzian oscillator with the conclusion that

> such a weak force couldn't be detected in our experience. But we can imagine that, impossibly, we have measuring devices so sensitive that we can measure such forces. We could then demonstrate that the principle of reaction is applicable not just to matter; and that would be confirmation of the theory of Lorentz.[18]

Having explained how Poincaré adapted the reaction principle to reconcile Lorentz's 1895 theory, let us take a look at the remaining task of accounting for the principle of relative motion.

7.6 THE APPARENT CONTRADICTION

According to the derivation in section 7.1, the reaction principle followed from the principle of relative motion when applied to an isolated system. Since Poincaré also had argued that the principle of reaction was only applicable to ordinary matter including the ether, it followed that the principle of relative motion should not be restricted to ordinary matter either. However, this conclusion seemed to lead to a paradox:

---

[17] Poincaré 2008, p. 5.

[18] Ibid., p. 9.

> [E]xperiments have been carried out to detect the motion of the Earth [through the ether] [...] Those experiments, it is true, have produced negative results [...] and the theory of Lorentz explains that negative result. It appears that the principle of relative motion, which is not true *a priori*, is verified *a posteriori* and that the principle of reaction should follow. Yet the principle of reaction does not hold; how can that be?"[19]

Here Poincaré drew attention to the fact that all experiments intended to detect the motion of the earth through the ether, such as those of Michelson and Morley, had produced negative results, implying they ultimately confirmed the validity of the principle of relative motion only for matter. Moreover, according to Poincaré, Lorentz was not only motivated by the challenge of finding a theoretical explanation for the null results, but he actually succeeded. However, the Dutchman did not take into account the paradoxical situation that arose between a void reaction principle (for matter alone) on the one hand and the experimental verification of the principle of relative motion (for matter alone) on the other. In particular, he did not deduce the phenomenon of unpaired material recoil, as in a Hertzian oscillator, as a consequence of his theory. The resolution of this apparent contradiction was the task that Poincaré set himself in the further course of his jubilee article.

7.7 RESOLVING THE APPARENT CONTRADICTION

He first moderated the alleged contradiction somewhat by clarifying that "the principle of relative motion has been verified only imperfectly, as shown by the theory of Lorentz."[20] As we have seen above, Lorentz's 1895 theory did not fully comply with the principle of relative motion applied to matter alone, and therefore this principle did not qualify as a convention in the context of Lorentz's theory. This meant that Poincaré had leeway to attempt to apply the principle of relative motion to both matter and the ether. The qualification was accompanied by a detailed sketch indicating the direction of Poincaré's solution to the paradox, the centerpiece of which was the compensation of multiple effects:

1. That compensation does not take place unless we neglect u^2 [...]

[19] Ibid., pp. 19–20.
[20] Ibid., p. 20.

2. For the compensation to work, we must relate the phenomena not to to true time t, but to a certain *local time* t' [...]
3. In relative motion, the propagation of the apparent energy follows the same laws as the real energy in absolute motion, but the apparent energy is not exactly equal to the corresponding real energy.
4. In relative motion, the bodies emitting the electromagnetic energy are subject to an apparent complementary force which does not exist in absolute motion.[21]

While Poincaré's first point merely emphasized that the solution by compensation of multiple effects did not work beyond the stated approximation, one can hardly overestimate the importance of the second point, which corresponded to a physical reinterpretation of Lorentz's theory of corresponding states, providing a new understanding of local time. The third and fourth points clarified how he wanted to make use of the fact that the principle of relative motion for matter alone was only approximately true within Lorentz's theory. We will return to this topic in detail a little later.

7.7.1 Physical Interpretation of Local Time

Let me first describe how Poincaré reinterpretated Lorentz's theory of corresponding states. As we have seen in detail above, this theory allowed Lorentz to infer physical consequences for a system of bodies $\overline{S}$ in uniform rectilinear motion by comparing it with a purely fictitious system S' brought to rest. In particular, Lorentz understood the vector fields $\mathbf{D}'$, $\mathbf{F}'$ and $\mathbf{H}'$ of S' as purely mathematical transformations of the fields $\overline{\mathbf{D}}$, $\overline{\mathbf{F}}$ and $\overline{\mathbf{H}}$ of $\overline{S}$. In 1900, however, Poincaré understood that the corresponding states were the 'apparent' states of the moving system, i.e. that the coordinates of the primed quantities represented the quantities measured by an observer of the moving system. In particular, Poincaré observed that Lorentz's local time $t' = t - \frac{ux}{c^2}$ corresponded to the readings of moving clocks synchronized in the following way:

> I suppose that observers placed in different points set their watches by means of optical signals; that they try to correct these signals by the transmission time, but that, ignoring their translational motion

[21] Poincaré 2008, p. 20.

> and thus believing that the signals travel at the same speed in both directions, they content themselves with crossing the observations by sending one signal from A to B, then another from B to A.[22]

As Darrigol has pointed out, "Poincaré only made this remark *en passant*, gave no proof, and did as if it had already been on Lorentz's mind."[23] Darrigol's version of the missing proof is as follows: Suppose two observers A and B move through the ether with velocity $\mathbf{u}$ such that $\mathbf{u}$ is parallel to AB and in the direction from A to B. Let us now assume that the observer A sends a light signal in the direction of B. When B receives the signal from A, he immediately resets his clock to zero and sends a light signal back to A. When A receives this signal, he reads his clock and notes that the time τ has elapsed since he sent his signal to B. "[B]elieving that the signals travel at the same speed in both directions," A sets his clock to $\frac{\tau}{2}$. "By doing so," Darrigol points out, "he commits an error $\frac{\tau}{2} - t_-$, where t_- is the time that light really takes to travel from B to A."[24] Applying the Galilean rule for the addition of velocities, Darrigol notes that

$$t_- = \frac{|AB|}{c+u}, \qquad t_+ = \frac{|AB|}{c-u}, \tag{7.13}$$

where t_+ is the time it takes for a light signal to travel from A to B. Thus, $\tau = t_+ + t_-$ and $t_- \leq \frac{\tau}{2} \leq t_+$. From this it follows that

$$\begin{aligned} \frac{\tau}{2} - t_- &= \frac{t_+ + t_-}{2} - t_- = \frac{t_+ - t_-}{2} \\ &= |AB|\frac{(c+u)-(c-u)}{c^2-u^2} = \frac{u|AB|}{c^2}\frac{1}{1-\frac{u^2}{c^2}}. \end{aligned} \tag{7.14}$$

Darrigol now remarks that, to first-order in $\frac{u}{c}$, the committed error is

$$\frac{\tau}{2} - t_- = \frac{u|AB|}{c^2}. \tag{7.15}$$

On this basis, he finally comes to the conclusion that "[a]t a given instant of the true time, the times indicated by the two clocks differ by $\frac{uAB}{c^2}$, in conformity [to first order in $\frac{u}{c}$] with Lorentz's expression of the local time."[25] In equation (5.46), using barred coordinates, I expressed the latter as $t' = \bar{t} - \frac{u\bar{x}}{c^2}$.

[22] Ibid., p. 20.
[23] Darrigol 2006, p. 14.
[24] Ibid., p. 14.
[25] Ibid., p. 14.

7.7.2 Preliminary Analysis of a Hertzian Oscillator

Let us now see how Poincaré applied this reinterpretation in the case of a Hertzian oscillator moving uniformly through space at velocity **v** relative to the ether while emitting a constant rate of light. Before going into the details of the calculations, he gave a preliminary comparative analysis of two cases in which a coordinate system was adapted to the Hertzian oscillator. In the first case, the Hertzian oscillator was at rest ($\mathbf{v} = \mathbf{0}$), in the second it moved uniformly and rectilinearly relative to the ether ($\mathbf{v} \neq \mathbf{0}$). The aim of this preliminary analysis was twofold. First, Poincaré wanted to clarify the "paradox" or "apparent contradiction" between a void reaction principle for matter alone and valid principles of energy conservation and relative motion in the case of the Hertzian oscillator. Secondly, in the light of this clarification, he wanted to indicate how his detailed analysis was meant to resolve the paradox.

In analyzing the first case, Poincaré referred to his analysis of momentum mentioned above:

> Initially at rest, the exciter emits som energy along the x axis, and that energy is exactly equal to that which is expended in the exciter. As we have seen, the device *recoils* and takes on a certain velocity.[26]

In this case, the recoil force did no work on the oscillator, simply because it was at rest before the emission of energy due to radiation. This implied that all the energy expended in the oscillator was spent on the emission. Elaborating the second case, Poincaré noted that

> [i]f we relate everything to the moving axes which are linked to the exciter, the apparent phenomena should be, except for the reservations mentioned above, the same as if the exciter were at rest; it will therefore radiate an *apparent* quantity of energy which is equal to the energy expended in the exciter.[27]

While the reservations related to the above list, Poincaré made another, less obvious assumption in this quotation: there would be no difference between the apparent and the real energy expended in the oscillator. This assumption was not addressed in the contribution. However, Darrigol has argued that "[f]or an observer moving with the emitter, the energy spent

[26] Poincaré 2008, p. 21.
[27] Ibid., p. 21.

by the emitter should be the same as for the observer at rest, since it is measured by permanent alterations of the energy source."[28] Thus (via point 3), the apparent quantity of energy would follow the same laws as in the first case, where all the energy expended in the oscillator was spent on the emission. In other words, the apparent radiated quantity of energy E' had to be equal to the apparent energy D' expended in the oscillator, which in turn was equal to the (real) energy D expended in the oscillator. However, if we consider the situation from the perspective of an observer at rest relative to the ether, Poincaré argued that the (real) force of recoil would do (real) work on the oscillator:

> On the other hand, it receives an impulse from the recoil, and since it is not stopped, but already has a nonzero velocity, that impulse does work on the device and its kinetic energy increases.[29]

It followed that the apparent quantity of radiated energy E' must differ from the real radiated energy E. For assuming that these two terms were equal, the "increase in kinetic energy of the device would be obtained without any corresponding consumption."[30] And this would contradict the principle of the conservation of energy. To see this, note that Poincaré's remark presupposes that the emitter moved in the direction opposite to that of the radiated light. Moreover, let us recall that $E' = D' = D$. From the perspective of an observer at rest, the oscillator would gain some (real) kinetic energy T due to the work done on it by the recoil force. To be consistent with the principle of the conservation of energy, the total energy spent in the oscillator would therefore have to exceed the energy actually radiated ($D = E+T$ or $E' = E+T$). Poincaré therefore felt justified in concluding that "the phenomena in relative motion will not be exactly the same as those in absolute motion."[31]

7.7.3 *Comprehensive Analysis of a Hertzian Oscillator*

Having explained Poincaré's design for his solution to the apparent contradiction, let us follow him and work out the details. In his comprehensive analysis, he did not assume that the moving system was adapted to the oscillator. More precisely, he introduced moving coordinates traveling at a

[28] Darrigol 1995, p. 26.
[29] Poincaré 2008, p. 21.
[30] Ibid., p. 21.
[31] Ibid., p. 21. In this context, absolute motion referred to motion relative to the ether.

constant velocity **u**, which was not necessarily equal to the velocity **v** of the oscillator. He also assumed that all three velocities (**u**, **v** and the velocity of the emitted radiation) were parallel to the x-axis of the ether frame. Presuming that the emitted radiation had the form of a linearly polarized plane wave, he wrote down the following equations in component form, adapted to the ether frame:

$$f = h = \alpha = \beta = 0, \tag{7.16}$$

$$V\frac{d\gamma}{dx} + \frac{d\gamma}{dt} = 0, \tag{7.17}$$

$$4\pi\frac{dg}{dt} = -\frac{d\gamma}{dx}, \tag{7.18}$$

$$-\frac{1}{4\pi V^2}\frac{d\gamma}{dt} = \frac{dg}{dx}. \tag{7.19}$$

where V denoted the speed of light. Although Poincaré himself did not say this explicitly, we must interpret these equations on a macroscopic level in order to apply Lorentz's theory of corresponding states. The equations (7.16) and (7.17) represented the polarized plane wave. In vector notation, they take the form

$$\mathbf{D} = D_y\hat{\mathbf{y}}, \qquad \mathbf{H} = H_z\hat{\mathbf{z}}, \tag{7.20}$$

$$(c\hat{\mathbf{z}} \cdot \nabla)\mathbf{H} + \frac{\partial \mathbf{H}}{\partial t} = \mathbf{0}.$$

The equations (7.18) and (7.19) were obtained by simplifying Lorentz's second and fourth equations of (5.45) using (7.16) and the ether relation $\mathbf{F} = 4\pi c^2\mathbf{D}$ (which follows from $\epsilon = 1$). If we rewrite (7.18) and (7.19) in accordance with (7.20), we get

$$4\pi\frac{\partial D_y}{\partial t} = -\frac{\partial H_z}{\partial x},$$

$$-\frac{1}{4\pi c^2}\frac{\partial H_z}{\partial t} = \frac{\partial D_y}{\partial x}.$$

From these equations Poincaré obtained

$$H_z = 4\pi c D_y. \tag{7.21}$$

Inserting (7.21) into the equation for the (macroscopic) energy density, the result was

$$J_M = \frac{H_z^2}{8\pi} + 2\pi c^2 D_y^2 = 4\pi c^2 D_y^2. \tag{7.22}$$

He then continued by deriving the corresponding "equations for the apparent motion [...] relative to the moving axes."[32] Since Poincaré's representation of apparent motion in terms of primed coordinates obviated the need for barred coordinates, let me rewrite the primed variables of Lorentz's theory of corresponding states in terms of unprimed variables before we continue our exposition.

If we rewrite the variables from unprimed to barred using the Galilean transformation (5.12) and from barred to primed using (5.46), to first order in $\frac{u}{c}$ we obtain

$$\begin{aligned} x' &= x - ut, & y' &= y, \\ z' &= z, & t' &= t - \frac{ux}{c^2}. \end{aligned} \tag{7.23}$$

Furthermore, due to Lorentz's tacit assumption that

$$\overline{\mathbf{D}}(\bar{\mathbf{r}}, \bar{t}) = \mathbf{D}(\mathbf{r}, t)$$

and

$$\overline{\mathbf{H}}(\bar{\mathbf{r}}, \bar{t}) = \mathbf{H}(\mathbf{r}, t),$$

where

$$(\bar{\mathbf{r}}, \bar{t}) = (\mathbf{r} - t\mathbf{u}, t),$$

$\mathbf{D}'$ and $\mathbf{H}'$ of (5.48) can be expressed in terms of unprimed variables:

$$\mathbf{D}' = \mathbf{D} + \frac{1}{4\pi c^2}\,\mathbf{u} \times \mathbf{H}, \quad \mathbf{H}' = \mathbf{H} - 4\pi\,\mathbf{u} \times \mathbf{D}. \tag{7.24}$$

Since $\epsilon = 1$ in the ether, Lorentz's primed equations (5.49) finally have the form

$$\begin{aligned} \nabla' \cdot \mathbf{D}' &= 0, & \nabla' \times \mathbf{D}' &= -\frac{1}{4\pi c^2}\frac{\partial \mathbf{H}'}{\partial t}, \\ \nabla' \cdot \mathbf{H}' &= 0, & \nabla' \times \mathbf{H}' &= 4\pi \frac{\partial \mathbf{D}'}{\partial t'}. \end{aligned} \tag{7.25}$$

[32] Poincaré 2008, p. 21.

Let us return to Poincaré's account of apparent motion and explain it using these equations.

First, he wrote equations for the apparent electric and magnetic fields:

$$D'_y = D_y - \frac{u}{4\pi c^2} H_z, \quad H'_z = H_z - 4\pi u D_y.$$

Both equations result directly if (7.20) is inserted into (7.24). He then argued that, by neglecting u^2, but not uv, it was possible to express the apparent energy density in terms of the real fields:

$$J'_M = \frac{H'^2_z}{8\pi} + 2\pi c^2 D'^2_y = \frac{H_z^2}{8\pi} + 2\pi c^2 D_y^2 - 2u D_y H_z \tag{7.26}$$

Let us restate his argument in terms of neglecting second-order terms in $\frac{u}{c}$:

$$\begin{aligned} J'_M &= \frac{H'^2_z}{8\pi} + 2\pi c^2 D'^2_y \\ &= \frac{1}{8\pi}\left(H_z^2 - 8\pi u H_z D_y + 16\pi^2 u^2 D_y^2\right) + 2\pi c^2 \left(D_y^2 - \frac{u D_y H_z}{2\pi c^2} + \frac{u^2 H_z^2}{16\pi^2 c^4}\right) \\ &= \frac{H_z^2}{8\pi} - 2u H_z D_y + 2\pi c^2 D_y^2 \left(1 + \left(\frac{u}{c}\right)^2\right) + \frac{H_z^2}{8\pi}\left(\frac{u}{c}\right)^2 \\ &\approx \frac{H_z^2}{8\pi} + 2\pi c^2 D_y^2 - 2u D_y H_z \end{aligned}$$

By applying (7.21) he finally obtained

$$J'_M = 4\pi c^2 D_y^2 - 2u D_y H_z = 4\pi c^2 D_y^2 \left(1 - \frac{2u}{c}\right). \tag{7.27}$$

Secondly, he wrote down the apparent equations

$$4\pi \frac{\partial D'_y}{\partial t'} = -\frac{\partial H'_z}{\partial x'}, \quad -\frac{1}{4\pi c^2}\frac{\partial H'_z}{\partial t'} = \frac{\partial D'_y}{\partial x'}. \tag{7.28}$$

These equations result from Lorentz's theory if we insert (7.20) into the second and fourth equations of (7.25) respectively.

Thirdly, by denoting the oscillation period of the plane wave by T, Poincaré determined the real wavelength λ by calculating the real distance between two successive amplitudes. More precisely, he calculated the distance between the head and tail of the perturbation at a certain absolute time t.

Assuming that the head left the Hertzian oscillator at time 0 at the origin of S, it would be at ct at time t. Due to the definition of the period T, the tail would leave the device at time T. However, as the oscillator would move a distance vT during the period, at time t the tail would be at

$$x = vT + c(t - T). \tag{7.29}$$

Therefore, Poincaré argued, the real wavelength λ was

$$\lambda = ct - (vT + c(t - T)) = (c - v)T. \tag{7.30}$$

He then determined the apparent wavelength λ' by calculating the apparent distance between the pertubation's head and tail at local time t', which corresponded to the absolute time t and the absolute position x just mentioned. In apparent coordinates, the head exited at local time 0 and local position 0. From the theory of corresponding states (and in particular from the equations (7.28)) it followed that the apparent speed of light was equal to its real speed c. Therefore, at local time t' the head would be in the local position ct'. Let us now explain Poincaré's calculation of the tail's local position x' at local time t'. In absolute coordinates, as we have established, the tail left position vT at time T. Transforming these coordinates into apparent coordinates using (7.23), Poincaré obtained that the corresponding local event occurred at local time $T - \frac{uvT}{c^2} = T\left(1 - \frac{uv}{c^2}\right)$ and local position $(v - u)T$. Expressing the local time t' by the corresponding absolute time t and absolute position x, he got

$$t' = t - \frac{ux}{c^2}.$$

Isolating t in this expression and substituting the result into equation (7.29), he arrived at

$$x = vT + c\left(t' + \frac{ux}{c^2} - T\right).$$

Isolating x, the equation yielded

$$x = \frac{vT + c(t' - T)}{1 - \frac{u}{c}}. \tag{7.31}$$

Poincaré then argued that x could be written as

$$x = (vT + c(t' - T))\left(1 + \frac{u}{c}\right) \tag{7.32}$$

by neglecting u^2. To see this, we write the Taylor expansion for $\frac{1}{1-\frac{u}{c}}$ to first order in $\frac{u}{c}$:

$$\frac{1}{1-\frac{u}{c}} = 1 + \left(\frac{u}{c}\right) + \left(\frac{u}{c}\right)^2 + \left(\frac{u}{c}\right)^3 + \ldots \approx 1 + \frac{u}{c}. \tag{7.33}$$

If (7.33) is inserted into (7.31), the expression (7.32) for x now follows.

Poincaré went on by noticing that the local position x' could be written as

$$x' = x - ut', \tag{7.34}$$

which followed from (7.23) if the second-order term in $\frac{u}{c}$ was neglected:

$$x' = x - ut = x - u\left(t' + \frac{ux}{c^2}\right) = x - ut' - x\left(\frac{u}{c}\right)^2 \approx x - ut'.$$

Substituting (7.32) and

$$ct'\left(1 + \frac{u}{c}\right) - ut' = ct'$$

into (7.34) then resulted in

$$x' = (vT + c(t' - T))\left(1 + \frac{u}{c}\right) - ut' = (vT - cT)\left(1 + \frac{u}{c}\right) + ct'. \tag{7.35}$$

From equations (7.30) and (7.35) it then followed that the apparent wavelength λ' could be written as

$$\lambda' = ct' - x' = (c - v)T\left(1 + \frac{u}{c}\right) = \lambda\left(1 + \frac{u}{c}\right).$$

The immediate reason why Poincaré obtained expressions for the real wavelength λ and its apparent counterpart λ' was the need to calculate what he called the total real energy and the (total) apparent energy per unit area:

> The total real energy (per unit cross-section) is therefore
>
> $$\left(\frac{\gamma^2}{8\pi} + 2\pi V^2 g^2\right) L = 4\pi V^2 g^2 L,$$
>
> and the apparent energy
>
> $$\left(\frac{\gamma'^2}{8\pi} + 2\pi V^2 g'^2\right) L' = 4\pi V^2 g^2 L\left(1 - \frac{2v}{V}\right)\left(1 + \frac{v}{V}\right) = 4\pi V^2 g^2 L\left(1 - \frac{v}{V}\right).$$ [33]

In our notation, V corresponds to c, L to λ and v to u. Let me denote the total real energy by E and the total apparent energy by E' and reconcile the above equations with our notation:

$$E = \left(\frac{{H_z}^2}{8\pi} + 2\pi c^2 {D_y}^2\right)\lambda = 4\pi c^2 D_y^2 \lambda,$$
$$E' = \left(\frac{H_z'^2}{8\pi} + 2\pi c^2 {D_y'}^2\right)\lambda' = 4\pi c^2 D_y^2 \lambda \left(1 - \frac{u}{c}\right). \tag{7.36}$$

Stated otherwise, the total real energy per unit area was calculated as the product of the energy density J_M and the wavelength λ

$$E = J_M \lambda,$$

while the corresponding total apparent energy was the product of the apparent energy density J_M' and the apparent wavelength λ'

$$E' = J_M' \lambda'.$$

The naming and the equation therefore suggest that by "the total real energy" Poincaré meant the energy that flowed through a unit area over a complete wave cycle. To recognize this, in the expression for E we rewrite the wavelength λ as the product of its speed c and period T:

$$E = J_M c T. \tag{7.37}$$

Secondly, we calculate the magnitude of Poynting's energy flux vector using (7.9), (7.22) and (7.27):

$$S = |c^2(\mathbf{h} \times \mathbf{d})| = c^2 H_z D_y = 4\pi c^3 D_y^2 = J_M c, \tag{7.38}$$

Stated in words, the magnitude of the Poynting vector was the product of the wave's energy density and its speed. If we insert (7.38) into (7.37), we finally get

$$E = ST.$$

This means that Poincaré's total energy was equal to the energy flux per unit area times the period, which supports the above interpretation that by

[33] Poincaré 2008, p. 23.

total energy he understood the energy flowing through a unit area over a complete wave cycle. However, Poincaré's expression for E would only be correct if S had a constant value. To obtain the exact value of E, he would have to integrate the magnitude of the Poynting vector over a complete wave cycle. Another very plausible possibility is that he already understood the value of D_y to be an average or constant, from which it follows that S would also be an average or constant, rather than a fluctuating cosine-squared term. Since he was only interested in the relationship between the apparent and the real total energy, he obtained the correct result for the latter by comparing E and E' in equation (7.36):

$$E' = E\left(1 - \frac{u}{c}\right). \tag{7.39}$$

This relationship was necessary to explain the relationship between the real and apparent work of recoil, to which we now turn.

With the help of (7.39) Poincaré remarked to begin with that "if $J\mathrm{d}t$ represents the real radiated energy during time $\mathrm{d}t$, then $J\mathrm{d}t\left(1 - \frac{v}{V}\right)$ will represent the apparent energy."[34] To recognize this, we first note that the quantity J in this quotation does not correspond to the energy density J_M, but to the magnitude of the energy flux vector $S = J_M c$. If we represent "the real radiated energy during time dt" in our notation, we get

$$S\mathrm{d}t = J_M c\mathrm{d}t.$$

It follows that the corresponding definition for the apparent energy would be

$$S'\mathrm{d}t' = J'_M c\mathrm{d}t'.$$

Second, we note that "in the case of *light*, the wavelength is so short ($\sim 5 \times 10^{-7}$m), and the period so brief ($\sim 10^{-15}$s), that any macroscopic measurement will encompass many cycles,"[35] so that any measurement of $S\mathrm{d}t$ and $S'\mathrm{d}t'$ would amount to a certain product of E and E' respectively. With the help of (7.39) it follows that

$$S'\mathrm{d}t' = S\left(1 - \frac{u}{c}\right)\mathrm{d}t$$

as claimed by Poincaré.

[34] Poincaré 2008, p. 23.
[35] Griffiths 1998, p. 381.

Let us see how Poincaré used this result to explain the real and the apparent recoil of the oscillator. He began by stating that the force of recoil, multiplied by $\mathrm{d}t$, was equal in magnitude (but opposite in direction) to the increase in momentum of the fictitious fluid. This increase in momentum during $\mathrm{d}t$ (per unit area), Poincaré argued, was in turn equal to

$$\frac{S\mathrm{d}t}{c^2}\mathbf{c} = \frac{S\mathrm{d}t}{c}\hat{\mathbf{x}},$$

because $\frac{S\mathrm{d}t}{c^2}$ was the increase in mass of the fictitious fluid (per unit area) and $\mathbf{c} = c\hat{\mathbf{x}}$ was its veolcity. That is to say, applying the reaction principle for matter including the ether,

$$-\frac{S}{c}\hat{\mathbf{x}}$$

was the pressure of recoil exerted by the radiation on the oscillator. The work of recoil $\mathrm{d}W$ (per unit area) was therefore

$$\mathrm{d}W = \left(-\frac{S}{c}\hat{\mathbf{x}}\right)\cdot \mathrm{d}\mathbf{l} = -\frac{vS\mathrm{d}t}{c},$$

as the oscillator would move $\mathrm{d}\mathbf{l} = \mathbf{v}\mathrm{d}t$ during the time $\mathrm{d}t$.

For the apparent motion, Poincaré had to replace the real velocity of the oscillator $\mathbf{v}$ by the apparent $\mathbf{v}' = \mathbf{v} - \mathbf{u}$, and the real radiated energy $S\mathrm{d}t$ by the apparent $S'\mathrm{d}t' = S\big(1 - \frac{u}{c}\big)\mathrm{d}t$. Neglecting second-order terms in $\frac{u}{c}$, the apparent work of recoil was therefore

$$\mathrm{d}W' = -\frac{v-u}{c}S\left(1 - \frac{u}{c}\right)\mathrm{d}t = S\mathrm{d}t\left(-\frac{v}{c} + \frac{u}{c} + \frac{uv}{c^2}\right).$$

Assuming that the energy expended in the oscillator was $D\mathrm{d}t$, Poincaré was finally able to write down equations for the conservation of energy in real and apparent motion. For the real motion he obtained

$$S\mathrm{d}t - D\mathrm{d}t - \frac{vS\mathrm{d}t}{c} = 0 \quad \text{or} \quad S - D - \frac{vS}{c} = 0, \tag{7.40}$$

where "the first term represents the radiated energy, the second the expended energy, and the third the work of recoil."[36] In other words, the energy lost in the oscillator, together with the decrease in its kinetic energy from recoil, was converted into radiated energy.

[36] Poincaré 2008, p. 24.

For the apparent motion, Poincaré repeated the assumption that the energy expended in the oscillator was the same as for the real motion. Finally, he introduced the apparent complementary force:

> [I]n apparent motion, we must account for the apparent complementary force of which I spoke above (4.). That complementary force is equal to $-\frac{vJ}{V^2}$ and the work it does, neglecting v^2, is $-\frac{vv'}{V^2}Jd\tau$.[37]

In our notation, this apparent force is equal to

$$-\frac{uS}{c^2}. \tag{7.41}$$

Let us reconstruct the calculation of the work done by this force on the emitter. Since the emitter's relative speed is $v' = v - u$, it follows that the apparent work is

$$v'\mathrm{d}t'\left(-\frac{uS}{c^2}\right) = (v-u)\mathrm{d}t\left(-\frac{uS}{c^2}\right) = -\frac{vuS}{c^2}\mathrm{d}t + \frac{u^2S}{c^2}\mathrm{d}t, \tag{7.42}$$

where we have used that $\mathrm{d}t' = \mathrm{d}t$ to first order in $\frac{u}{c}$. If we also neglect the other second-order term in $\frac{u}{c}$, we get the value from the quote:

$$-\frac{vuS}{c^2}\mathrm{d}t. \tag{7.43}$$

With this in mind, it becomes clear why it makes sense to call this force an apparent force: it only performs work in apparent motion. To recognize this, observe that its contribution vanishes when $u = 0$.

Let us now look at the consequences of the introduction of this force for the apparent energy balance

$$S' - D' - \frac{v'S'}{c} - \frac{vuS}{c^2}, \tag{7.44}$$

where "[t]he first term represents the apparent radiated energy, the second term the energy expended, the third term the apparent work of recoil, and the fourth the work of the apparent complementary force."[38] If we express the four terms in absolute coordinates, we see that they sum to 0:

$$S\left(1-\frac{u}{c}\right) - D + S\left(-\frac{v}{c} + \frac{u}{c} + \frac{uv}{c^2}\right) - \frac{vuS}{c^2} = 0. \tag{7.45}$$

[37] Poincaré 2008, p. 24.
[38] Ibid., p. 25.

Poincaré went on to observe a correspondence between this equation and the equivalent equation (7.40) in absolute motion, implying that the two equations expressed the same physical system from two different perspectives: absolute and relative. In particular, equation (7.45) reduced to (7.40) for $u = 0$ (i.e. when the relative coordinates were at rest relative to the ether). But secondly, it is important to note that (7.45) also reduced to (7.40) for $u = v$, implying that the relative coordinates were adapted to the emitter. It followed that for an observer at rest relative to the emitter, the "radiated energy" appeared to have the value $S\left(1 - \frac{u}{c}\right) = D$ (per unit time). However, as the relationship between the equations (7.40) and (7.45) makes clear, this was only apparent. The real radiated energy was equal to S (per unit time). Therefore, Poincaré had finally "resolved the apparent contradiction."

On the one hand, this solution had its price. Unlike in the case of $u = v$, there was a difference between absolute and relative motions for $u \neq v$, which could in principle be detected empirically. It followed that the principle of reaction for matter alone was invalid in Lorentz's theory. On the other hand, Poincaré had succeeded in reformulating both the principle of reaction and the principle of relative motion in such a way that they were valid for matter including the ether:

> Thus, according to the theory of Lorentz, the principle of reaction must not apply solely to matter; the principle of relative motion also must not apply solely to matter. It's important to note that there is a intimate and necessary connection between these two facts.
>
> *It is therefore sufficient to establish experimentally either of the two, from which the other will be established, ipso facto.* It would no doubt be less difficult to demonstrate the second; but even that is nearly impossible since, for example, Liénard has calculated that with a machine of 100 kW, the apparent complementary force would be no more than $\frac{1}{600}$ dyne.[39]

Since both principles applied to matter including the ether, it followed that their application to matter alone was empirically void. The failure of the principle of relative motion for matter alone was most clearly demonstrated by the circumstance that the "apparent complementary force" depended on its relative velocity with respect to the stationary ether.

[39] Ibid., pp. 25–6.

One question that remains to be answered is how Poincaré evaluated his maneuvers to resolve the apparent contradiction in Lorentz's theory from the perspective of his physics of principles. Were they fruitful corrections that brought electrodynamics back into line with the list of general principles by clarifying the proper role of the ether within Lorentz's theory? Or should they rather be seen as fleeting ad hoc safeguards for the principles of reaction and relative motion, heralding an inevitable forthcoming crisis? Only at the end of his paper did Poincaré indicate that he agreed with the second position. Let us quote the last sentence of his article, which sums up the dilemma:

> Thus, all theories which respect that principle [the principle of reaction applied to matter alone] are condemned, *at least unless we consent to profoundly modify all our ideas regarding electrodynamics.*[40]

If one chose to extend the principle of reaction to the ether in order to preserve the unity of science, one had to discard all previous theories respecting the original principle. If instead one refused to abandon the original principle of reaction, one had to question the foundations of electrodynamics. The emphasis conveyed Poincaré's own stance.

7.8 CONCLUSION

Although Poincaré judged Lorentz's theory to give the most satisfactory account of electrical and optical phenomena available at the beginning of the 20th century, he no longer believed it to be true after his jubilee lecture. Due to the fact that all previous experiments on the influence of the earth's motion had been negative, he was convinced that it was much more reasonable to "believe [...] that more exact observations will ever make evident anything else but the relative displacement of material bodies."[41] Initially, experiments were thought of that should have revealed an influence to the first order in $\frac{u}{c}$. But the results were negative. Lorentz provided an explanation that implied "that the terms of the first order should cancel each other, but not the terms of the second order."[42] More precise experiments were carried out. In 1887, Michelson and Morley improved the accuracy of the 1881 interferometer experiment, but the results were still negative. By

[40] Poincaré 2008, p. 26.
[41] Poincaré 1952, p. 172.
[42] Ibid., p. 172.

introducing the molecular force hypothesis, Lorentz was able to find an explanation for the absence of second-order effects in $\frac{u}{c}$. Poincaré's objection to this form of reasoning was that every time new negative results came to light, the need to introduce a new hypothesis might arise. In his lecture, for example, he himself had to introduce the apparent complementary force in order to resolve the apparent contradiction in Lorentz's theory. Instead, an explanation was needed that applied to all orders in $\frac{u}{c}$, meaning that the alleged theory should fulfill the principle of relative motion when applied only to matter. In 1900, Poincaré was certain that such corrections to electrodynamics would imply major changes to its foundations rather than the rejection of general principles. In particular, because of the null result of all ether drift experiments, he believed that both the principle of relative motion and the principle of reaction had to apply to matter only. That is to say, he was still convinced that a correct theory of electrodynamic phenomena would have to rigorously affirm these principles and thus maintain their status as conventions.

CHAPTER 8

Improving the Theory of Corresponding States

In his article "Electromagnetic Phenomena in a System Moving With Any Velocity Smaller Than That of Light" (1904), Lorentz attempted to meet the objection raised by Poincaré

> that, in order to explain Michelson's negative result, the introduction of a new hypothesis has been required, and that the same necessity may occur each time new facts will be brought to light.[1]

He continued:

> Surely, this course of inventing special hypotheses for each new experimental result is somewhat artificial. It would be more satisfactory, if it were possible to show, by means of certain fundamental assumptions, and without neglecting terms of one order of magnitude or another, that many electromagnetic actions are entirely independent of the system. Some years ago, I have already sought to frame a theory of this kind [1)]. I believe now to be able to treat the subject with a better result. The only restriction as regards the velocity will be that it be smaller than that of light.[2]

Despite Poincaré's reinterpretation of Lorentz's theory of corresponding states, it was only after the publication of Einstein's paper in 1905 that Lorentz realized that the effective coordinate t' was not just a mathematical artifact, but corresponded to the reading of a moving clock. In 1904, he formulated a new version of his theory in which the invariance of optical phenomena under uniform rectilinear motion applied to any order in $\frac{u}{c}$. For other phenomena, he still believed that it might be possible to detect motion with respect to the ether, so he was careful to state that "*many* [as opposed to any] electromagnetic actions are entirely independent of the system."[3] As we shall see below, Lorentz thus did not follow Poincaré in considering all quantities measured in a moving system as the apparent states of that system, but continued to reason in terms of a fictitious system brought to rest.

[1] Lorentz 1937a, pp. 173–4.
[2] Ibid., p. 174.
[3] My emphasis and comment.

8.1 PRIMED VARIABLES

Lorentz began by rewriting his microscopic equations (5.8) for a coordinate system (x, y, z, t) adapted to the ether frame S into a system of units that he had introduced in his article “Maxwells elektromagnetische Theorie” (1904) and which we now call Heaviside-Lorentz units:

$$\begin{aligned} \nabla\cdot\mathbf{d} &= \rho, \qquad \nabla\cdot\mathbf{h} = 0,\\ \nabla\times\mathbf{h} &= \frac{1}{c}\left(\rho\mathbf{v} + \frac{\partial\mathbf{d}}{\partial t}\right),\\ \nabla\times\mathbf{d} &= -\frac{1}{c}\frac{\partial\mathbf{h}}{\partial t},\\ \mathbf{f} &= \mathbf{d} + \frac{1}{c}\mathbf{v}\times\mathbf{h}.^{4} \end{aligned} \tag{8.1}$$

He went on to give his version of Maxwell's equations (5.15) and (5.16) for a coordinate system $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ adapted to a frame $\overline{S}$ that moved through the ether at constant velocity $\mathbf{u}$ in the new units:

$$\begin{aligned} \overline{\nabla}\cdot\overline{\mathbf{d}} &= \overline{\rho}, \qquad \overline{\nabla}\cdot\overline{\mathbf{h}} = 0,\\ \overline{\nabla}\times\overline{\mathbf{h}} &= \frac{1}{c}\diamond\overline{\mathbf{d}} + \frac{\overline{\rho}}{c}(\overline{\mathbf{v}} + \mathbf{u}),\\ \overline{\nabla}\times\overline{\mathbf{d}} &= -\frac{1}{c}\diamond\overline{\mathbf{h}},\\ \overline{\mathbf{f}} &= \overline{\mathbf{d}} + \frac{1}{c}\overline{\mathbf{v}}\times\overline{\mathbf{h}} + \frac{1}{c}\mathbf{u}\times\overline{\mathbf{h}}, \end{aligned} \tag{8.2}$$

where $\diamond$ denotes the scalar operator $\frac{\partial}{\partial\overline{t}} - u\frac{\partial}{\partial\overline{x}}$.[5] Lorentz then introduced new independent primed variables:

$$\begin{aligned} x' &= \gamma l\,\overline{x}, & y' &= l\,\overline{y},\\ z' &= l\,\overline{z}, & t' &= \frac{l}{\gamma}\overline{t} - \gamma l\frac{u}{c^2}\overline{x}, \end{aligned} \tag{8.3}$$

where he understood l as a function of u whose value was equal to 1 for $u = 0$ and differed from unity by no more than a second-order quantity for

[4] For details see Lorentz 1904, pp. 83–8.

[5] This means that he continued to presuppose the validity of the Galilean transformations (5.12). As before, $\overline{S}$ is chosen so that the $\overline{x}$-axis moves along the x-axis of S at speed u.

small values of u. He also defined two new primed vectors $\mathbf{d}'$ and $\mathbf{h}'$:

$$\mathbf{d}' = (l^{-2}, \gamma l^{-2}, \gamma l^{-2}) \left(\overline{\mathbf{d}} + \frac{1}{c}\mathbf{u} \times \overline{\mathbf{h}}\right),$$
$$\mathbf{h}' = (l^{-2}, \gamma l^{-2}, \gamma l^{-2}) \left(\overline{\mathbf{h}} - \frac{1}{c}\mathbf{u} \times \overline{\mathbf{d}}\right), \tag{8.4}$$

which could also be written

$$\overline{\mathbf{d}} = (l^{2}, \gamma l^{2}, \gamma l^{2}) \left(\mathbf{d}' - \frac{1}{c}\mathbf{u} \times \mathbf{h}'\right),$$
$$\overline{\mathbf{h}} = (l^{2}, \gamma l^{2}, \gamma l^{2}) \left(\mathbf{h}' + \frac{1}{c}\mathbf{u} \times \mathbf{d}'\right), \tag{8.5}$$

where $(l^{-2}, \gamma l^{-2}, \gamma l^{-2})$ and $(l^{2}, \gamma l^{2}, \gamma l^{2})$ each denote a 3×3 diagonal matrix.[6] Next, he introduced an auxiliary entity ρ' and a new vector $\mathbf{v}'$:

$$\rho' = \frac{\overline{\rho}}{\gamma l^3}, \tag{8.6}$$
$$\mathbf{v}' = (\gamma^2, \gamma, \gamma)\overline{\mathbf{v}}. \tag{8.7}$$

Based on these definitions, he arrived at the equations (8.2) in terms of the primed coordinates:

$$\nabla' \cdot \mathbf{d}' = \left(1 - \frac{u v_1'}{c^2}\right)\rho', \qquad \nabla' \cdot \mathbf{h}' = 0,$$
$$\nabla' \times \mathbf{h}' = \frac{1}{c}\left(\frac{\partial \mathbf{d}'}{\partial t'} + \rho' \mathbf{v}'\right), \tag{8.8}$$
$$\nabla' \times \mathbf{d}' = -\frac{1}{c}\frac{\partial \mathbf{h}'}{\partial t'}.^{7}$$

Before proceeding, we should remark that, if l equals 1 everywhere, the transformations between x', y', z', t' and x, y, z, t obtained by combining the equations (5.12) and (8.3) would correspond to the Lorentz transformations as we know them today:

$$x' = \gamma(x - ut), \qquad y' = y,$$
$$z' = z, \qquad t' = \gamma\left(t - \frac{u}{c^2}x\right). \tag{8.9}$$

[6] Lorentz 1937a, p. 176. However, I follow Janssen 1995, section 3.3.1, and write the transformations in matrix-vector notation.

[7] Lorentz 1937a, p. 176.

However, although Lorentz had a way to derive $l = 1$ (which we will discuss later), we should emphasize two things: Lorentz still believed that the primed variables were purely mathematical constructions with no real physical content. Secondly, he did neither obtain the (purely mathematical) Lorentz invariance of Maxwell's equations.

8.2 LORENTZ AND THE INVARIANCE OF MAXWELL'S EQUATIONS

The reason for this is that he did not define ρ' or $\mathbf{v}'$ in a way that made such a demonstration possible. On this basis, Zahar argues that Lorentz made a mistake:

> Although the *equation* $\nabla' \cdot \mathbf{d}' = \left(1 - \frac{uv_1'}{c^2}\right)\rho'$ is correct, the *correct transform* of the charge density is not ρ' but $\sigma' = \left(1 - \frac{uv_1'}{c^2}\right)\rho'$. In fact:
>
> $$\sigma' = \frac{1}{\gamma}\left(1 - \frac{uv_1'}{c^2}\right)\rho = \frac{1}{\gamma}\left(1 - \frac{u}{c^2}\gamma^2(v_1 - u)\right)\rho = \gamma\left(1 - \frac{uv_1}{c^2}\right)\rho,$$
>
> which is the value given by Einstein.
>
> Here, Lorentz seems to have made an easily corrigible mistake.[8]

Beyond setting $l = 1$, Zahar's calculations make use of expressions for ρ' and $\mathbf{v}'$ in terms of coordinates adapted to the ether frame S:

$$\rho' = \frac{\overline{\rho}}{\gamma} = \frac{\rho}{\gamma}, \tag{8.10}$$

$$\mathbf{v}' = (\gamma^2, \gamma, \gamma)\overline{\mathbf{v}} = (\gamma^2, \gamma, \gamma)(v_1 - u, v_2, v_3), \tag{8.11}$$

where we have applied (5.14) in (8.10) and the Galilean rule for the transformation of velocities in (8.11).

According to Zahar, the significance of Lorentz's mistake in the transformation of ρ' stems from the "difficulties presented by the physical interpretation of local time."[9] In particular, $\rho' = \frac{\overline{\rho}}{\gamma}$ was valid for a moving electrostatic system (it followed from (8.6) for $l = 1$), but invalid in all cases where ρ' depended on local time. Zahar goes on to write:

[8] Zahar 1989, p. 73. Zahar's own notation is slightly different from mine. I have modified it accordingly. This remark applies to all quotations to this reference.

[9] Ibid., p. 73.

> Yet, the transformed equations obtained by Lorentz in 1904 are so similar to Maxwell's that one wonders why he did not simply postulate $\left(1-\frac{uv_1'}{c^2}\right)\rho' = \sigma'$ = transformed density, and thus obtain the full covariance of Maxwell's equations. However, if he had taken this step, he would still have had to face the problem posed by equation (iii') [my equation (8.8)], which now becomes
>
> $$\nabla'\times\mathbf{h}' = \frac{1}{c}\left(\frac{\partial\mathbf{d}'}{\partial t'} + \sigma'\frac{\mathbf{v}'}{1-\frac{uv_1'}{c^2}}\right).$$
>
> He would have had to interpret $\mathbf{v}'/(1-\frac{uv_1'}{c^2})$ as an *effective velocity*, and for this he needed to have a general kinematical framework.[10]

I think this way of putting it is somewhat unsatisfactory. Zahar himself emphasizes that Lorentz introduced the primed variables as mathematical tools for the sole purpose of theoretical calculations. In the case of a moving system, these entities had therefore no direct physical interpretation for him. For this reason alone, it is strange to argue that his *definitions* were wrong. Yet, it is correct that Lorentz's definitions of ρ' and $\mathbf{v}'$ were not in conformity with his definition of the primed variables in terms of the barred ones. To see this in the case of $\mathbf{v}'$, we write the differentials for the transformations (8.3):

$$\begin{aligned} dx' &= \gamma l d\bar{x}, & dy' &= l d\bar{y}, \\ dz' &= l d\bar{z}, & dt' &= \frac{l}{\gamma}d\bar{t} - \gamma l\frac{u}{c^2}d\bar{x}. \end{aligned} \tag{8.12}$$

[10] Zahar 1989, pp. 73–4. The concept of covariance has both a narrow and a broad sense. In its narrow sense, it refers to the requirement that the laws of physics preserve their form in inertial frames of reference. In its broad sense, general covariance requires that the laws of physics be expressed in such a way that they preserve their form in all coordinate systems, not just inertial ones. In practice, this is naturally achieved by formulating physical laws in tensorial form. In the quote, Zahar uses the concept of covariance in its narrow sense. In other words, it corresponds to what Minkowski (1909, p. 75) called the invariance of form under uniform rectilinear motion. To avoid confusion with the later usage in general relativity, I keep following Minkowski's wording.

If we define $v_1^* = \frac{dx'}{dt'}$, it follows that

$$v_1^* = \frac{dx'/d\bar{t}}{dt'/d\bar{t}} = \frac{\gamma l d\bar{x}/d\bar{t}}{\gamma^{-1} l d\bar{t}/d\bar{t} - \gamma l u c^{-2} d\bar{x}/d\bar{t}}$$
$$= \frac{\gamma \bar{v}_1}{\gamma^{-1} - \gamma u c^{-2} \bar{v}_1} = \frac{\gamma^2 \bar{v}_1}{1 - \gamma^2 \frac{u\bar{v}_1}{c^2}}. \tag{8.13}$$

In the same way, we find expressions for $v_2^* = \frac{dy'}{dt'}$ and $v_3^* = \frac{dz'}{dt'}$:

$$v_2^* = \frac{\gamma \bar{v}_2}{1 - \gamma^2 \frac{u\bar{v}_1}{c^2}}, \qquad v_3^* = \frac{\gamma \bar{v}_3}{1 - \gamma^2 \frac{u\bar{v}_1}{c^2}}. \tag{8.14}$$

If we compare the components of $\mathbf{v}^*$ and $\mathbf{v}'$, we get the relation

$$\mathbf{v}^* = \frac{\mathbf{v}'}{1 - \frac{u v_1'}{c^2}}. \tag{8.15}$$

This means that $\mathbf{v}^*$, unsurprisingly, corresponds to Zahar's effective velocity. What this derivation emphasizes is that Lorentz did not introduce and therefore did not think of the primed scalars and primed vector fields such as ρ', $\mathbf{d}'$, $\mathbf{h}'$ and $\mathbf{v}'$ in terms of how their barred counterparts $\bar{\rho}$, $\bar{\mathbf{d}}$, $\bar{\mathbf{h}}$ and $\bar{\mathbf{v}}$ transform under the change of the independent coordinates from $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$ to (x', y', z', t').

8.3 RETARDED POTENTIALS AND ELECTRIC DIPOLE MOMENT

In lectures, which Lorentz gave at Columbia University in 1906 and which were later published under the title *The Theory of Electrons* (1909), he explained how he had actually defined the primed variables:

> Our aim must again be to reduce, at least as far as possible, the equations for a moving system to the form of the ordinary formulae that hold for a system at rest. It is found that the transformations needed for this purpose may be left indeterminate to a certain extent; our formulae will contain a numerical coefficient l, of which we shall provisionally assume only that it is a function of the velocity of translation u, whose value is equal to unity for $u = 0$.[11]

[11] Lorentz 1916, p. 196.

Lorentz thus tried to find transformations that brought the field equations for a moving system (8.2) back into their form for a stationary system (8.1). However, as the last section shows, he did not know how to determine the invariance group of a set of equations.[12] If one compares Lorentz's version of Maxwell's equations (8.1) with the corresponding equations for the primed variables (8.8) that Lorentz was able to obtain by means of the above definitions, it becomes clear that he was only successful in the case of vacuum (space devoid of ponderable matter, but filled with ether) and electrostatics. In the general case he was mistakenly of the opinion that a "reduction" was impossible (as the quotation makes clear). In other words, his definitions were not optimal in his own terms. Despite this fact, as we shall see in detail, he was able to demonstrate the invariance of optical phenomena under uniform rectilinear motion.

The first step Lorentz took in this direction was to represent $\mathbf{d}'$ and $\mathbf{h}'$ by means of a scalar potential φ' and a vector potential $\mathbf{A}'$, which satisfied the equations

$$\left(\nabla'^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\right)\varphi' = -\rho', \tag{8.16}$$

$$\left(\nabla'^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\right)\mathbf{A}' = -\frac{1}{c^2}\rho'\mathbf{v}', \tag{8.17}$$

and made it possible to express $\mathbf{d}'$ and $\mathbf{h}'$ as

$$\mathbf{d}' = -\frac{1}{c}\frac{\partial \mathbf{A}'}{\partial t'} - \nabla'\varphi' + \frac{u}{c}\nabla' A_1', \tag{8.18}$$

$$\mathbf{h}' = \nabla' \times \mathbf{A}'.^{13} \tag{8.19}$$

He then obtained solutions for the two potentials in terms of the primed coordinates: for a definite local time t' the potentials at a point (x', y', z') were given by

$$\varphi' = \frac{1}{4\pi}\int \frac{[\rho']}{r'}\mathrm{d}\tau', \tag{8.20}$$

$$\mathbf{A}' = \frac{1}{4\pi c}\int \frac{[\rho'\mathbf{v}']}{r'}\mathrm{d}\tau', \tag{8.21}$$

[12] Therefore, strictly speaking, he did not commit the said mistake.

[13] Lorentz 1937a, p. 177.

where the integration took place over all volume elements $d\tau'$ of the primed coordinates, r' denoted the distance between (x', y', z') and $d\tau'$, and the brackets indicated that an enclosed expression should be evaluated at the time $t' - \frac{r'}{c}$.[14] Lorentz applied these equations to two special cases: 1) charges at rest relative to the moving frame and 2) a polarized particle.

In the electrostatic case, where $\overline{\mathbf{v}} = \mathbf{0}$, (8.7) reduced to $\mathbf{v}' = \mathbf{0}$. The equations (8.16), (8.18), (8.19) and (8.21) reduced to

$$\begin{aligned} \nabla'^2\varphi' &= -\rho', & \mathbf{A}' &= \mathbf{0}, \\ \mathbf{d}' &= -\nabla'\varphi', & \mathbf{h}' &= \mathbf{0}, \end{aligned} \tag{8.22}$$

while $\overline{\mathbf{f}}$ in (8.2) became

$$\overline{\mathbf{f}} = (l^2, \gamma^{-1}l^2, \gamma^{-1}l^2)\mathbf{d}'.^{15} \tag{8.23}$$

From this followed

$$\nabla' \cdot \mathbf{d}' = -\nabla'^2\varphi' = \rho', \tag{8.24}$$

so that the primed set of equations (8.8) reduced to the equations for an electrostatic system at rest. As in 1895, Lorentz therefore compared the moving electrostatic system with a second, fictitious electrostatic system at rest. While he denoted the former by Σ and the latter by Σ', I will continue to distinguish between three systems, which I denote by Σ, $\overline{\Sigma}$ and Σ'. Σ denotes the electrostatic system $\overline{\Sigma}$ in the special case that $u = 0$ $(\Sigma = \overline{\Sigma}(0))$. The dimensions of the fictitious system Σ' were obtained from $\overline{\Sigma}$ by multiplying the dimension of $\overline{\Sigma}$ parallel to the $\overline{x}$-axis by γl and the dimension parallel to the $\overline{y}$-axis or $\overline{z}$-axis by l:

$$\dim(\Sigma') = (\gamma l, l, l)\dim(\overline{\Sigma}). \tag{8.25}$$

Lorentz then defined the charge density of Σ' as equal to $\rho' = \frac{\overline{\rho}}{\gamma}$ "so that the charges of corresponding elements of volume and of corresponding electrons are the same in $\overline{\Sigma}$ and Σ'."[16] He had thus defined a fictitious electrostatic system Σ' at rest with the help of the moving electrostatic system $\overline{\Sigma}$. He was therefore led to the following interpretation of formula (8.23): The components of the force acting on an electron of the moving system

[14] Ibid., p. 178.
[15] Ibid., p. 179.
[16] Ibid., p. 179. I have changed Σ to $\overline{\Sigma}$ according to my modification of Lorentz's notation.

$\overline{\Sigma}$ could be obtained by multiplying the components of the corresponding force in Σ' by $(l^2, \gamma^{-1}l^2, \gamma^{-1}l^2)$:

$$\overline{\mathbf{f}} = (l^2, \gamma^{-1}l^2, \gamma^{-1}l^2)\mathbf{f}'. \tag{8.26}$$

In the second case, Lorentz used the equations (8.20) and (8.21) to determine the potentials φ' and $\mathbf{A}'$ at a point (x', y', z', t') in Σ', which corresponded to a spatial point $(\overline{x}, \overline{y}, \overline{z})$ at a finite distance from a polarized particle in $\overline{\Sigma}$ at a time $\overline{t}$. The polarized particle, in turn, was characterized as a particle with total charge $\int \overline{\rho}\, \mathrm{d}\overline{\tau} = 0$ and an electric dipole moment

$$\overline{\mathbf{p}} = \int \overline{\rho}\, \overline{\mathbf{r}}\, \mathrm{d}\overline{\tau} \tag{8.27}$$

deviating from $\mathbf{0}$. Introducing the corresponding vector

$$\mathbf{p}' = (\gamma l, l, l)\overline{\mathbf{p}} \tag{8.28}$$

and neglecting the squares and products of the components of $\overline{\mathbf{r}}$ and $\overline{\mathbf{v}} = \frac{\mathrm{d}\overline{\mathbf{r}}}{\mathrm{d}\overline{t}}$, Lorentz was then able to arrive at equations for φ' and $\mathbf{A}'$ in terms of $\mathbf{p}'$:

$$\varphi' = \frac{u}{4\pi c^2 r'} \frac{\partial [p'_1]}{\partial t'} - \frac{1}{4\pi} \nabla' \cdot \left(\frac{[\mathbf{p}']}{r'} \right), \tag{8.29}$$

$$\mathbf{A}' = \frac{1}{4\pi c r'} \frac{\partial [\mathbf{p}']}{\partial t'}, \tag{8.30}$$

where all enclosed quantities had to be evaluated at time $t' - \frac{r'}{c}$ and r' denoted the distance between (x', y', z') and the point in Σ' that corresponded to the center of the particle in $\overline{\Sigma}$.[17] Lorentz did not elucidate whether he assumed that the system Σ' would have a fictitious physical interpretation like in the electrostatic case. There are reasons to be skeptical about this: it is somehow striking that he did nothing to emphasize the possibility of rendering $\mathbf{p}'$ in terms of primed quantities:

$$\mathbf{p}' = (\gamma l, l, l)\overline{\mathbf{p}} = \int \rho' \mathbf{r}'\, \mathrm{d}\tau'. \tag{8.31}$$

This formula suggests that $\mathbf{p}'$ can be thought of as the dipole moment of a particle in Σ'. However, this interpretation is problematic because $\overline{\mathbf{p}}$ was

[17] Lorentz 1937a, pp. 180–2. Lorentz did not explain how to determine the center $\overline{\mathbf{r}}_c$ of a particle. As will become evident later, he could have defined $\overline{\mathbf{r}}_c = \frac{\int \overline{\rho}\, \overline{\mathbf{r}}\, \mathrm{d}\overline{\tau}}{\int \overline{\rho}\, \mathrm{d}\overline{\tau}}$.

meant to vary over time. The latter implied the existence of moving charge, so that the primed equations (8.8) did not reduce to the form corresponding to the field equations for an electromagnetic system at rest. In the second and more general case, therefore, there was no immediate fictitious physical interpretation.

Despite this fact, we should bear in mind that Lorentz only wanted to demonstrate the invariance of optical phenomena under rectilinear motion. For this reason alone, it was legitimate for him to introduce certain approximations. We have already noted that he neglected the squares and products of the components of $\overline{\mathbf{r}}$ and $\overline{\mathbf{v}}$ in his derivation of (8.29) and (8.30). Considering his overall goal, these approximations seem reasonable. First, the radius of the particles was supposed to be negligible compared to the distance between the particles (i.e. the distance at which the force exerted by one particle on another was evaluated). Secondly, it was assumed that intra-atomic velocities were small compared to the velocity of light. This information is very important because it follows that the components of $\frac{\mathbf{v}'}{c}$ were also negligible, so that the set of equations (8.8) reduced to the form of the general field equations (8.1) of a system at rest. It follows that Zahar's corrected formula for the transformation of the charge density reduces to

$$\sigma' = \rho', \tag{8.32}$$

whereas my comparison between the correct velocity transformation and Lorentz's own definition results in

$$\mathbf{v}^* = \mathbf{v}'. \tag{8.33}$$

Under the above assumptions, Lorentz's definitions thus prove to be consistent with the correct transformation formulas, implying that Σ' also has a fictitious physical interpretation in the second case.

Instead of arguing with the primed equations (8.8), Lorentz himself chose a more indirect way by obtaining solutions for $\mathbf{d}'$ and $\mathbf{h}'$ in terms of $\mathbf{p}'$. For example, by means of (8.18),

$$\mathbf{d}' = -\frac{1}{4\pi c^2}\frac{\partial^2}{\partial t'^2}\left(\frac{[\mathbf{p}']}{r'}\right) + \frac{1}{4\pi}\nabla'\left(\nabla'\cdot\left(\frac{[\mathbf{p}']}{r'}\right)\right), \tag{8.34}$$

while $\mathbf{h}'$ could be expressed in terms of $\mathbf{p}'$ by inserting (8.30) into (8.19). Lorentz noticed that (8.23) remained valid under the above approximations,

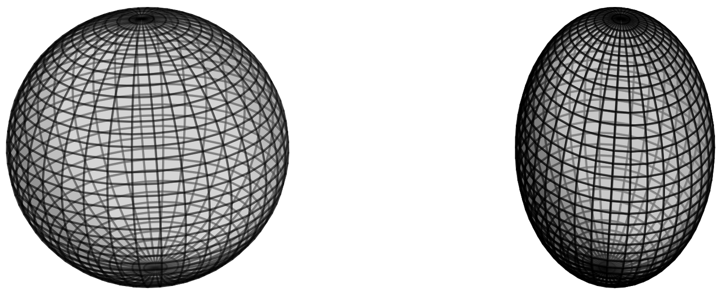


Fig. 8.1: Lorentz's contractile electron model.

and together with (8.34) he used it to determine the force exerted by a polarized particle in $\overline{\Sigma}$ on another at some distance.[18] The formulas he obtained in this way allowed him to argue both at the molecular and intra-atomic level.

8.4 LORENTZ CONTRACTION REVISITED

Compared to his article of 1895, Lorentz had not yet introduced any new physical assumptions, but only new auxiliary variables to which he assigned a fictitious physical interpretation. In section 8, however, he made the hypothesis that electrons, which he assumed to be spheres of radius R in the rest state with charge e evenly distributed over the surface, "have their dimensions changed by the effect of a translation, the dimensions in the direction of motion becoming γl times and those in perpendicular directions l times smaller."[19] After Lorentz had introduced this contractile electron model, he applied it to an electrostatic system $\overline{\Sigma}$ moving at a constant velocity $\mathbf{u}$. By hypothesis, the electrons of this system would be flattened ellipsoids in $\overline{\Sigma}$. In accordance with the previous section, it was now possible to define a second system Σ' by means of $\overline{\Sigma}$, so that the electrons would be spherical objects with radius R in Σ' and the electrostatic forces in Σ' and $\overline{\Sigma}$ would satisfy the relationship (8.26). So far, he was therefore able to draw the following conclusion: Imagine an electron in the system Σ at rest and exerting a force $\mathbf{f}$ on an electron placed at a certain distance $\mathbf{r}$ from it. If brought into uniform rectilinear motion, the force $\overline{\mathbf{f}}$ that the electron would exert on an electron placed at a distance $\overline{\mathbf{r}} = (\gamma l, l, l)\mathbf{r}$ from it would correspond to $\overline{\mathbf{f}} = (l^2, \gamma^{-1}l^2, \gamma^{-1}l^2)\mathbf{f}$.

[18] Lorentz 1937a, p. 182.

[19] Ibid., pp. 182–3.

Lorentz went on to add another hypothesis:

> In the second place I shall suppose *that the forces between uncharged particles, as well as those between such particles and electrons, are influenced by a translation in quite the same way as the electric forces in an electrostatic system.* In other terms, whatever be the nature of the particles composing a ponderable body, so long as they do not move relatively to each other, we shall have between the forces acting in a system (Σ') without, and the same system ($\overline{\Sigma}$) with a translation, the relation specified in (21) [(8.26) in my text] if, as regards the relative position of the particles, Σ' is got from $\overline{\Sigma}$ by the deformation $(\gamma l, l, l)$, or $\overline{\Sigma}$ from Σ' by the deformation $(\frac{1}{\gamma l}, \frac{1}{l}, \frac{1}{l})$.[20]

What he assumed in this passage was a bold generalization of the transformation law for electrostatic forces (under uniform rectilinear motion) to *all* forces in nature: Whatever the nature of the forces acting between particles or electrons, a force acting on a particle or electron in $\overline{\Sigma}$ would always, if the relative position between them remained fixed, have the relation (8.26) to its corresponding force in Σ'. This second hypothesis enabled him to derive the following result:

> We see by this that, as soon as the resulting force is 0 for a particle in Σ', the same must be true for the corresponding particle in $\overline{\Sigma}$. Consequently, if, neglecting the effects of molecular motion, we suppose each particle of a solid body to be in equilibrium under the action of the attractions and repulsions exerted by its neighbours, and if we take for granted that there is but one configuration of equilibrium, we may draw the conclusion that the system Σ', if the velocity u is imparted to it, will *of itself* change into the system $\overline{\Sigma}$. In other terms, the translation will *produce* the deformation $(\frac{1}{\gamma l}, \frac{1}{l}, \frac{1}{l})$.[21]

Although Lorentz did not emphasize it, this argument only applied at the molecular level. On the basis of the generalized hypothesis, it predicted how the distances between the particles of a system would change if we imposed a translation on it. In terms of the distinction between Σ, $\overline{\Sigma}$ and Σ', he concluded that Σ' turned out to be the system we would get if we really brought $\overline{\Sigma}$ to rest: $\Sigma = \overline{\Sigma}(0) = \Sigma'$. At first glance, the only difference between

[20] Ibid., p. 183. I have replaced Σ by $\overline{\Sigma}$ and k by γ.

[21] Ibid., p. 183. I have replaced k by γ, w by u, and Σ by $\overline{\Sigma}$.

this argument and the 1895 account of the Michelson-Morley experiments seems to be the introduction of the ambiguity in the deformation, which did not ruin the argument since the relative deformations remained the same. However, such a remark overlooks a crucial difference. As the title of Lorentz's 1904 article was meant to emphasize, the new account, unlike the earlier one, remained valid for "a system moving with any velocity smaller than that of light" and thus for all orders in $\frac{u}{c}$. It was thus Lorentz's first major step towards the fulfillment of Poincaré's requirement.

8.5 OPTICS REVISITED

The second decisive step in this direction was Lorentz's attempt (so glaringly absent in 1895) to explain how the contraction of bodies would affect optical phenomena in a moving system of dielectrics. The hypothesis that all electrons were bound in the dielectrics, which Lorentz had used in 1895 to set the macroscopic density of free charge equal to 0, was now reformulated as the assumption that the total charge of each particle was equal to 0 and that there was no charge outside the particles. This enabled him to study optical phenomena in a dielectric by means of a system of polarized particles. First and foremost, he neglected molecular motion (understood as the relative motion between the particles of the dielectric). He also assumed that "in each particle, the charge is concentrated in a certain number of separate electrons"[22] and added a further contraction hypothesis:

> Regarding the interior state of the atoms, we shall assume that the configuration of a particle A in $\overline{\Sigma}$ at a certain time may be derived by means of the deformation $(\frac{1}{\gamma l}, \frac{1}{l}, \frac{1}{l})$ from the configuration of the corresponding particle in Σ', such as it is at the corresponding instant. Insofar as this assumption relates to the form of the electrons themselves, it is implied in the first hypothesis of §8.[23]

Previously, Lorentz had derived a change in distance between particles at rest relative to each other on the basis of the generalized force hypothesis, which implied that the derivation did not allow any conclusions to be drawn about the internal state of the particles.[24] In the passage just quoted, he added the hypothesis that during the transition from rest to uniform

[22] Lorentz 1937a, p. 186.
[23] Ibid., pp. 186–7.
[24] Lorentz spoke alternately of atoms and particles.

rectilinear motion, not only the shape of the electrons but the interior of a particle as a whole would undergo a deformation by $(\frac{1}{\gamma l}, \frac{1}{l}, \frac{1}{l})$. More precisely, in addition to the change in molecular distances, Lorentz assumed that the state of the interior of a particle placed at $(\bar{x}, \bar{y}, \bar{z})$ in $\bar{\Sigma}$ at time $\bar{t}$ would be given by a deformation by $(\frac{1}{\gamma l}, \frac{1}{l}, \frac{1}{l})$ from the configuration of the corresponding particle placed at (x', y', z') in Σ' at the corresponding time t'. Taken together, Lorentz's assumptions were meant to imply that, if we put Σ in uniform rectilinear motion, it would transform in such a way that the resulting system $\bar{\Sigma}$ would have a system identical to Σ as its corresponding system Σ'. As a result, the relationships that obtained between the barred and primed quantities would also hold between the quantities in $\bar{\Sigma}$ and Σ.[25] Starting from the state of a system of particles whose centers remained at rest relative the ether, apart from the indeterminacy of the function l, Lorentz had thus completely defined the resulting state of the moving system.

8.6 ELECTROMAGNETIC MASS OF AN ELECTRON

In order to eliminate the indeterminacy of l, Lorentz derived conditions which, in his opinion, had to be fulfilled in order to bring his contractile electron model in conformity with Newton's second law of motion. As already indicated, he assumed that a single electron moving at constant velocity **u** was itself a moving electrostatic system: a flattened ellipsoid with its smaller axis in the direction of motion. Follwing Abraham, he went on to define its electromagnetic momentum in consonance with the force that an electrostatic system experienced due to its self-field:

$$\mathbf{F}_{\text{self}} = -\frac{\mathrm{d}\mathbf{G}}{\mathrm{d}t}, \tag{8.35}$$

where

$$\mathbf{G} = \frac{1}{c}\int \mathbf{d} \times \mathbf{h}\, \mathrm{d}\tau.^{26} \tag{8.36}$$

If we compare this equation with (7.7), it becomes clear that Abraham's definition was inspired by Poincaré's contribution to the jubilee volume. Abraham in fact explicitly referred to Poincaré:

---

[25] As we have already noted, it was only possible to interpret Σ' as an electromagnetic system at rest under certain conditions. To satisfy the latter, Lorentz assumed that in a system at rest or in uniform rectilinear motion, the atoms remained at rest relative to each other and that the velocities of the electrons oscillating in the atoms were small compared to the speed of light.

[26] See Lorentz 1904, pp. 162–63, Lorentz 1937a, p. 179, Abraham 1902 and Abraham 1903.

> The position of Lorentz's theory with respect to Newton's third axiom was discussed in detail by H. Poincaré; he emphasized that the law of conservation of momentum retains its validity if a specific momentum is attributed to the electromagnetic field.[27]

Accordingly, Abraham did not realize in 1902 that Poincaré only introduced the ether as a fictitious, momentum carrying fluid in order to argue that it was a step in the wrong direction.

Since $\mathbf{h}' = \mathbf{0}$ for the corresponding electrostatic system at rest in Σ', the x-component of $\mathbf{G}$ reduced by (8.5) to

$$G_1 = \frac{1}{c}\int d_2 h_3 - h_2 d_3 \,\mathrm{d}\tau = \frac{\gamma l u}{c^2}\int {d'_2}^2 + {d'_3}^2 \,\mathrm{d}\tau',^{28}$$

where we have used that

$$\mathrm{d}\tau = \mathrm{d}\overline{\tau} = \frac{1}{\gamma}\mathrm{d}\tau'. \tag{8.37}$$

Due to the spherical symmetry of the electron in Σ',

$$G_1 = \frac{\gamma l u}{c^2}\int \frac{2}{3} d'^2 \,\mathrm{d}\tau' = \frac{4}{3}\frac{\gamma l u}{c^2}\left(\int \frac{1}{2} d'^2 \,\mathrm{d}\tau'\right) = \frac{4}{3}\frac{\gamma l u}{c^2}E', \tag{8.38}$$

where E' denotes the electromagnetic energy of an electron at rest with respect to the ether. Since it was assumed that the spherical charge distribution of an electron at rest was uniformly distributed,

$$E' = \frac{1}{2}\int d'^2 \,\mathrm{d}\tau' = \frac{e^2}{8\pi}\int_R^\infty \frac{\mathrm{d}r'}{r'^2} = \frac{e^2}{8\pi R}. \tag{8.39}$$

It followed that

$$G_1 = \frac{e^2}{6\pi c^2 R}\gamma l u.^{29} \tag{8.40}$$

Remarking that the integrals over the y-axis and the z-axis would vanish due to the symmetry of the charge distribution, Lorentz concluded

$$\mathbf{G} = \frac{e^2}{6\pi c^2 R}\gamma l \mathbf{u}.^{30} \tag{8.41}$$

---

[27] Abraham 1902, p. 99. My translation.

[28] Lorentz 1937a, pp. 179–80.

[29] Ibid., p. 184.

This result was only valid for uniform rectilinear motion. However, following Abraham, Lorentz emphasized that it could be extended by approximation to so-called *quasi-stationary* translations, for which the changes in the state of motion were sufficiently slow.[31] In particular, "the motion of an electron may be treated as quasi-stationary if it changes very little during the time a light-wave takes to travel over a distance equal to the diameter."[32] This condition was fulfilled for optical phenomena "because the diameter of an electron is extremely small in comparison with the wavelength."[33] Consequently, Lorentz went on to apply his formula (8.41) for the electromagnetic momentum of an electron to this particular case of quasi-stationary translation.

Decomposing the acceleration $\mathbf{a}$ of the electron into the components tangential and normal to its trajectory,

$$\mathbf{a} = \mathbf{a}_{\|} + \mathbf{a}_{\perp}, \tag{8.42}$$

the Dutch physicist observed that the external force $\mathbf{F}_{\text{ext}}$ on the electron consisted of the following two components:

$$\mathbf{F}_{\text{ext}} = m_1 \mathbf{a}_{\|} + m_2 \mathbf{a}_{\perp}, \tag{8.43}$$

where

$$m_1 = \frac{e^2}{6\pi c^2 R} \frac{\mathrm{d}(\gamma l |\mathbf{u}|)}{\mathrm{d}|\mathbf{u}|}, \qquad m_2 = \frac{e^2}{6\pi c^2 R} \gamma l.^{34} \tag{8.44}$$

Lorentz himself concluded the following at this point:

> Hence, in phenomena in which there is an acceleration in the direction of motion, the electron behaves as if it had a mass m_1, in those in which the acceleration is normal to the path, as if the mass were m_2. These quantities m_1 and m_2 may therefore properly be called the "longitudinal" and "transverse" electromagnetic masses of the electron. I shall suppose *that there is no other, no "true" or "material" mass.*[35]

[30] Ibid., p. 184. For more details, see Janssen 1995, section 3.4.2.
[31] Lorentz 1937a, p. 185.
[32] Ibid., p. 189.
[33] Ibid., p. 189.
[34] Ibid., p. 185.
[35] Ibid., p. 185.

Because equation (8.43) resembled Newton's second law of motion, Lorentz interpreted the quantities m_1 and m_2 as measures of an electron's inertia arising from its interaction with its own electromagnetic field. In this way, and in line with Abraham's electromagnetic worldview—which sought to ground mechanics entirely in electrodynamics—Lorentz assumed that an electron's inertia was of purely electromagnetic origin.

Let me reconstruct how Lorentz arrived at this interpretation by examining his account of the differentiation of **G**, given at a lecture series at Columbia University in the spring of 1906.[36] In particular, I would like to clarify why he had to assume that the Newtonian mass m_N of an electron was equal to 0. Noticing that **G** and **u** have the same direction, he rewrote the former as

$$\mathbf{G} = \frac{G(|\mathbf{u}|)}{|\mathbf{u}|}\mathbf{u}. \tag{8.45}$$

In combination with (8.41), it follows that

$$G(|\mathbf{u}|) = \frac{e^2}{6\pi c^2 R}\gamma l|\mathbf{u}|.$$

Differentiating **G** with respect to time, he obtained

$$\mathbf{F}_{\text{self}} = -\frac{\mathrm{d}\mathbf{G}}{\mathrm{d}t} = -\frac{G}{|\mathbf{u}|}\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} - \frac{\mathrm{d}(G/|\mathbf{u}|)}{\mathrm{d}t}\mathbf{u} = -\frac{G}{|\mathbf{u}|}\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} - \frac{\mathrm{d}(G/|\mathbf{u}|)}{\mathrm{d}|\mathbf{u}|}\frac{\mathrm{d}|\mathbf{u}|}{\mathrm{d}t}\mathbf{u}.$$

He then noticed that

$$\mathbf{u}\frac{\mathrm{d}|\mathbf{u}|}{\mathrm{d}t} = |\mathbf{u}|\mathbf{a}_{\parallel},$$

which we can check by differentiating $|\mathbf{u}| = (\mathbf{u}\cdot\mathbf{u})^{\frac{1}{2}}$ with respect to time:

$$\frac{\mathrm{d}|\mathbf{u}|}{\mathrm{d}t} = \frac{1}{2}(\mathbf{u}\cdot\mathbf{u})^{-\frac{1}{2}}\frac{\mathrm{d}(\mathbf{u}\cdot\mathbf{u})}{\mathrm{d}t} = \frac{1}{2}\frac{1}{|\mathbf{u}|}\left(\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\cdot\mathbf{u} + \mathbf{u}\cdot\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}\right) = \frac{\mathbf{a}\cdot\mathbf{u}}{|\mathbf{u}|}.$$

If we insert this result into the expression for $\mathbf{a}_{\parallel}$, we get the desired formula:

$$\mathbf{a}_{\parallel} = \left(\frac{\mathbf{a}\cdot\mathbf{u}}{|\mathbf{u}|}\right)\frac{\mathbf{u}}{|\mathbf{u}|} = \frac{\mathrm{d}|\mathbf{u}|}{\mathrm{d}t}\frac{\mathbf{u}}{|\mathbf{u}|}.$$

He could therefore rewrite the expression for $\mathbf{F}_{\text{self}}$ as follows:

$$\mathbf{F}_{\text{self}} = -\frac{G}{|\mathbf{u}|}(\mathbf{a}_{\parallel} + \mathbf{a}_{\perp}) - |\mathbf{u}|\frac{\mathrm{d}(G/|\mathbf{u}|)}{\mathrm{d}|\mathbf{u}|}\mathbf{a}_{\parallel} = -\frac{\mathrm{d}G}{\mathrm{d}|\mathbf{u}|}\mathbf{a}_{\parallel} - \frac{G}{|\mathbf{u}|}\mathbf{a}_{\perp}.$$

[36] Lorentz 1916, p. 250.

If we calculate the two components, we get

$$\frac{\mathrm{d}G}{\mathrm{d}|\mathbf{u}|} = \frac{e^2}{6\pi c^2 R}\frac{\mathrm{d}(\gamma l|\mathbf{u}|)}{\mathrm{d}|\mathbf{u}|}, \qquad \frac{G}{|\mathbf{u}|} = \frac{e^2}{6\pi c^2 R}\gamma l.$$

Introducing

$$m_1 = \frac{\mathrm{d}G}{\mathrm{d}|\mathbf{u}|}, \quad m_2 = \frac{G}{|\mathbf{u}|}, \tag{8.46}$$

it follows that

$$\mathbf{F}_{\text{self}} = -m_1\mathbf{a}_{\|} - m_2\mathbf{a}_{\perp}.$$

Now we only need to reformulate this formula with regard to $\mathbf{F}_{\text{ext}}$ to obtain equation (8.43).

Lorentz also demonstrated how to do this in his lecture series.[37] He started from the assumption that the electron had a certain Newtonian mass m_N. It followed that the total force on an electron was the sum of the force due to its self-field and any external forces acting on it:

$$\mathbf{F}_{\text{tot}} = \mathbf{F}_{\text{ext}} + \mathbf{F}_{\text{self}}.$$

Rewriting the total force using Newton's second law of motion

$$\mathbf{F}_{\text{tot}} = m_N\,\mathbf{a} = m_N(\mathbf{a}_{\|} + \mathbf{a}_{\perp}),$$

he obtained an expression for $\mathbf{F}_{\text{ext}}$:

$$\mathbf{F}_{\text{ext}} = \mathbf{F}_{\text{tot}} - \mathbf{F}_{\text{self}} = (m_N + m_1)\mathbf{a}_{\|} + (m_N + m_2)\mathbf{a}_{\perp}.$$

This expression brings to the fore why it makes sense to interpret m_1 and m_2 as the components of the electromagnetic inertia of an electron. If we set $m_N = 0$, Lorentz's equation (8.43) follows.

Given that γ and l differed from unity only at second order in $\frac{u}{c}$, Lorentz found for $u \ll c$ that

$$m_0 = m_1 = m_2 = \frac{e^2}{6\pi c^2 R}. \tag{8.47}$$

This was "the mass with which we are concerned, if there are small vibratory motions of the electrons in a system without translation."[38] In the case

[37] Ibid., p. 38.
[38] Lorentz 1937a, p. 185.

of a system moving with velocity $\mathbf{u}$, he could therefore write the mass m of an electron in the form

$$m = \left(\frac{\mathrm{d}(\gamma l |\mathbf{u}|)}{\mathrm{d}|\mathbf{u}|}, \gamma l, \gamma l \right) m_0. \tag{8.48}$$

Furthermore, Lorentz concluded with the help of (8.44) and (8.47) for electrons oscillating in the rest system Σ that

$$\mathbf{F}_{\text{ext}} = m_0 \mathbf{a}\,, \tag{8.49}$$

in accordance with Newton's second law of motion. Let us now clarify how Lorentz applied his formulas for the electromagnetic inertia of an electron in order to determine the function l.

8.7 DETERMINATION OF *L*

Assuming that the relative distances between the particles in which the electrons were located remained unchanged in Σ despite the oscillations of the latter, Lorentz recalled the hypothesis that this would imply that, whatever its nature, *every* force $\overline{\mathbf{F}}$ acting on a particle or an electron in $\overline{\Sigma}$ would be "influenced by a translation in quite the same way as the electric forces in an electrostatic system."[39] In other words, the relation of $\overline{\mathbf{F}}$ in $\overline{\Sigma}$ to its corresponding force $\mathbf{F}$ in Σ was

$$\overline{\mathbf{F}} = (l^2, \gamma^{-1} l^2, \gamma^{-1} l^2)\mathbf{F}.$$

By means of the hypothesis, it followed that this correspondence "may also be regarded as indicating the relation between the total forces, acting on corresponding electrons, at corresponding instants:"[40]

$$\overline{\mathbf{F}}_{\text{ext}} = (l^2, \gamma^{-1} l^2, \gamma^{-1} l^2)\mathbf{F}_{\text{ext}},$$

which is equivalent to

$$\mathbf{F}_{\text{ext}} = (l^{-2}, \gamma l^{-2}, \gamma l^{-2})\overline{\mathbf{F}}_{\text{ext}}. \tag{8.50}$$

According to his own statements, Lorentz then obtained the following relationship between $\mathbf{a}$ and $\overline{\mathbf{a}}$ by means of the transformations (8.3) between primed and barred coordinates:

$$\mathbf{a} = (\gamma^3 l^{-1}, \gamma^2 l^{-1}, \gamma^2 l^{-1})\overline{\mathbf{a}}. \tag{8.51}$$

---

[39] Lorentz 1937a, p. 183.

[40] Ibid., p. 188.

We should note that this relationship is much more complicated in the general case. However, if we assume that $\frac{\bar{v}}{c}$ is small, we already obtained

$$\frac{\mathrm{d}\mathbf{r}'}{\mathrm{d}t'} = (\gamma^2, \gamma, \gamma)\frac{\mathrm{d}\bar{\mathbf{r}}}{\mathrm{d}\bar{t}}.^{41} \tag{8.52}$$

As Lorentz later emphasized,[42] given that $\frac{\bar{v}}{c} \ll 1$ and $\frac{u}{c} < 1$ implied $\frac{|u\bar{v}_1|}{c^2} < \frac{u\bar{v}}{c^2} \ll 1$, it was possible to approximate

$$\frac{\mathrm{d}t'}{\mathrm{d}\bar{t}} = \frac{l}{\gamma} - \gamma l \frac{u}{c^2}\bar{v}_1$$

by

$$\frac{\mathrm{d}t'}{\mathrm{d}\bar{t}} \approx \frac{l}{\gamma},$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t'} \approx \frac{\gamma}{l}\frac{\mathrm{d}}{\mathrm{d}\bar{t}}.$$

Together with (8.52), it followed that

$$\frac{\mathrm{d}^2\mathbf{r}'}{\mathrm{d}t'^2} = (\gamma^3 l^{-1}, \gamma^2 l^{-1}, \gamma^2 l^{-1})\frac{\mathrm{d}^2\bar{\mathbf{r}}}{\mathrm{d}\bar{t}^2}. \tag{8.53}$$

This proved formula (8.51), given Lorentz's identification of Σ with Σ'.

Inserting (8.50) and (8.51) into (8.49) then yielded

$$\bar{\mathbf{F}}_{\mathrm{ext}} = m_0(\gamma^3 l, \gamma l, \gamma l)\bar{\mathbf{a}}. \tag{8.54}$$

Stated otherwise, the electromagnetic inertia m of an electron oscillating in the moving system had to depend on the speed of the moving system in the following way:

$$m = (\gamma^3 l, \gamma l, \gamma l)m_0. \tag{8.55}$$

By comparing this result with (8.48), he derived a condition for l:

$$\frac{\mathrm{d}(\gamma l|\mathbf{u}|)}{\mathrm{d}|\mathbf{u}|} = \gamma^3 l. \tag{8.56}$$

41 Cf. equation (8.33).

42 Lorentz 1916, pp. 202–3, 327–8.

Since

$$\frac{d(\gamma|\mathbf{u}|)}{d|\mathbf{u}|} = \gamma^3, \tag{8.57}$$

he arrived at the formula

$$\frac{dl}{d|\mathbf{u}|} = 0, \tag{8.58}$$

which implied that l had to be a constant. By assumption, $l = 1$ for $|\mathbf{u}| = 0$, so that $l = 1$. Removing the indeterminancy of l in this way, it was found, among other things, that the transformations (8.3) were reduced to the Lorentz transformations. Lorentz himself drew the following conclusion at this point:

> We are therefore led to suppose *that the influence of a translation on the dimensions (of the separate electrons and of a ponderable body as a whole) is confined to those that have the direction of the motion, these becoming γ times smaller than they are in the state of rest.* If this hypothesis is added to those we have already made, we may be sure that the two states, the one in the moving system, the other in the same system while at rest, corresponding as stated above, may both be possible.[43]

This passage was intended to emphasize that by adding the assumption $l = 1$, the improved theory of corresponding states was plausible in the sense that the conjunction of all hypotheses would not contradict any fundamental law of nature (like Lorentz's modified version (8.54) of Newton's second law of motion).

8.8 LORENTZ'S FORCE FORMULA AND RELATIVISTIC MOMENTUM

Before I continue, I would like to point out the equivalence between Lorentz's modification of the second law of motion (8.54)

$$\overline{\mathbf{F}}_{\text{ext}} = (\gamma^3, \gamma, \gamma) m_0 \overline{\mathbf{a}}$$

and the result obtained by differentiating the relativistic formula for the momentum of an electron

$$\mathbf{p} = \gamma m_0 \mathbf{u}. \tag{8.59}$$

[43] Lorentz 1937a, pp. 188–9. I have changed k into γ.

Without loss of generality, setting

$$\mathbf{u} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t}\right) = (u, 0, 0), \tag{8.60}$$

we determine that

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \gamma^3 \frac{u}{c^2} \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} \tag{8.61}$$

by means of the chain rule. In view of

$$\gamma^2 \frac{u^2}{c^2} + 1 = \gamma^2, \tag{8.62}$$

we thus obtain

$$\begin{aligned}
\mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= \frac{\mathrm{d}\gamma}{\mathrm{d}t} m_0 \mathbf{u} + \gamma m_0 \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} \\
&= m_0 \left(\gamma^2 \frac{u^2}{c^2} + 1, \gamma, \gamma\right) \frac{\mathrm{d}^2 \mathbf{r}}{\mathrm{d}t^2} \\
&= (\gamma^3, \gamma, \gamma) m_0 \mathbf{a}
\end{aligned} \tag{8.63}$$

in accordance with (8.54).

8.9 INVARIANCE OF OPTICAL PHENOMENA

Let us now return to Lorentz's 1904 paper to explain how he used his improved theory of corresponding states to derive the invariance of optical phenomena under uniform rectilinear motion for all orders of $\frac{u}{c}$. Since he assumed that the moving system $\overline{\Sigma}$ had a system identical to Σ as its corresponding system Σ', it followed that $\mathbf{p}$, $\mathbf{d}$ and $\mathbf{h}$ would have exactly the same functional relationship to the coordinates (x, y, z, t) as $\mathbf{p}'$, $\mathbf{d}'$ and $\mathbf{h}'$ to (x', y', z', t'). This circumstance allowed him to improve his argument for the invariance of patterns of light and darkness under uniform rectilinear motion from the first order to all orders in $\frac{u}{c}$. Noting that $\mathbf{p}' = \mathbf{0}$, $\mathbf{d}' = \mathbf{0}$ and $\mathbf{h}' = \mathbf{0}$ if and only if $\overline{\mathbf{p}} = \mathbf{0}$, $\overline{\mathbf{d}} = \mathbf{0}$ and $\overline{\mathbf{h}} = \mathbf{0}$, he concluded that the patterns remained invariant and "[i]t will therefore be impossible to detect an influence of the Earth's motion on any optical experiment, made with a terrestrial light source, in which the geometrical distribution of light and darkness is observed."[44] We should emphasize that he still held on to the

[44] Ibid., pp. 189–90.

tacit assumption that the observed positions of dark fringes in the moving frame $\overline{S}$ coincided with those produced by $\overline{\mathbf{p}}$, $\overline{\mathbf{d}}$ and $\overline{\mathbf{h}}$. He then observed that most experiments belonged to this class. Moreover, he noted that his theory could also explain the negative results of Rayleigh and Brace as well as Trouton and Noble.[45]

8.10 COMPLETING THE THEORY OF CORRESPONDING STATES

Lorentz still had to comment on cases in which molecular motion was present. In such cases, he claimed, it was permissible to assume that the "bodies [...] undergo the same deformation as the systems of particles of constant relative position of which alone we have spoken till now."[46] To support his argument, he stated that in all situations where the velocities of the molecular motions in a moving system $\overline{\Sigma}$ were small compared to c, the equations of the fictitious system Σ' would correspond to those of a system at rest. It also followed that it was permissible to assume the relationship (8.51) between the accelerations of the particles that corresponded to each other in $\overline{\Sigma}$ and Σ'. Furthermore, he made the assumption that molecular forces did not depend on the velocities of the molecular motions, but only on the relative position between the particles and the common velocity $\mathbf{u}$. Assuming that these forces were limited to such small distances that differences in local time could be neglected, Lorentz argued that "one of the particles, together with those which lie in its sphere of attraction or repulsion, will form a system which undergoes the often mentioned deformation."[47] That is, due to the negligible differences in local time between a particle P in the moving system $\overline{\Sigma}$ and those located in its sphere of attraction or repulsion $V(P)$, Lorentz was able to approximate the positions that the corresponding particles in Σ' occupied at the time t' corresponding to the local time of the particle P by means of the positions that each particle in Σ' occupied at the local time of its own corresponding particle in $\overline{\Sigma}$. Since the positions of corresponding particles at corresponding times satisfied the relationship (8.3), Lorentz could then apply his general force hypothesis to obtain that the force exerted on a particle in $V(P)$ at time $\overline{t}$ due to its interaction with the other particles in $V(P)$ was related to the force that the corresponding particle P' experienced from the other particles in $V'(P')$ at time t'. The forces

[45] Lorentz 1937a, p. 190.
[46] Ibid., p. 191.
[47] Ibid., p. 191.

in the two systems therefore behaved as in the electrostatic case. This must be the reason why, despite the fact that the particles never reached a state of equilibrium, Lorentz felt justified in assuming the "often mentioned contraction," so that $\overline{\Sigma}(0) = \Sigma'$. However, due to the relation (8.51) obtaining between corresponding accelerations, this was only possible "if we suppose *that the masses of all particles are influenced by a translation to the same degree as the electromagnetic masses of the electrons*."[48] By adding this assumption, Lorentz finally extended his theory to optical phenomena involving molecular motion. Although he did not demonstrate how all mass could be reduced to electromagnetic inertia, this assumption was nevertheless firmly grounded in Abraham's electromagnetic worldview, which sought to base mechanics entirely on electrodynamics.

8.11 CONCLUSION

In 1904 Lorentz published his improved theory of corresponding states, in which the invariance of optical and electrostatic phenomena under uniform rectilinear motion applied to every order of $\frac{u}{c}$. From a modern point of view, however, he did not prove the mathematical invariance of the general Maxwell equations under the Lorentz transformations. Accordingly, the Dutchman still believed that it might be possible to detect motion with respect to the ether. Despite Poincare's physical interpretation of corresponding states in terms of apparent measurements, Lorentz continued to argue in terms of a fictitious system brought to rest. He introduced the concepts of "longitudinal mass" and "transverse mass" of the electron, which could be thought of as the inertia of an electron due to the interaction with its self-field. Finally, he extended his theory to optical phenomena involving molecular motion by assuming that the masses of all particles were affected by translation to the same extent as the electromagnetic masses of electrons.

---

[48] Ibid., p. 191.

CHAPTER 9

Poincaré's Theory of Relativity

Shortly after the publication of Lorentz's 1904 paper, Max Abraham revealed a major flaw in the Dutchman's theory. In 1903, Abraham had noticed that a contractile electron model would be incompatible with the electromagnetic worldview because it required non-electromagnetic adhesive forces to hold the separate parts of the electron charge together. However, such stabilizing forces could be safely neglected in a rigid spherical electron model, since in this case the forces that counterbalanced the Coulomb repulsion did not displace the surface and consequently performed no work. In other words, despite the fact that such forces were also needed in the case of his own spherical electron model, Abraham argued that he could treat the rigidity of the electron as a purely kinematical constraint within the framework of analytical mechanics.[1] In 1905, Abraham extended his criticism of Lorentz's contractile electron model by making some calculations that proved that the model was inconsistent. As we will see in this chapter, Poincaré both removed this inconsistency by adding an additional potential to Lorentz's model and demonstrated that the resulting theory satisfied what he then called the postulate of relativity.

9.1 ABRAHAM'S 1905 CRITICISM OF LORENTZ'S ELECTRON MODEL

Before analyzing Poincaré's complicated treatise in detail, I would like to briefly address Abraham's 1905 criticism of Lorentz's electron model. In 1903, assuming a charge distribution in uniform rectilinear motion with velocity $\mathbf{u}$, he was able to identify the Lagrangian of the generated electromagnetic field as

$$L = T - U, \tag{9.1}$$

where T denoted the field's magnetic energy and U its electric energy:

$$\begin{aligned} T &= \frac{1}{2}\int h^2\,\mathrm{d}\tau, \\ U &= \frac{1}{2}\int d^2\,\mathrm{d}\tau.^2 \end{aligned}$$

[1] Abraham 1903, pp. 108–9.

[2] Ibid., pp. 128, 143. Abraham expressed his 1903 and 1905 results in Gaussian units; for comparison with Lorentz's 1904 article, I recast them in Heaviside–Lorentz units.

This meant that, in the framework of Lagrangian mechanics, the magnetic and electrical energy of the field corresponded to its kinetic and potential energy respectively. Next, Abraham was able to calculate the momentum of the field $\mathbf{G}$ and its total electromagnetic energy E according to the standard formulas of Lagrangian mechanics:

$$\mathbf{G} = \frac{\partial L}{\partial \mathbf{u}} = \frac{1}{c}\int \mathbf{d}\times\mathbf{h}\,\mathrm{d}\tau, \tag{9.2}$$

$$E = \mathbf{G}\cdot\mathbf{u} - L = \frac{1}{2}\int d^2 + h^2\,\mathrm{d}\tau, \tag{9.3}$$

where

$$\frac{\partial}{\partial \mathbf{u}} = \left(\frac{\partial}{\partial u_1}, \frac{\partial}{\partial u_2}, \frac{\partial}{\partial u_3}\right) = \nabla_{\mathbf{u}}.$$ [3]

In 1905, he went on to demonstrate that the momentum of Lorentz's electron model in terms of its Lagrangian, $\mathbf{G} = \frac{\partial L}{\partial \mathbf{u}}$, differed from $\mathbf{G} = \frac{1}{c}\int \mathbf{d}\times\mathbf{h}\,\mathrm{d}\tau$. Applying the latter formula, in accordance with equation (8.45) for $l = 1$, he obtained

$$\mathbf{G} = \frac{4}{3}\frac{E'\gamma}{c^2}\mathbf{u} = \frac{e^2}{6\pi c^2 R}\gamma\mathbf{u}. \tag{9.4}$$

When calculating the Lagrangian L of the flattened electron using its corresponding spherical state in S', he showed that

$$L = -\frac{E'}{\gamma} = -\frac{e^2}{8\pi R\gamma}. \tag{9.5}$$

The differentiation of (9.5) then resulted in

$$\frac{\partial L}{\partial \mathbf{u}} = \frac{e^2}{8\pi c^2 R}\gamma\mathbf{u}. \tag{9.6}$$

When he finally compared (9.4) with (9.6), he obtained the inconsistency.[4]

Let us reconstruct Abraham's derivation of (9.6) using our account of Lorentz's theory of corresponding states. The moving ellipsoidal electron corresponds to a static sphere in S'. It follows that $\mathbf{h}' = \mathbf{0}$. If we insert the latter into (8.5) for $l = 1$, we obtain

$$\begin{aligned} \mathbf{d} &= \left(d'_1, \gamma d'_2, \gamma d'_3\right), \\ \mathbf{h} &= \left(0, \tfrac{u\gamma}{c} d'_3, \tfrac{u\gamma}{c} d'_2\right). \end{aligned} \tag{9.7}$$

[3] Ibid., p. 145.

[4] Abraham 1905, pp. 203–4.

The substitution of (9.7) into the expression for the Lagrangian then results in (9.5):

$$L = \frac{1}{2}\int h^2 - d^2 \,\mathrm{d}\tau = \frac{1}{2\gamma}\int \frac{u^2}{c^2}\gamma^2 d'^2 - \gamma^2 d'^2 \,\mathrm{d}\tau'$$
$$= -\frac{1}{\gamma}\left(\frac{1}{2}\int d'^2 \,\mathrm{d}\tau'\right) = -\frac{E'}{\gamma} = -\frac{e^2}{8\pi R\gamma}.$$

The calculation of the gradient of $\frac{1}{\gamma}$ with respect to $\mathbf{u}$ gives

$$\frac{\partial}{\partial \mathbf{u}}\left(\frac{1}{\gamma}\right) = -\frac{\gamma}{c^2}\mathbf{u}.$$

Since E' is independent of $\mathbf{u}$, (9.6) follows:

$$\frac{\partial L}{\partial \mathbf{u}} = -E'\frac{\partial}{\partial \mathbf{u}}\left(\frac{1}{\gamma}\right) = \frac{E'}{c^2}\gamma\mathbf{u} = \frac{e^2}{8\pi c^2 R}\gamma\mathbf{u}.$$

Abraham then made clear why this inconsistency in Lorentz's electron model was due to a missing non-electromagnetic energy term.[5] He calculated the total electromagnetic energy E of the electron by substituting (9.5) and (9.4) into (9.3) and obtained

$$E = \frac{E'}{\gamma} + \frac{4}{3}\frac{u^2\gamma}{c^2}E'.$$

Next, he observed that "[t]he increase in electromagnetic energy of Lorentz's electron due to acceleration is greater than the work done by external force."[6] This meant that the (instantaneous) work-energy theorem

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \mathbf{F}_{\text{ext}}\cdot\mathbf{u}$$

did not hold for Lorentz's electron model, and it was therefore necessary to endow the electron with a non-electromagnetic internal energy E_c in order to retain it:

$$\frac{\mathrm{d}(E + E_c)}{\mathrm{d}t} = \mathbf{F}_{\text{ext}}\cdot\mathbf{u}. \quad (9.8)$$

[5] Abraham 1905, pp. 205–8.

[6] Ibid., p. 205. My translation.

To determine E_c, Abraham remarked that $\mathbf{F}_{\text{ext}}$ was equal to the time derivative of the electromagnetic momentum. We can see this by substituting $\mathbf{F}_{\text{ext}}$ into

$$\mathbf{F}_{\text{tot}} = \mathbf{F}_{\text{ext}} + \mathbf{F}_{\text{self}} = \mathbf{0}$$

and expressing $\mathbf{F}_{\text{self}}$ by $-\frac{\mathrm{d}\mathbf{G}}{\mathrm{d}t}$:

$$\mathbf{F}_{\text{ext}} = -\mathbf{F}_{\text{self}} = \frac{\mathrm{d}\mathbf{G}}{\mathrm{d}t}.^{7} \tag{9.9}$$

The combination of (9.8) and (9.9) then resulted in

$$\frac{\mathrm{d}(E + E_c)}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{G}}{\mathrm{d}t} \cdot \mathbf{u}. \tag{9.10}$$

Since

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{G} \cdot \mathbf{u}) = \frac{\mathrm{d}\mathbf{G}}{\mathrm{d}t} \cdot \mathbf{u} + \mathbf{G} \cdot \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t},$$

Abraham could rewrite (9.10) as

$$\mathbf{G} \cdot \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{G} \cdot \mathbf{u} - E - E_c).$$

Expressing L by means of (9.3) as

$$L = \mathbf{G} \cdot \mathbf{u} - E,$$

he obtained

$$\mathbf{G} \cdot \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{\mathrm{d}(L - E_c)}{\mathrm{d}t}. \tag{9.11}$$

Since he investigated the increase in the energy of the electron as a result of acceleration along a fixed direction of motion, it followed that both $\mathbf{G}$ and $\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}$ were parallel to this direction. From this followed

$$\left| \mathbf{G} \cdot \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} \right| = |\mathbf{G}| \frac{\mathrm{d}|\mathbf{u}|}{\mathrm{d}t}. \tag{9.12}$$

Rewriting the time derivative of $L - E_c$ as

$$\frac{\mathrm{d}(L - E_c)}{\mathrm{d}t} = \frac{\mathrm{d}(L - E_c)}{\mathrm{d}|\mathbf{u}|} \frac{\mathrm{d}|\mathbf{u}|}{\mathrm{d}t} \tag{9.13}$$

[7] $\mathbf{F}_{\text{tot}} = \mathbf{0}$ followed from $m_N = 0$.

and combining (9.11), (9.12) and (9.13) then resulted in

$$|\mathbf{G}| = \frac{\mathrm{d}(L - E_c)}{\mathrm{d}|\mathbf{u}|}. \tag{9.14}$$

Differentiating L with respect to $|\mathbf{u}|$,

$$\frac{\mathrm{d}L}{\mathrm{d}|\mathrm{u}|} = -E' \frac{\mathrm{d}}{\mathrm{d}|\mathbf{u}|} \left(\sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}} \right) = \frac{E'}{c^2} \gamma |\mathbf{u}|, \tag{9.15}$$

and calculating the norm of $\mathbf{G}$ using (9.4),

$$|\mathbf{G}| = \frac{4}{3} \frac{E'}{c^2} \gamma |\mathbf{u}|, \tag{9.16}$$

it followed from (9.14), (9.15) and (9.16) that

$$\frac{\mathrm{d}E_c}{\mathrm{d}|\mathbf{u}|} = \frac{\mathrm{d}L}{\mathrm{d}|\mathbf{u}|} - |\mathbf{G}| = -\frac{1}{3} \frac{E'}{c^2} \gamma |\mathbf{u}| = -\frac{1}{3} \frac{\mathrm{d}L}{\mathrm{d}|\mathbf{u}|}. \tag{9.17}$$

Integrating (9.17) he was finally able to express E_c as

$$E_c = E_c' - \frac{1}{3}(L - L') = E_c{}' + \frac{1}{3}\left(\frac{E'}{\gamma} - E'\right) = E_c' - \frac{e^2}{24\pi R}\left(1 - \frac{1}{\gamma}\right), \tag{9.18}$$

where E_c' was the complementary energy of the electron at rest.[8] Abraham had therefore fixed E_c up to its value for the stationary electron. On the significance of the need to add a non-electromagnetic energy term to Lorentz's electron model, he remarked that this negated a purely electromagnetic worldview. For it meant that, contrary to Abraham's rigid spherical electron model, part of the inertia of Lorentz's electron was non-electromagnetic in origin.[9]

9.2 DYNAMICS OF THE ELECTRON 1905-6

Poincaré was well aware of Abraham's criticism of Lorentz's electron model. But instead of criticizing Lorentz's shortcomings, Poincaré praised Lorentz's 1904 theory because it enabled him to fulfill the relativity postulate in a rigorous way by achieving exact invariance of electrodynamic

[8] Abraham 1905, p. 207.
[9] Ibid., p. 208.

phenomena with respect to uniform rectilinear motion. Lorentz attempted, at least to some extent, to do justice to Poincaré's criticism by formulating a new version of his theory in which the invariance of optical phenomena held at every order in $\frac{u}{c}$. For other phenomena, as we have already noted, Lorentz still believed that it might be possible to detect motion with respect to the ether. The latter was enough to arouse Poincaré's enthusiasm. The results of his efforts were first published in 1905 as a short note entitled "Sur la Dynamique de l'Électron" and the following year as a comprehensive article under the same title. In the introduction to this article, Poincaré reflected on the epistemological status of the null results of all ether drift experiments:

> It seems that this impossibility of demonstrating an experimental evidence for absolute motion of the Earth is a general law of nature; we are naturally led to admit this law, which we will call the *Postulate of Relativity* and admit it without restriction. This postulate, which is up to now in accord with experiments, may be either confirmed or disproved later by more precise experiments, it is in any case interesting to see which consequences follow from it.[10]

Although he called it the postulate of relativity, I agree with Miller that it is identical with the principle of relativity,[11] i.e. it corresponds to what Poincaré in 1900 called the principle of relative motion applied to matter alone. According to Miller, Poincaré's willingness to discard this principle in the future on the basis of more precise experiments "constitutes further proof for my assertion that the principle of relativity is not a convention."[12] But despite its controversial status, it seems to me more reasonable to argue the opposite. The French mathematician tells us in the quote that the principle is a general law of nature, which he was going to *admit without restrictions* in order to see what consequences follow from it. These restrictions referred to the ether and to earlier constraints on the validity of the principle up to a certain order of approximation. In other words, for the sake of argument, Poincaré provisionally decided to elevate the principle of relativity to the status of a convention in the context of electrodynamics. His cautious optimism was based on Lorentz's theory from 1904:

---

[10] Poincaré 2021, p. 129.
[11] Miller 1986, p. 246.
[12] Ibid., p. 246.

> [T]he hypothesis of a contraction [. . .] would give an account of the experiment of Michelson and all those which were carried out up to now. The hypothesis would become insufficient, however, if one were to assume the postulate of relativity in all its generality. Lorentz sought to supplement and modify it in order to put it in perfect agreement with this postulate. He succeeded in doing so in his article entitled *Electromagnetic phenomena in a system moving with any velocity smaller than that of light.* [. . .] The idea of Lorentz can be summarized as follows: if we can bring the whole system to a common translation, without modification of any of the apparent phenomena, it is because the equations of electromagnetic medium are not altered by certain transformations, which we will call *LORENTZ transformation*; two systems, one motionless, the other in translation, thus becoming exact images of one another.[13]

If we compare this statement with our findings from the previous chapter, it becomes clear that Poincaré exaggerated Lorentz's achievement with regard to the exact invariance of all electrodynamic phenomena under rectilinear motion. In 1904, Lorentz only proved the invariance of optical phenomena for any order. For other phenomena, he still believed that it might be possible to detect motion with respect to the ether. It follows that the Dutchman still regarded the principle of relativity as a consequence within a theory for particular electrodynamic phenomena. For Poincaré, however, the principle of relativity was fundamental and applied to all phenomena:

> The hypothesis of Lorentz and Fitz-Gerald will appear most extraordinary at first sight. All that can be said in its favour for the moment is that it is merely the immediate interpretation of Michelson's experimental result, if we *define* distances by the time taken by light to traverse them.
>
> However that be, it is impossible to escape the impression that the Principle of Relativity is a general law of Nature, and that we shall never succeed, by any imaginable method, in demonstrating any but relative velocities; and by this I mean not merely in relation to the ether, but the velocities to bodies in relation to each other. So many different experiments have given similar results that we cannot but

[13] Poincaré 2021, pp. 129–130.

> feel tempted to attribute to this Principle of Relativity a value comparable, for instance, to that of the Principle of Equivalence.[14]

Poincaré's strategy was therefore to reverse the logic of Lorentz's approach, as he elaborated in a section entitled "Conséquences du Principe de la Relativité" (Consequences of the Principle of Relativity) in his book *La Science et Méthode* (1908):

> [T]he deformation of the electrons seems extremely hypothetical. But the matter can be presented differently, so as to avoid taking this hypothesis of deformation as the basis of the argument. Let us imagine the electrons as material points, and enquire how their mass ought to vary as a function of the velocity so as not to violate the Principle of Relativity. Or rather let us further enquire what should be their acceleration under the influence of an electric or magnetic field, so that the principle should not be violated and that we should return to the ordinary laws when we imagine the velocity very low. We shall find that the variations of this mass or of these accelerations must occur *as if* the electron underwent Lorentz's deformation.[15]

Neither in "Sur la Dynamique de l'Électron" nor in *Science et Méthode* did Poincaré clarify the physical interpretation of the Lorentz transformations in great detail. As we shall see, his Sorbonne lectures of 1906 support the reading that he still understood the transformed variables in S' as corresponding to the apparent measurements of a moving observer.[16]

9.3 RECTIFYING INVARIANCE

The starting point of Poincaré's investigation was Lorentz's 1904 version of Maxwell's equations, adapted to the ether frame S and written in Heaviside-Lorentz units, with the units of length and time chosen so that the speed of

[14] Poincaré 2003, p. 221.
[15] Ibid., pp. 227–228.
[16] We will return to this question in section 9.10.

light was equal to 1:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0, \qquad \frac{\partial\varphi}{\partial t} + \nabla\cdot\mathbf{A} = 0, \tag{9.19}$$

$$\Box\varphi = -\rho, \qquad \Box\mathbf{A} = -\rho\mathbf{v}, \tag{9.20}$$

$$\mathbf{h} = \nabla\times\mathbf{A}, \qquad -\frac{\partial\mathbf{A}}{\partial t} - \nabla\varphi = \mathbf{d}, \tag{9.21}$$

$$\nabla\cdot\mathbf{d} = \rho, \qquad \frac{\partial\mathbf{h}}{\partial t} = -\nabla\times\mathbf{d}, \tag{9.22}$$

$$\mathbf{i} = \frac{\partial\mathbf{d}}{\partial t} + \rho\mathbf{v} = \nabla\times\mathbf{h}, \tag{9.23}$$

where **i** denoted the total current and $\Box$ the partial differential operator

$$\Box = \nabla^2 - \frac{\partial^2}{\partial t^2}.$$ [17]

Poincaré introduced the Lorenz force per volume element $d\tau$

$$\mathbf{f} = \rho\mathbf{d} + \rho\mathbf{v}\times\mathbf{h} \tag{9.24}$$

and the Lorentz transformations with scale factor l between the frame S and a frame S' moving with a constant translational speed ϵ with respect to S:

$$t' = \gamma l(t + \epsilon x), \qquad x' = \gamma l(x + \epsilon t),$$
$$y' = ly, \qquad z' = lz, \tag{9.25}$$

where l denoted Lorentz's scale factor.[18]

He then proved the exact invariance of the Lorentz-Maxwell equations under the Lorentz transformations with scale factor. Provided that the form of the field equations and the charge of an electron remained the same under the coordinate transformations (9.25) from S to S', he showed how to improve Lorentz's transformation equation for charge by considering a sphere of volume

$$V_e = \frac{4}{3}\pi r^3, \tag{9.26}$$

[17] Poincaré 2021, pp. 132–3.

[18] Ibid., p. 133. Poincaré's Lorentz transformations with scale factor l make clear that the Frenchman reversed the relative velocity between the frames ($\epsilon = -u$).

which was carried along by an electron with a constant translational velocity $\mathbf{v} = (v_1, v_2, v_3)$. The transformations (9.25), he noted, would change the equation of the moving sphere into an ellipsoid. Evaluating the motion of the ellipsoid for $t' = 0$, it was possible to determine its volume:

$$V_e' = \frac{4}{3}\pi r^3 \frac{l^3}{\gamma(1 + v_1\epsilon)}. \tag{9.27}$$

The conservation of the charge enclosed in the ellipsoid then gave rise to the transformation formula for the new charge density ρ':

$$\rho' = \rho \frac{V_e}{V_e'} = \frac{\gamma}{l^3}(\rho + \epsilon\rho v_1). \tag{9.28}$$

The transformation formula for the new velocity $\mathbf{v}' = (v_1', v_2', v_3')$ was obtained according to the calculations

$$\begin{aligned} v_1' &= \frac{\mathrm{d}x'}{\mathrm{d}t'} = \frac{\mathrm{d}(x + \epsilon t)}{\mathrm{d}(t + \epsilon x)} = \frac{v_1 + \epsilon}{1 + \epsilon v_1}, \\ v_2' &= \frac{\mathrm{d}y'}{\mathrm{d}t'} = \frac{\mathrm{d}y}{\gamma \mathrm{d}(t + \epsilon x)} = \frac{v_2}{\gamma(1 + \epsilon v_1)}, \\ v_3' &= \frac{\mathrm{d}z'}{\mathrm{d}t'} = \frac{\mathrm{d}z}{\gamma \mathrm{d}(t + \epsilon x)} = \frac{v_3}{\gamma(1 + \epsilon v_1)}. \end{aligned} \tag{9.29}$$

Poincaré's next step was to calculate the transformation formulas for the convection current using (9.28) and (9.29):

$$\rho' v_1' = \frac{\gamma}{l^3}(\rho v_1 + \epsilon\rho), \qquad \rho' v_2' = \frac{1}{l^3}\rho v_2, \qquad \rho' v_3' = \frac{1}{l^3}\rho v_3. \tag{9.30}$$

Comparing (9.28) and (9.30) with the equivalent transformation equations between S and S' in Lorentz's theory for the charge density and the convection current (obtained using the Galilean rules for the transformation of charge and velocities as well as (8.6) and (8.7)),

$$\rho' = \frac{\rho}{\gamma}, \qquad \rho' v_1' = \frac{\gamma}{l^3}(\rho v_1 + \epsilon\rho), \qquad \rho' v_2' = \frac{1}{l^3}\rho v_2, \qquad \rho' v_3' = \frac{1}{l^3}\rho v_3,$$

Poincaré found that only the charge transformation differed. He went on to note that a major advantage of his transformations was that they satisfied the continuity equation:

$$\frac{\partial \rho'}{\partial t'} + \nabla' \cdot (\rho' \mathbf{v}') = 0. \tag{9.31}$$

To prove (9.31), he introduced an undetermined quantity λ and noted that the determinant D of the Jacobian of the set of quantities

$$t + \lambda\rho, \quad x + \lambda\rho v_1, \quad y + \lambda\rho v_2, \quad z + \lambda\rho v_3 \tag{9.32}$$

with respect to (t, x, y, z) could be rewritten in the form

$$D = D_0 + D_1\lambda + D_2\lambda^2 + D_3\lambda^3 + D_4\lambda^4, \tag{9.33}$$

where

$$D_0 = 1, \quad D_1 = \frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0. \tag{9.34}$$

Introducing $\lambda' = l^4\lambda$, it followed that the set of quantities

$$t' + \lambda'\rho', \quad x' + \lambda'\rho' v_1', \quad y' + \lambda'\rho' v_2', \quad z' + \lambda'\rho' v_3' \tag{9.35}$$

was related to (9.32) by the same relationships as (t', x', y', z') to (t, x, y, z). For example,

$$t' + \lambda'\rho' = \gamma l(t + \epsilon x) + l^4\lambda\frac{\gamma}{l^3}(\rho + \epsilon\rho v_1) = \gamma l(t + \lambda\rho + \epsilon(x + \lambda\rho v_1)).$$

In addition, the determinant D' of the Jacobian of (9.35) with respect to (t', x', y', z') could be written in the form

$$D' = D_0' + D_1'\lambda' + D_2'\lambda'^2 + D_3'\lambda'^3 + D_4'\lambda'^4, \tag{9.36}$$

where

$$D_0' = 1, \quad D_1' = \frac{\partial\rho'}{\partial t'} + \nabla'\cdot(\rho'\mathbf{v}'). \tag{9.37}$$

He then concluded the proof by noticing that

$$D' = D, \quad D_1' = l^{-4}D_1 = 0.$$ [19]

The last step in the derivation is not quite trivial, so let us add more details. Since (9.35) was related to (9.32) as (t', x', y', z') to (t, x, y, z), we have

$$\frac{\partial(t' + \lambda'\rho', \ldots, z' + \lambda'\rho' v_3')}{\partial(t + \lambda\rho, \ldots, z + \lambda\rho v_3)} = \frac{\partial(t', x', y', z')}{\partial(t, x, y, z)}$$

[19] Poincaré 2021, pp. 134–5.

and therefore

$$\begin{aligned}
D' &= \left| \frac{\partial(t'+\lambda'\rho', \ldots, z'+\lambda'\rho' v_3')}{\partial(t',x',y',z')} \right| \\
&= \left| \frac{\partial(t'+\lambda'\rho', \ldots, z'+\lambda'\rho' v_3')}{\partial(t+\lambda\rho, \ldots, z+\lambda\rho v_3)} \cdot \frac{\partial(t+\lambda\rho, \ldots, z+\lambda\rho v_3)}{\partial(t,x,y,z)} \cdot \frac{\partial(t,x,y,z)}{\partial(t',x',y',z')} \right| \\
&= \left| \frac{\partial(t'+\lambda'\rho', \ldots, z'+\lambda'\rho' v_3')}{\partial(t+\lambda\rho, \ldots, z+\lambda\rho v_3)} \right| \cdot \left| \frac{\partial(t+\lambda\rho, \ldots, z+\lambda\rho v_3)}{\partial(t,x,y,z)} \right| \cdot \left| \frac{\partial(t,x,y,z)}{\partial(t',x',y',z')} \right| \\
&= \left| \frac{\partial(t',x',y',z')}{\partial(t,x,y,z)} \right| \cdot D \cdot \left| \frac{\partial(t,x,y,z)}{\partial(t',x',y',z')} \right| \\
&= D \cdot \left| \frac{\partial(t,x,y,z)}{\partial(t',x',y',z')} \right|^{-1} \cdot \left| \frac{\partial(t,x,y,z)}{\partial(t',x',y',z')} \right| \\
&= D.
\end{aligned}$$

We now compare (9.33) and (9.36) by applying $D' = D$ and $\lambda' = l^4\lambda$:

$$\begin{aligned}
D' &= D_0' + D_1' l^4\lambda + D_2' l^8\lambda^2 + D_3' l^{12}\lambda^3 + D_4' l^{16}\lambda^4 \\
&= D_0 + D_1\lambda + D_2\lambda^2 + D_3\lambda^3 + D_4\lambda^4.
\end{aligned}$$

This results in

$$D_0' = D_0, \quad D_1' l^4 = D_1. \tag{9.38}$$

The continuity equation (9.31) now follows by comparing (9.34), (9.37) and (9.38).

Poincaré went on to introduce primed versions of scalar and vector potentials (9.20) in accordance with the principle of relativity:

$$\Box'\varphi' = -\rho', \quad \Box'\mathbf{A}' = -\rho'\mathbf{v}'.^{20} \tag{9.39}$$

Inverting the Lorentz transformations with scale factor,

$$\begin{aligned}
t &= \frac{\gamma}{l}(t' - \epsilon x'), & x &= \frac{\gamma}{l}(x' - \epsilon t'), \\
y &= \frac{1}{l}y', & z &= \frac{1}{l}z',
\end{aligned} \tag{9.40}$$

[20] Ibid., p. 135.

the relationships between primed and unprimed temporal and spatial derivatives could easily be calculated using the chain rule:

$$\frac{\partial}{\partial t'} = \frac{\gamma}{l}\left(\frac{\partial}{\partial t} - \epsilon\frac{\partial}{\partial x}\right), \qquad \frac{\partial}{\partial x'} = \frac{\gamma}{l}\left(\frac{\partial}{\partial x} - \epsilon\frac{\partial}{\partial t}\right),$$
$$\frac{\partial}{\partial y'} = \frac{1}{l}\frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial z'} = \frac{1}{l}\frac{\partial}{\partial z}.$$

Thus,

$$\Box' = \nabla'^2 - \frac{\partial^2}{\partial t'^2} = l^{-2}\left(\nabla^2 - \frac{\partial^2}{\partial t^2}\right) = l^{-2}\Box. \tag{9.41}$$

Without further ado, he then wrote down the transformation equations for the primed potentials:

$$\varphi' = \frac{\gamma}{l}(\varphi + \epsilon A_1), \qquad A_1' = \frac{\gamma}{l}(A_1 + \epsilon\varphi),$$
$$A_2' = \frac{1}{l}A_2, \qquad A_3' = \frac{1}{l}A_3.^{21} \tag{9.42}$$

The most common way to obtain these equations would be to substitute unprimed for primed quantities in (9.39). For example, by using (9.28) and (9.30),

$$\begin{aligned}
\Box'\varphi' &= -\rho' \\
\Rightarrow l^{-2}\Box\varphi' &= -\frac{\gamma}{l^3}(\rho + \epsilon\rho v_1) \\
\Rightarrow \Box\varphi' &= \frac{\gamma}{l}(\Box\varphi + \epsilon\Box A_1) \\
\Rightarrow \Box\varphi' &= \Box\frac{\gamma}{l}(\varphi + \epsilon A_1).
\end{aligned}$$

Therefore, it is possible to fix the transformation for the scalar φ' by (9.42). Similarly,

$$\begin{aligned}
\Box' A_1' &= -\rho' v_1' \\
\Rightarrow l^{-2}\Box A_1' &= -\frac{\gamma}{l^3}(\rho v_1 + \epsilon\rho) \\
\Rightarrow \Box A_1' &= \frac{\gamma}{l}(\Box A_1 + \epsilon\Box\varphi) \\
\Rightarrow \Box A_1' &= \Box\frac{\gamma}{l}(A_1 + \epsilon\varphi).
\end{aligned}$$

[21] Poincaré 2021, p. 135.

However, Poincaré's juxtaposition of the transformation equations for scalar and vector potential (9.39) suggests a different method. We know from (9.28) and (9.30) that the set of quantities $(\rho', \rho' v_1', \rho' v_2', \rho' v_3')$ has the same relationship to $(\rho, \rho v_1, \rho v_2, \rho v_3)$ as (t', x', y', z') has to (t, x, y, z); i.e. they are related by the Lorentz transformations with scale factor (9.25). It thus follows from (9.39) and (9.41) that it is possible to prescribe the transformations for the scalar and vector potential by (9.42).

In the case of the field transformations, there was no such shortcut. Introducing primed versions of the fields (9.21) in accordance with the principle of relativity,

$$\mathbf{h}' = \nabla' \times \mathbf{A}', \quad \mathbf{d}' = -\frac{\partial \mathbf{A}'}{\partial t'} - \nabla' \varphi', \tag{9.43}$$

Poincaré concluded without proof that they had to obey the following transformations:

$$d_1' = \frac{1}{l^2} d_1, \qquad d_2' = \frac{\gamma}{l^2}(d_2 + \epsilon h_3), \qquad d_3' = \frac{\gamma}{l^2}(d_3 - \epsilon h_2), \tag{9.44}$$

$$h_1' = \frac{1}{l^2} h_1, \qquad h_2' = \frac{\gamma}{l^2}(h_2 - \epsilon d_3), \qquad h_3' = \frac{\gamma}{l^2}(h_3 + \epsilon d_2).^{22} \tag{9.45}$$

Let us demonstrate how to obtain the first rule:

$$\begin{aligned}
d_1' &= -\frac{\partial A_1'}{\partial t'} - \frac{\partial \varphi'}{\partial x'} \\
&= \frac{\gamma^2}{l^2}\left(-\frac{\partial}{\partial t} + \epsilon \frac{\partial}{\partial x}\right)(A_1 + \epsilon\varphi) + \frac{\gamma^2}{l^2}\left(-\frac{\partial}{\partial x} + \epsilon \frac{\partial}{\partial t}\right)(\varphi + \epsilon A_1) \\
&= \frac{\gamma^2}{l^2}\left(-\frac{\partial}{\partial t} + \epsilon^2 \frac{\partial}{\partial t}\right) A_1 + \frac{\gamma^2}{l^2}\left(\epsilon^2 \frac{\partial}{\partial x} - \frac{\partial}{\partial x}\right)\varphi \\
&= -\frac{\gamma^2}{l^2}(1-\epsilon^2)\frac{\partial A_1}{\partial t} - \frac{\gamma^2}{l^2}(1-\epsilon^2)\frac{\partial \varphi}{\partial x} \\
&= -\frac{1}{l^2}\frac{\partial A_1}{\partial t} - \frac{1}{l^2}\frac{\partial \varphi}{\partial x} \\
&= \frac{1}{l^2} d_1.
\end{aligned}$$

Next, he noted that the primed version of the second equation in (9.19),

$$\frac{\partial \varphi'}{\partial t'} + \nabla' \cdot \mathbf{A}' = 0, \tag{9.46}$$

[22] Ibid., p. 135.

followed from (9.39). It would be possible to establish (9.46) by substitution. Poincaré's remark that (9.46) resembled the continuity condition[23] strongly suggests that he had a proof by means of an undetermined quantity λ in mind. To achieve the latter, we note that the determinant D of the Jacobian of the set of quantities

$$t+\lambda\varphi, \quad x+\lambda A_1, \quad y+\lambda A_2, \quad z+\lambda A_3 \tag{9.47}$$

with respect to (t, x, y, z) can be rewritten in the form

$$D = D_0 + D_1\lambda + D_2\lambda^2 + D_3\lambda^3 + D_4\lambda^4,$$

where

$$D_0 = 1, \quad D_1 = \frac{\partial\varphi}{\partial t} + \nabla\cdot\mathbf{A} = 0.$$

From the introduction of $\lambda' = l^2\lambda$ it follows that the set of quantities

$$t'+\lambda'\varphi', \quad x'+\lambda' A_1', \quad y'+\lambda' A_2', \quad z'+\lambda' A_3' \tag{9.48}$$

is related to (9.47) by the Lorentz transformations with scale factor l. For example,

$$\begin{aligned} t'+\lambda'\varphi' &= \gamma l(t+\epsilon x) + \lambda l^2\frac{\gamma}{l}(\varphi+\epsilon A_1) \\ &= \gamma l(t+\lambda\varphi+\epsilon(x+\lambda A_1)) \end{aligned}$$

and

$$\begin{aligned} x'+\lambda' A_1' &= \gamma l(x+\epsilon t) + \lambda l^2\frac{\gamma}{l}(A_1+\epsilon\varphi) \\ &= \gamma l(x+\lambda A_1+\epsilon(t+\lambda\varphi)). \end{aligned}$$

In addition, the determinant D' of the Jacobian of (9.48) with respect to (t', x', y', z') can be written in the form

$$D' = D_0' + D_1'\lambda' + D_2'\lambda'^2 + D_3'\lambda'^3 + D_4'\lambda'^4,$$

where

$$D_0' = 1, \quad D_1' = \frac{\partial\varphi'}{\partial t'} + \nabla'\cdot\mathbf{A}'.$$

[23] Poincaré 2021, p. 135.

We can now complete the proof by establishing

$$D' = D, \quad D_1' = l^{-2} D_1 = 0$$

in a similar way as above.

After the derivation of (9.46), the continuity condition (9.31) and transformation formulas for potentials and field quantities, Poincaré noted that the primed counterparts of (9.22) and (9.23) followed:

$$\begin{aligned} \frac{\partial \mathbf{d}'}{\partial t'} + \rho' \mathbf{v}' &= \nabla' \times \mathbf{h}', \\ \frac{\partial \mathbf{h}'}{\partial t'} &= -\nabla' \times \mathbf{d}', \\ \nabla' \cdot \mathbf{d}' &= \rho'.^{24} \end{aligned}$$

Let us prove the last equation by substitution:

$$\begin{aligned} \nabla' \cdot \mathbf{d}' &= \frac{\partial}{\partial x'} \left(-\frac{\partial A_1'}{\partial t'} - \frac{\partial \varphi'}{\partial x'} \right) + \frac{\partial}{\partial y'} \left(\frac{-\partial A_2'}{\partial t'} - \frac{\partial \varphi'}{\partial y'} \right) + \frac{\partial}{\partial z'} \left(\frac{-\partial A_3'}{\partial t'} - \frac{\partial \varphi'}{\partial z'} \right) \\ &= -\nabla'^2 \varphi' - \frac{\partial}{\partial t'} (\nabla' \cdot \mathbf{A}') \\ &= -\left(\nabla'^2 - \frac{\partial^2}{\partial t'^2} \right) \varphi' - \frac{\partial}{\partial t'} \left(\nabla' \cdot \mathbf{A}' + \frac{\partial \varphi'}{\partial t'} \right) \\ &= -\Box' \varphi' \\ &= \rho'. \end{aligned}$$

In order to obtain the complete invariance of the Lorentz-Maxwell equations under the Lorentz transformations with scale factor l, Poincaré then only had to establish the transformation equations for the Lorentz force. Introducing the primed version of the Lorentz force law (9.24) in accordance with the principle of relativity,

$$\mathbf{f}' = \rho' \mathbf{d}' + \rho' \mathbf{v}' \times \mathbf{h}', \tag{9.49}$$

he noticed that it was possible to obtain the following transformation equations:

$$f_1' = \frac{\gamma}{l^5} (f_1 + \epsilon\, \mathbf{f} \cdot \mathbf{v}), \quad f_2' = \frac{1}{l^5} f_2, \quad f_3' = \frac{1}{l^5} f_3.^{25} \tag{9.50}$$

[24] Ibid., p. 136.
[25] Ibid., p. 136.

Let us demonstrate the first rule:

$$\begin{aligned}
f_1' &= \rho' d_1' + \rho'(v_2' h_3' - v_3' h_2') \\
&= \frac{\gamma}{l^5}(\rho + \epsilon\rho v_1)d_1 + \frac{\gamma}{l^5}\rho(v_2(h_3 + \epsilon d_2) - v_3(h_2 - \epsilon d_3)) \\
&= \frac{\gamma}{l^5}(\rho d_1 + \rho(v_2 h_3 - v_3 h_2)) + \epsilon\rho(d_1 v_1 + d_2 v_2 + d_3 v_3) \\
&= \frac{\gamma}{l^5}(f_1 + \epsilon\rho\, \mathbf{d}\cdot\mathbf{v}) \\
&= \frac{\gamma}{l^5}(f_1 + \epsilon\, \mathbf{f}\cdot\mathbf{v}).
\end{aligned}$$

9.4 PRINCIPLE OF LEAST ACTION AND MECHANICAL EXPLANATION

Poincaré continued his investigation by showing how Lorentz's theory could be formulated within the framework of Lagrangian mechanics. This implied the possibility of a mechanical explanation of electricity (according to his concept of mechanical explanation). In doing so, Poincaré wrote the principle of least action in the form

$$J = \int \frac{d^2}{2} + \frac{h^2}{2} - \mathbf{A}\cdot\mathbf{i}\ \mathrm{d}\tau \mathrm{d}t, \tag{9.51}$$

where the volume elements $\mathrm{d}\tau$ were evaluated over the entire space and the time increments $\mathrm{d}t$ between definite limits, and he constrained his variational principle by the following three equations from (9.21)-(9.23):

$$\mathbf{h} = \nabla \times \mathbf{A}, \tag{9.52}$$

$$\nabla\cdot \mathbf{d} = \rho, \tag{9.53}$$

$$\mathbf{i} = \frac{\partial \mathbf{d}}{\partial t} + \rho\mathbf{v}. \tag{9.54}$$

To complete the demonstration, he therefore had to derive the Lorentz force law (9.24) and the missing Lorentz-Maxwell equations:

$$\nabla \times \mathbf{h} = \frac{\partial \mathbf{d}}{\partial t} + \rho\mathbf{v}, \tag{9.55}$$

$$\nabla \times \mathbf{d} = -\frac{\partial \mathbf{h}}{\partial t}.^{26} \tag{9.56}$$

[26] Taking the gradient of (9.52) implies $\nabla\cdot \mathbf{h} = \mathbf{0}$.

Taking variations gave

$$\delta J = \int \mathbf{d}\cdot\delta\mathbf{d} + \mathbf{h}\cdot\delta\mathbf{h} - \delta\mathbf{A}\cdot\mathbf{i} - \mathbf{A}\cdot\delta\mathbf{i}\ \mathrm{d}\tau\mathrm{d}t. \tag{9.57}$$

Using the constraint $\mathbf{h} = \nabla \times \mathbf{A}$, the admissible variations satisfied

$$\delta\mathbf{h} = \nabla \times \delta\mathbf{A}.$$

Hence,

$$\int \mathbf{h}\cdot\delta\mathbf{h}\,\mathrm{d}\tau = \int \mathbf{h}\cdot(\nabla \times \delta\mathbf{A})\,\mathrm{d}\tau.$$

Rewriting the integrand by means of the vector calculus identity

$$\mathbf{h}\cdot(\nabla \times \delta\mathbf{A}) = \nabla\cdot(\delta\mathbf{A} \times \mathbf{h}) + (\nabla \times \mathbf{h})\cdot\delta\mathbf{A}$$

then resulted in

$$\int \mathbf{h}\cdot\delta\mathbf{h}\,\mathrm{d}\tau = \int \nabla\cdot(\delta\mathbf{A} \times \mathbf{h})\,\mathrm{d}\tau + \int (\nabla \times \mathbf{h})\cdot\delta\mathbf{A}\,\mathrm{d}\tau.$$

By the divergence theorem the first term was equal to a surface integral,

$$\int \nabla\cdot(\delta\mathbf{A} \times \mathbf{h})\,\mathrm{d}\tau = \oint (\delta\mathbf{A} \times \mathbf{h})\cdot\mathrm{d}\sigma,$$

which vanished by supposing that the fields fell off sufficiently fast at infinity. Poincaré could therefore write the variation of J with respect to virtual displacements of $\mathbf{A}$ as

$$\delta J = -\int (\mathbf{i} - \nabla \times \mathbf{h})\cdot\delta\mathbf{A}\ \mathrm{d}\tau\mathrm{d}t\ =\ 0.$$

Setting the coefficients of $\delta\mathbf{A}$ to zero, he obtained

$$\mathbf{i} = \nabla \times \mathbf{h}, \tag{9.58}$$

which implied (9.55). His next step was to substitute (9.58) into the source term of the action:

$$\int \mathbf{A}\cdot\mathbf{i}\,\mathrm{d}\tau = \int \mathbf{A}\cdot(\nabla \times \mathbf{h})\,\mathrm{d}\tau.$$

Applying the identity

$$\nabla\cdot(\mathbf{A}\times\mathbf{h}) = \mathbf{h}\cdot(\nabla\times\mathbf{A}) - \mathbf{A}\cdot(\nabla\times\mathbf{h}),$$

the integrand could be expressed as

$$\mathbf{A}\cdot(\nabla\times\mathbf{h}) = \mathbf{h}\cdot(\nabla\times\mathbf{A}) - \nabla\cdot(\mathbf{A}\times\mathbf{h}).$$

Since $\mathbf{h} = \nabla\times\mathbf{A}$ by (9.52), the first term became

$$\mathbf{h}\cdot(\nabla\times\mathbf{A}) = \mathbf{h}\cdot\mathbf{h} = h^2,$$

which lead to

$$\mathbf{A}\cdot(\nabla\times\mathbf{h}) = h^2 - \nabla\cdot(\mathbf{A}\times\mathbf{h}).$$

Upon integrating this relation over all space, he obtained

$$\int \mathbf{A}\cdot(\nabla\times\mathbf{h})\,\mathrm{d}\tau = \int h^2\,\mathrm{d}\tau - \int \nabla\cdot(\mathbf{A}\times\mathbf{h})\,\mathrm{d}\tau.$$

By the divergence theorem, the last term was equal to a surface integral:

$$\int \nabla\cdot(\mathbf{A}\times\mathbf{h})\,\mathrm{d}\tau = \oint (\mathbf{A}\times\mathbf{h})\cdot\mathrm{d}\sigma.$$

Repeating the assumption that the fields vanished at spatial infinity, this surface term was zero. Therefore,

$$\int \mathbf{A}\cdot\mathbf{i}\,\mathrm{d}\tau = \int h^2\,\mathrm{d}\tau. \tag{9.59}$$

If (9.59) was substituted into (9.51), Poincaré could then finally write his action principle in the form

$$J = \frac{1}{2}\int d^2 - h^2\,\mathrm{d}\tau\mathrm{d}t.^{27} \tag{9.60}$$

The Frenchman went on to analyze the variation of J (9.57) with respect to virtual displacements of $\mathbf{d}$. Since $\mathbf{A}$ and $\mathbf{h}$ were independent of $\mathbf{d}$, the variation of J reduced to

$$\delta J = \int \mathbf{d}\cdot\delta\mathbf{d} - \mathbf{A}\cdot\delta\mathbf{i}\,\mathrm{d}\tau\mathrm{d}t.$$

[27] Unlike Abraham and Lorentz, Poincaré preferred to write the Lagrangian of the electromagnetic field as $U - T$ instead of $T - U$.

Because the variation of J with respect to $\mathbf{d}$ had to satisfy the second constraint (9.53), he added to δJ the variation of (9.53) multiplied by a Lagrange multiplier φ:

$$\delta J = \int \mathbf{d}\cdot\delta\mathbf{d} - \mathbf{A}\cdot\delta\mathbf{i} - \varphi\big(\nabla\cdot\delta\mathbf{d} - \delta\rho\big)\ \mathrm{d}\tau\mathrm{d}t.$$

Using the product rule for the divergence of a scalar times a vector,

$$-\int \varphi\,\nabla\cdot\delta\mathbf{d}\,\mathrm{d}\tau = -\int \nabla\cdot(\varphi\,\delta\mathbf{d})\,\mathrm{d}\tau + \int(\nabla\varphi)\cdot\delta\mathbf{d}\,\mathrm{d}\tau.$$

By the divergence theorem the first term became once again equal to a surface integral $\oint \varphi\,\delta\mathbf{d}\cdot\mathrm{d}\sigma$, which vanished at infinity under the standard boundary conditions. Hence

$$-\int \varphi\,\nabla\cdot\delta\mathbf{d}\,\mathrm{d}\tau = \int(\nabla\varphi)\cdot\delta\mathbf{d}\,\mathrm{d}\tau.$$

Therefore,

$$\delta J = \int (\mathbf{d} + \nabla\varphi)\cdot\delta\mathbf{d} - \mathbf{A}\cdot\delta\mathbf{i} + \varphi\,\delta\rho\,\mathrm{d}\tau\mathrm{d}t. \tag{9.61}$$

Poincaré still needed to impose the third constraint $\mathbf{i} = \frac{\partial\mathbf{d}}{\partial t} + \rho\mathbf{v}$, whose variation was

$$\delta\mathbf{i} \ = \ \frac{\partial}{\partial t}(\delta\mathbf{d}) + \delta(\rho\mathbf{v}).$$

Substituting this into the $\mathbf{A}$–term of (9.61) gave

$$-\mathbf{A}\cdot\delta\mathbf{i} = -\mathbf{A}\cdot\frac{\partial}{\partial t}(\delta\mathbf{d}) - \mathbf{A}\cdot\delta(\rho\mathbf{v}).$$

By integrating the first term by parts in time, he obtained

$$-\int \mathbf{A}\cdot\frac{\partial}{\partial t}(\delta\mathbf{d})\,\mathrm{d}t = \int \frac{\partial\mathbf{A}}{\partial t}\cdot\delta\mathbf{d}\,\mathrm{d}t.$$

Collecting all contributions, the variation of J took the form

$$\delta J = \int \left(\mathbf{d} + \frac{\partial\mathbf{A}}{\partial t} + \nabla\varphi\right)\cdot\delta\mathbf{d} + \varphi\delta\rho - \mathbf{A}\cdot\delta(\rho\mathbf{v})\ \mathrm{d}\tau\mathrm{d}t \ = \ 0.$$

Setting the coefficients of $\delta\mathbf{d}$ equal to zero then yielded

$$\mathbf{d} = -\frac{\partial\mathbf{A}}{\partial t} - \nabla\varphi. \tag{9.62}$$

Poincaré did not even bother to mention it, but if we take the curl of both sides of (9.62), we get (9.56) due to (9.52):

$$\nabla \times \mathbf{d} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{A}) = -\frac{\partial \mathbf{h}}{\partial t}.$$

In addition, the Lagrange multiplier corresponded to the scalar potential.

After obtaining the relation (9.62), the previous expression for δJ reduced to

$$\delta J = \int \varphi\, \delta\rho - \mathbf{A}\cdot\delta(\rho\mathbf{v})\, \mathrm{d}\tau. \tag{9.63}$$

In this form δJ depended only on terms involving variations of the charge density ρ and the convection current $\rho\mathbf{v}$. In order to make the mechanical analogy with D'Alembert's principle explicit, Poincaré expressed these variations in terms of a virtual displacement $\delta\mathbf{r}$ of the charge distribution itself:

$$\delta\rho = -\nabla\cdot\,(\rho\,\delta\mathbf{r}), \qquad \delta(\rho\mathbf{v}) = \frac{\partial}{\partial t}(\rho\,\delta\mathbf{r}) - \nabla\times(\rho\,\delta\mathbf{r}\times\mathbf{v}).$$

Substituting these expressions into (9.63), he was able to rewrite (9.63) as

$$\delta J = \int \rho\,\big(-\mathbf{d} - \mathbf{v}\times\mathbf{h}\big)\cdot\delta\mathbf{r}\, \mathrm{d}\tau\, \mathrm{d}t. \tag{9.64}$$

This expression had precisely the structure of D'Alembert's principle of virtual work,

$$\delta J = -\int \mathbf{f}\cdot\delta\mathbf{r}\ \mathrm{d}\tau\mathrm{d}t, \tag{9.65}$$

where $\mathbf{f}$ denoted the force density acting on the displaced matter. Equating the coefficients of $\delta\mathbf{r}$ in (9.64) and (9.65) therefore yielded the Lorentz force law

$$\mathbf{f} = \rho\,(\mathbf{d} + \mathbf{v}\times\mathbf{h})$$

and completed his demonstration.

9.5 PRINCIPLE OF LEAST ACTION AND INVARIANCE

Poincaré opened the third paragraph with the statement that "the principle of least action gives us the reason for the success of the Lorentz transformation."[28] By this he meant that the principle of least action could account

[28] Poincaré 2021, p. 142.

for the invariance of the Lorentz-Maxwell equations under Lorentz transformations with scale factor l. To justify his announcement, he returned to his variational principle in the form

$$J = \frac{1}{2}\int d^2 - h^2 \,\mathrm{d}\tau\mathrm{d}t.$$

He noted that

$$\mathrm{d}\tau'\mathrm{d}t' = l^4\mathrm{d}\tau\mathrm{d}t, \tag{9.66}$$

which follows from the Jacobian determinant

$$\left|\frac{\partial(x', y', z', t')}{\partial(x, y, z, t)}\right| = l^4$$

together with the general rule

$$\mathrm{d}\tau'\mathrm{d}t' = \left|\frac{\partial(x', y', z', t')}{\partial(x, y, z, t)}\right| \mathrm{d}\tau\mathrm{d}t.$$

Defining J' in accordance with the principle of relativity,

$$J' = \frac{1}{2}\int d'^2 - h'^2 \,\mathrm{d}\tau'\mathrm{d}t',$$

it followed from (9.66) together with the field transformation formulas (9.44) and (9.45) that

$$J' = J. \tag{9.67}$$

In other words, the the action J was invariant under Lorentz transformations with scale factor l, This invariance implied that the variational method would produce (Lagrange) equations of the same form in both frames S and S'. The Lagrangian formulation of electrodynamics therefore provided a direct explanation for the Lorentz invariance of the Lorentz-Maxwell equations: the invariance of the action ($J = J'$) ensured the invariance of the equations derived from it. We will see later how Poincaré applied this insight in his investigation of Lorentz's electron model.

In the remainder of §3 Poincaré showed how to obtain the transformation formulas (9.50) for the Lorentz force by varying the action. From (9.67) it followed that

$$\delta J = -\int \mathbf{f}\cdot\delta\mathbf{r}\,\mathrm{d}\tau\mathrm{d}t$$

had to be equal to

$$\delta J' = -\int \mathbf{f}' \cdot \delta \mathbf{r}' \, \mathrm{d}\tau' \mathrm{d}t'.$$

The substitution of (9.66) in (9.5) resulted in

$$\delta J = -\int \frac{1}{l^4} \mathbf{f} \cdot \delta \mathbf{r} \, \mathrm{d}\tau' \mathrm{d}t'. \tag{9.68}$$

By rewriting the displacements $\delta\mathbf{r}$ as linear relations of $\delta\mathbf{r}'$ and substituting the result into (9.68), he was then able to obtain the transformation formulas (9.50) by equating the coefficients of $\delta\mathbf{r}'$ in his expressions for δJ and $\delta J'$. From this he concluded that "[t]he principle of least action leads us to the same result as the analysis of §1."[29]

9.6 LORENTZ GROUP

Poincaré devoted the fourth paragraph to a brief investigation of the group-theoretical properties of the set of Lorentz transformations with scale factor l:

> It is important to note that the Lorentz transformations form a group.
>
> Indeed, if we set:
>
> $$x' = \gamma l(x + \epsilon t), \quad y' = ly, \quad z' = lz, \quad t' = \gamma l(t + \epsilon x),$$
>
> and in addition
>
> $$x'' = \gamma' l'(x' + \epsilon' t'), \quad y'' = l' y', \quad z'' = l' z', \quad t'' = \gamma' l'(t' + \epsilon' x'),$$
>
> with
>
> $$\gamma^{-2} = 1 - \epsilon^2, \quad \gamma'^{-2} = 1 - \epsilon'^2$$
>
> it follows:
>
> $$x'' = \gamma'' l''(x + \epsilon'' t), \quad y'' = l'' y, \quad z'' = l'' z, \quad t'' = \gamma'' l''(t + \epsilon'' x),$$
>
> with
>
> $$\epsilon'' = \frac{\epsilon + \epsilon'}{1 + \epsilon\epsilon'}, \quad l'' = ll', \quad \gamma'' = \gamma\gamma'(1 + \epsilon\epsilon') = \frac{1}{\sqrt{1 - \epsilon''^2}}.$$ [30]

[29] Poincaré 2021, p. 144.
[30] Ibid., pp. 144–145.

This proved that Lorentz transformations were closed under composition. From the composition formula it was clear that the Lorentz transformation with $\epsilon = 0$ and $l = 1$ was the identity transformation, while

$$x' = \frac{\gamma}{l}(x - \epsilon t), \quad y' = \frac{1}{l}y, \quad z' = \frac{1}{l}z, \quad t' = \frac{\gamma}{l}(t - \epsilon x) \tag{9.69}$$

corresponded to the inverse. Poincaré went on to obtain Lie generators for the Lorentz group, on the basis of which he concluded:

> Any transformation of this group can always be decomposed into a transformation of the form:
>
> $$x' = lx, \quad y' = ly, \quad z' = lz, \quad t' = lt$$
>
> and a linear transformation which does not change the quadratic form
>
> $$x^2 + y^2 + z^2 - t^2.$$[31]

The Lorentz transformations with arbitrary l could therefore be regarded as rotations around the origin of a four-dimensional space with an imaginary fourth coordinate. Poincaré remarked, however, that for the purposes of his article it was only necessary to consider that part of the transformations for which l could be regarded as a function of ϵ, making its elements a subgroup.

He then argued that this condition implied $l = 1$ for all ϵ and named the corresponding subgroup P. The argument was as follows. Assuming that the transformation

$$x' = \gamma l(x + \epsilon t), \quad y' = ly, \quad z' = lz, \quad t' = \gamma l(t + \epsilon x) \tag{9.70}$$

belonged to the subgroup, he rotated the two frames S and S' by 180° around the y-axis. This reversal flipped the signs of the x and z coordinates (and likewise of x' and z'), while leaving y and t unchanged. The rotated transformation therefore took the form

$$x' = \gamma l(x - \epsilon t), \quad y' = ly, \quad z' = lz, \quad t' = \gamma l(t - \epsilon x). \tag{9.71}$$

[31] Ibid., p. 146.

Poincaré observed that this new transformation must also belong to the same subgroup. The key point was that the only change produced by the spatial rotation was the reversal $\epsilon \mapsto -\epsilon$. Therefore the scale factor l could not distinguish between positive and negative values of ϵ; it could only depend on $|\epsilon|$. By assumption, the inverse of (9.70) was also an element of the subgroup. The latter therefore contained two transformations with the same ϵ. This meant that (9.71) and (9.69) had to be identical. The comparison of the two then resulted in

$$l = \frac{1}{l},$$

which proved the point.

Before proceeding, we should clarify the nature of Poincaré's group-theoretical assumptions for the elements of the Lorentz group applicable to electromagnetic phenomena. As we have seen in section 9.3, all Lorentz transformations with scale factor l preserved the form of the Lorentz-Maxwell equations. However, for each possible value of ϵ, Poincaré told us that only one could be considered for the application mentioned. It followed that l should be a function of ϵ. But what exactly determined the value of $l(\epsilon)$? And why did these transformations have to form a subgroup?

To answer these questions, let us assume that S and S' are identical reference frames adapted to the ether. Secondly, imagine that S' is set into a uniform rectilinear motion $-\boldsymbol{\epsilon} = (-\epsilon, 0, 0)$ with respect to S. I would like to argue that Poincaré intended the value of $l(\epsilon)$ to be uniquely determined by the requirement that

$$\begin{aligned} t' &= \gamma l(\epsilon)(t + \epsilon x), & x' &= \gamma l(\epsilon)(x + \epsilon t), \\ y' &= l(\epsilon)y, & z' &= l(\epsilon)z \end{aligned} \tag{9.72}$$

should provide us with the relationships between corresponding measurements in S and S'. Under this assumption it follows from the principle of relativity that (apparent) length measurements in S' of objects at rest with respect to S should only depend on the relative velocity between S and S'. For example, imagine a meter stick placed parallel to the x-direction in S. Measuring its length in S' should then give the same result, regardless of whether S' moves along the x-axis of S with speed ϵ or in the opposite direction (with exactly the same speed), because contrary results would contradict the isotropy of (apparent) space and thus the principle of relativity. Therefore, $l(\epsilon) = l(-\epsilon)$.

Let us consider the assumption that these transformations form a subgroup P. For this purpose, we introduce another frame S'' that moves with uniform velocity $-\boldsymbol{\epsilon}' = (-\epsilon', 0, 0)$ with respect to S':

$$t'' = \gamma' l(\epsilon')(t' + \epsilon' x'), \qquad x'' = \gamma' l(\epsilon')(x' + \epsilon' t'),$$
$$y'' = l(\epsilon')y', \qquad z'' = l(\epsilon')z'. \tag{9.73}$$

The principle of relativity then requires that S'' is related to S by the composition of (9.72) with (9.73):

$$t'' = \gamma'' l(\epsilon'')(t + \epsilon'' x), \qquad x'' = \gamma'' l(\epsilon'')(x + \epsilon'' t),$$
$$y'' = l(\epsilon'')y, \qquad z'' = l(\epsilon'')z,$$

where

$$\epsilon'' = \frac{\epsilon + \epsilon'}{1 + \epsilon\epsilon'}, \quad l(\epsilon'') = l(\epsilon)l(\epsilon'), \quad \gamma'' = \frac{1}{\sqrt{1 - \epsilon''^2}}.$$

For $\epsilon' = -\epsilon$ in particular, the composition should result in the identity transformation. It follows that

$$1 = l(\epsilon)l(-\epsilon) = l(\epsilon)^2,$$

so that the positive scale factor l must be equal to 1 for all values of ϵ. The fulfillment of the principle of relativity therefore implies that Lorentz transformations applicable to electromagnetic phenomena form a subgroup.

9.7 INVESTIGATING ELECTRON MODELS

In §6 Poincaré addressed the inconsistency in Lorentz's electron model that had been pointed out by Abraham. Instead of referring to Abraham explicitly, he reformulated the problem in terms of the equilibrium condition for the electron. For a charged body at rest, the Lorentz force was given by $\mathbf{f} = \rho\mathbf{d} + \rho\mathbf{v} \times \mathbf{h}$. For a stationary electron, $\mathbf{v} = 0$, so equilibrium required $\mathbf{f} = \rho\mathbf{d} = \mathbf{0}$. The only way this could hold, given Gauss's law $\nabla \cdot \mathbf{d} = \rho$, was if $\rho = 0$. This was the paradox: Lorentz's model described an electron endowed with charge, yet its own equilibrium condition appeared to imply the absence of charge. Abraham therefore concluded that the model was inconsistent. To avoid this difficulty, Abraham had proposed to treat the electron as absolutely rigid, thereby turning the condition into a purely kinematical constraint rather than a dynamical one. Langevin and Bucherer proposed another alternative: instead of treating the electron as absolutely rigid, they assumed it to be deformable but incompressible. In this way, the stresses required to balance the Coulomb forces were not introduced as new forces, but rather absorbed into the kinematical constraints of the model itself. Poincaré took a different path. Rather than discarding Lorentz's electron in favor of one of its alternatives, he made it his task to determine electron models that 1) were consistent and 2) transformed so as to preserve the (ether) rest shape in the comoving frame S'. The requirement for consistency entailed that he had to find the missing adhesive forces in Lorentz's model "so that the equilibrium of the electron is not disturbed by the tranformation,"[32] while the second requirement was necessary to fulfill the principle of relativity.

Poincaré began his investigation by applying the theory of corresponding states to the electron models of Abraham, Lorentz and Langevin-Bucherer, embedding them in his Lagrangian formulation of Lorentz's theory in order to facilitate the calculation of their fields. 1) In Abraham's model, "the electrons are spherical and not deformable," while the corresponding ideal "electron becomes a perfect ellipsoid"[33] under the Lorentz transformations with scale factor l. 2) In the models of Lorentz and Langevin-Bucherer "the moving electrons are deformed, so that the real electron would become an ellipsoid, while the ideal electron is still always a perfect sphere."[34] Lorentz argued that a moving electron was contracted in the direction of motion. However, Langevin and Bucherer assumed that the deformable electron was incompressible, i.e. its contraction in the direction of motion was compensated by corresponding expansions in the two directions transverse to the motion.

Poincaré explained how to calculate the corresponding state of the moving Abraham electron, but left the calculations for the other two models to the reader as an exercise. In the following, we will generalize his calculations so that they can easily be applied to all three cases.

Let us assume that an ellipsoidal electron with half-axes (a, b, c) moves along the x-axis in S at velocity $\mathbf{v} = (v, 0, 0)$:

$$\frac{(x - vt)^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{9.74}$$

[32] Poincaré 2021, p. 136.

[33] Ibid., p. 152.

[34] Ibid., p. 152.

To transform the electron's equation of motion from the ether frame S to its corresponding state in S', we express (t, x, y, z) as functions of (t', x', y', z') using the inverse (9.69) and insert the result into (9.74):

$$\frac{((1+v\epsilon)x' - (v+\epsilon)t')^2}{(\gamma^{-1}la)^2} + \frac{y'^2}{(lb)^2} + \frac{z'^2}{(lc)^2} = 1. \tag{9.75}$$

If we require that the electron in S' is at rest, we must set $\epsilon = -v$, which in turn gives

$$1 + v\epsilon = 1 - \epsilon^2 = \frac{1}{\gamma^2}. \tag{9.76}$$

The substitution of (9.76) in (9.75) then results in

$$\frac{x'^2}{(\gamma la)^2} + \frac{y'^2}{(lb)^2} + \frac{z'^2}{(lc)^2} = 1. \tag{9.77}$$

The moving ellipsoid with half-axes (a, b, c) thus corresponds in S' to a stationary ellipsoid with half-axes $(\gamma la, lb, lc)$. While the electron models of Abraham, Lorentz and Langevin-Bucherer agreed that an electron at rest with respect to the ether had the shape of a sphere of radius R, they disagreed on the shape of an electron in motion with respect to the ether (as already mentioned). According to Abraham, the moving electron was a sphere of radius R. By (9.77) its corresponding electron therefore had the shape of an ellipsoid with half-axes $(\gamma lR, lR, lR)$. According to Lorentz, the moving electron was an ellipsoid with half-axes $\left(\frac{R}{\gamma l}, \frac{R}{l}, \frac{R}{l}\right)$. In S' it therefore corresponded to a sphere of radius R. Langevin-Bucherer had argued that the moving electron was an ellipsoid with half-axes $\left(\frac{R}{\gamma^{2/3}}, \gamma^{1/3}R, \gamma^{1/3}R\right)$. The corresponding electron was therefore a sphere of radius $\gamma^{1/3}lR$.

Since Poincaré had already argued that l must be equal to 1 (see section 9.6), we should expect him to apply this condition to any Lorentz transformation between S and S'. However, this was not the case. While he left the value of l undetermined in the case of Abraham's electron model, he chose $l = 1$ in the case of Lorentz, but $l = \gamma^{-1/3}$ in the case of Langevin-Bucherer. So how did he determine the value of the scale factor? Although he did not explain his choice, I would argue that it was based on the requirement that the equation of motion of the electron in S should correspond to the equation of a sphere of radius R in S'. In the case of Lorentz's electron, this requirement was already fulfilled by design (so Poincaré could follow Lorentz and set $l = 1$), while the scale factor of the moving Langevin-Bucherer electron had to be equal to $\gamma^{-1/3}$. Since it was impossible to fulfill this condition for Abraham's electron model, Poincaré left its scale factor undetermined. The physical significance of this impossibility was that the moving Abraham electron could not have the shape of a sphere in a comoving frame S'. It would therefore be possible to detect its absolute motion with respect to the ether. Similarly to the group-theoretical requirement (which implied $l = 1$), the constraint that electrons had a spherical shape in their rest frame S' was thus a relativity condition. And since only Lorentz's electron model could fulfill both conditions, only this model could be considered for a viable explanation of the relativity of motion.

At this point, however, Poincaré emphasized that Lorentz's electron model had its own problems by generalizing Abraham's calculations. As we shall see, this analysis led to the conclusion that Langevin-Bucherer's model was the only deformable electron model compatible with a Lagrangian defined only by its self-field. This is because only for a model with $l = \gamma^{-1/3}$ did the energy E, the Lagrangian L and the momentum $\mathbf{G} = (G, 0, 0)$ of the moving electron agree with the standard relations between Hamiltonian, Lagrangian and generalized momentum:

$$E = L - G\epsilon, \tag{9.78}$$

$$G = \frac{\partial L}{\partial \epsilon}.\text{[35]} \tag{9.79}$$

To arrive at this conclusion, Poincaré calculated the above quantities by means of the electron's corresponding electrostatic state in S'. In this case, the equations (9.21) reduced to

$$\begin{aligned} \mathbf{h}' &= \mathbf{0}, \\ \mathbf{d}' &= -\nabla\varphi'. \end{aligned} \tag{9.80}$$

The substitution of (9.80) into (9.44) and (9.45) resulted in:

$$\begin{aligned} \mathbf{h} &= (0, \epsilon\gamma l^2 d_3', -\epsilon\gamma l^2 d_2'), \\ \mathbf{d} &= (l^2 d_1', \gamma l^2 d_2', \gamma l^2 d_3'). \end{aligned} \tag{9.81}$$

[35] Note the reversal of sign due to Poincaré's definition of the Lagrangian.

Poincaré then defined three new quantities to facilitate his calculations:

$$A = \frac{1}{2}\int d_1^2 \,\mathrm{d}\tau, \tag{9.82}$$

$$B = \frac{1}{2}\int d_2^2 + d_3^2 \,\mathrm{d}\tau, \tag{9.83}$$

$$C = \frac{1}{2}\int h_2^2 + h_3^2 \,\mathrm{d}\tau, \tag{9.84}$$

where A denoted the longitudinal electric energy, B the transverse electric energy and C the transverse magnetic energy. Since $h_1 = h_1' = 0$, it was not necessary to introduce a longitudinal magnetic energy. Letting A', B' and C' denote the corresponding quantities in S', we first note that $\mathbf{h}' = \mathbf{0}$ by assumption. Hence there was no magnetic field in the electron's rest frame and therefore

$$C' = \tfrac{1}{2}\int (h_2'^2 + h_3'^2)\,\mathrm{d}\tau' = 0.$$

In the moving frame S, Poincaré readily obtained

$$C = \frac{1}{2}\int \epsilon^2 d_2^2 + \epsilon^2 d_3^2 \,\mathrm{d}\tau = \epsilon^2 B \tag{9.85}$$

by means of (9.81).

He then noted that

$$\mathrm{d}\tau' = \gamma l^3 \mathrm{d}\tau, \tag{9.86}$$

which we can obtain by combining (9.66) and

$$\mathrm{d}t' = \frac{l}{\gamma}\mathrm{d}t.$$

To prove the latter, we note that the electron was comoving with S':

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\epsilon.$$

The equation now follows by differentiating t' with respect to t:

$$\frac{\mathrm{d}t'}{\mathrm{d}t} = \gamma l\left(1 + \epsilon\frac{\mathrm{d}x}{\mathrm{d}t}\right) = \gamma l(1-\epsilon^2) = \frac{l}{\gamma}.$$

Using (9.81) and (9.86) it was easy to obtain

$$A = \frac{l}{\gamma}A', \quad B = \gamma l B'. \tag{9.87}$$

Poincaré went on to limit his investigation to deformable electron models that satisfied what he called "Lorentz's hypothesis:"[36]

$$B' = 2A'. \tag{9.88}$$

It restricted the deformable electron into a spherical shape in S', implying that the half-axes (a, b, c) of the moving ellipsoidal electron satisfied $\gamma a = b = c$ (by means of (9.77)).

On the basis of the above results, the French mathematician was finally able to express E, L and G in terms of A':

$$E = A + B + C = (3 + \epsilon^2) l\gamma A', \tag{9.89}$$

$$L = A + B - C = \frac{3l}{\gamma} A', \tag{9.90}$$

$$G = -4l\gamma\epsilon A'. \tag{9.91}$$

The first relation (9.78) was satisfied for all values of l:

$$\begin{aligned}
(3 + \epsilon^2) l\gamma A' &= \frac{3l}{\gamma} A' + 4\epsilon^2 l\gamma A' \\
\Rightarrow \quad l\gamma(1 - \epsilon^2) &= \frac{l}{\gamma} \\
\Rightarrow \quad \gamma &= \frac{1}{\sqrt{1 - \epsilon^2}}.
\end{aligned}$$

Poincaré claimed that the second relation (9.79) was only satisfied for $l = \gamma^{-1/3}$. Let us see how to obtain this result. The substitution of (9.90) and (9.91) into (9.79) results in

$$\begin{aligned}
-4l\gamma A'\epsilon &= \frac{\partial}{\partial \epsilon}\left(\frac{3A'l}{\gamma}\right) \\
\Rightarrow \quad -\frac{4}{3} l\gamma\epsilon &= \frac{1}{\gamma}\frac{\partial l}{\partial \epsilon} + l\frac{\partial}{\partial \epsilon}\left(\frac{1}{\gamma}\right) \\
\Rightarrow \quad -\frac{4}{3} l\gamma\epsilon &= \frac{1}{\gamma}\frac{\partial l}{\partial \epsilon} - l\gamma\epsilon \\
\Rightarrow \quad -\frac{1}{3} l\gamma^2\epsilon &= \frac{\partial l}{\partial \epsilon}.
\end{aligned}$$

[36] Poincaré 2021, p. 153.

Since $\gamma^2 = (1-\epsilon^2)^{-1}$ and $\frac{\partial}{\partial\epsilon}(1-\epsilon)^2 = -2\epsilon$, it is quite obvious to hint at a solution of the form

$$l(\epsilon) = \left(1-\epsilon^2\right)^n.$$

If we insert the latter into the differential equation, we get

$$\begin{aligned} -\frac{1}{3}\left(1-\epsilon^2\right)^n\left(1-\epsilon^2\right)^{-1}\epsilon &= \frac{\partial}{\partial\epsilon}\left(1-\epsilon^2\right)^n \\ \Rightarrow \quad -\frac{1}{3}\left(1-\epsilon^2\right)^{n-1}\epsilon &= -2n\left(1-\epsilon^2\right)^{n-1}\epsilon \\ \Rightarrow \quad \frac{1}{3} &= 2n \\ \Rightarrow \quad n &= \frac{1}{6}. \end{aligned}$$

Therefore, $l = (1-\epsilon^2)^{1/6} = \gamma^{-1/3}$. It followed that Langevin-Bucherer's model was the only deformable electron model that fulfilled the second relationship. Poincaré was thus in a dilemma between the contradictory results of §4 and §6.

9.8 RECTIFYING LORENTZ'S ELECTRON MODEL

To find a solution, he continued to search for a general account of electron models using the theory of corresponding states:

> Let us consider any hypothesis, which may be either that of Lorentz, or that of Abraham, or that of Langevin, or an intermediate hypothesis.
>
> Let
>
> $$r, \theta r, \theta r$$
>
> the three axes of the real electron; that of the ideal electron will be:
>
> $$\gamma l r, \theta l r, \theta l r$$
>
> Then $A' + B'$ is the electrostatic energy of an ellipsoid with axes $\gamma l r$, $\theta l r$, $\theta l r$.
>
> Let us suppose that the electricity is spread on the surface of the electron as it is known from an inductor, or uniformly distributed within the electron; than this energy will be of the form:
>
> $$A' + B' = \frac{\phi\left(\frac{\theta}{\gamma}\right)}{\gamma l r},$$

where ϕ is a known function.[37]

Poincaré himself provided no argument for this decisive formula. So let us see if we can prove it by applying a result from Abraham.

The German had obtained a formula for the value of the potential on the surface of a stationary ellipsoidal electron with half-axes (a_0, b_0, c_0) and uniformly distributed surface charge e:

$$\varphi_0 = \frac{e}{8\pi}\int_0^\infty \left((a_0^2+s)(b_0^2+s)(c_0^2+s)\right)^{-1/2} \mathrm{d}s.^{38} \tag{9.92}$$

Since the potential had a constant value on the electron’s surface, the electrostatic energy of the electron was given by the following surface integral:

$$E_0 = \frac{1}{2}\int \varsigma_0 \varphi_0 \,\mathrm{d}\sigma = \frac{1}{2} e\varphi_0, \tag{9.93}$$

where ς_0 was the constant surface charge density.

To account for Poincaré’s assumption, we have to distinguish between two cases: 1) $a_0 = b_0 = c_0$, and 2) $a_0 \neq b_0 = c_0$. For both cases, we have to show that $E_0 = E'$ can be written in the desired form for $a_0 = \gamma lr$, $b_0 = \theta lr$ and $c_0 = \theta lr$.

The first case $a_0 = b_0 = c_0$ corresponds to $\theta = \gamma$. If we calculate (9.92), we get:

$$\varphi_0 = \frac{e}{4\pi\gamma lr}. \tag{9.94}$$

If (9.94) is substituted into (9.93), the result is

$$E' = \frac{e^2}{8\pi\gamma lr}. \tag{9.95}$$

Comparing this result with Poincaré’s general formula

$$E' = \frac{\phi\left(\frac{\theta}{\gamma}\right)}{\gamma lr}, \tag{9.96}$$

we see that

$$\phi(1) = \frac{e^2}{8\pi}. \tag{9.97}$$

[37] Poincaré 2021, p. 154.
[38] Abraham 1905, p. 166.

The second case $a_0 \neq b_0 = c_0$ corresponds to $\theta \neq \gamma$. By applying (9.92) to the prolate ellipsoid ($\frac{b_0}{a_0} = \frac{\theta}{\gamma} < 1$), Abraham obtained

$$\varphi_0 = \frac{e}{4\pi} \frac{\ln\left(\frac{a_0+\sqrt{a_0^2-b_0^2}}{b_0}\right)}{\sqrt{a_0^2 - b_0^2}}.^{39} \tag{9.98}$$

By substituting (9.98) into (9.93) and using $\frac{b_0}{a_0} = \frac{\theta}{\gamma}$, we then get

$$E' = \frac{e^2}{8\pi\gamma l r} \frac{\ln\left(\left(\frac{\theta}{\gamma}\right)^{-1} + \sqrt{\left(\frac{\theta}{\gamma}\right)^{-2} - 1}\right)}{\sqrt{1 - \left(\frac{\theta}{\gamma}\right)^2}}.$$

Comparing our result with Poincaré's general formula (9.96), it follows that

$$\phi\left(\frac{\theta}{\gamma}\right) = \frac{e^2}{8\pi} \frac{\ln\left(\left(\frac{\theta}{\gamma}\right)^{-1} + \sqrt{\left(\frac{\theta}{\gamma}\right)^{-2} - 1}\right)}{\sqrt{1 - \left(\frac{\theta}{\gamma}\right)^2}}, \qquad \gamma > \theta \tag{9.99}$$

A similar result can be obtained for the oblate ellipsoid ($\frac{b_0}{a_0} = \frac{\theta}{\gamma} > 1$), proving Poincaré's general formula.

Let us return to Poincaré's application of his general formula to the three electron models. Based on the relationships between the spatial dimensions of the moving electron and its corresponding stationary electron, Poincaré noted that all three models adhered to a constraint of the form

$$r = b\theta^m. \tag{9.100}$$

Let us fill in some details. The rigidity of Abraham's electron implied $\theta = 1$ and

$$r = R,$$

while the spherical shape of the electron in S' results in $\theta = \gamma$ for the other two models. It follows that Lorentz's and Langevin-Bucherer's electron

[39] Ibid., p. 179.

models both corresponded to the first case ($\theta = \gamma$), whereas Abraham's model belonged to the second case ($\frac{\theta}{\gamma} = \frac{1}{\gamma} < 1$). Lorentz assumed $l = 1$ and

$$\theta r = R,$$

Langevin and Bucherer chose $l = \gamma^{-1/3}$ and thus

$$\theta^{2/3} r = R.$$

Isolating r, we obtain that all three models fulfill a condition of the form

$$r = R\theta^m. \tag{9.101}$$

In particular, $m = 0$ for Abraham's model, $m = -1$ for Lorentz's model and $m = -\frac{2}{3}$ for Langevin-Bucherer's model. Comparing our result with Poincaré's, we further see that his constant b must be equal to the radius R of an electron at rest.

Next, Poincaré rewrote the formula for the Lagrangian of the moving electron. From (9.85) and (9.87) he obtained

$$L = A + B - C = \frac{l}{\gamma}(A' + B') = \frac{l}{\gamma}E'. \tag{9.102}$$

The substitution of (9.96) and (9.100) into (9.102) then resulted in

$$L = \frac{\phi\left(\frac{\theta}{\gamma}\right)}{R\theta^m \gamma^2}. \tag{9.103}$$

He went on to note that

$$\frac{\partial L}{\partial \theta} = 0 \tag{9.104}$$

determined the shape of the moving electron under the condition that "we do not suppose the involvement of forces other than the binding forces."[40] To recognize this, we were told to remember the following general equations:

$$\frac{\mathrm{d}\mathrm{G}}{\mathrm{d}t} = -\int \mathbf{f}\,\mathrm{d}\tau, \tag{9.105}$$

$$J = \int L\,\mathrm{d}t, \tag{9.106}$$

$$\delta J = -\int \mathbf{f}\cdot\delta\mathbf{r}\,\mathrm{d}\tau\mathrm{d}t, \tag{9.107}$$

[40] Poincaré 2021, p. 155.

where **f** denoted the force produced by the electron's self-field. Since L was a function of θ and ϵ, it followed from (9.106) that the variation of J could be written as

$$\delta J = \int \frac{\partial L}{\partial \theta}\delta\theta + \frac{\partial L}{\partial \epsilon}\delta\epsilon \, \mathrm{d}t. \tag{9.108}$$

By assumption, the electron moved with velocity $\mathbf{v} = (-\epsilon, 0, 0)$, so that

$$\epsilon = -\frac{\mathrm{d}x}{\mathrm{d}t}$$

and

$$\delta\epsilon = -\frac{\mathrm{d}\delta x}{\mathrm{d}t}.$$

The substitution of (9.105) into (9.107) and integration by parts then resulted in

$$\delta J = \int \frac{\mathrm{d}\mathbf{G}}{\mathrm{d}t}\cdot\delta\mathbf{r}\, \mathrm{d}t = -\int G\frac{\mathrm{d}\delta x}{\mathrm{d}t}\,\mathrm{d}t = \int G\delta\epsilon\,\mathrm{d}t. \tag{9.109}$$

From the comparison of (9.108) and (9.109) Poincaré then concluded

$$\frac{\partial L}{\partial \theta} = 0, \qquad \frac{\partial L}{\partial \epsilon} = G. \tag{9.110}$$

Assuming that the ideal electron had the shape of a sphere corresponding to $\frac{\theta}{\gamma} = 1$, Poincaré determined the shape of the electron by differentiating (9.103) with respect to θ and solving the resulting differential equation. By differentiating (9.103) he obtained

$$-m\theta^{-m-1}\phi + \theta^{-m}\gamma^{-1}\phi' = 0,$$

which could also be written in the form

$$\frac{\phi'}{\phi} = \frac{m\gamma}{\theta}.$$

Applying the requirement $\frac{\theta}{\gamma} = 1$ resulted in

$$\frac{\phi'}{\phi} = m. \tag{9.111}$$

He then noticed that Abraham had obtained the following expression for ϕ:

$$\phi\left(\tfrac{1}{\gamma}\right) = \frac{a}{\epsilon}\ln\left(\frac{1+\epsilon}{1-\epsilon}\right), \tag{9.112}$$

where a was a constant. We can easily prove this by means of (9.99):

$$\phi\left(\frac{1}{\gamma}\right) = \frac{e^2}{8\pi}\frac{\ln\left(\left(\frac{1}{\gamma}\right)^{-1} + \sqrt{\left(\frac{1}{\gamma}\right)^{-2} - 1}\right)}{\sqrt{1 - \left(\frac{1}{\gamma}\right)^2}} = \frac{e^2}{16\pi\epsilon}\ln\left(\frac{1+\epsilon}{1-\epsilon}\right).$$

If we compare the two expressions for $\phi\left(\frac{1}{\gamma}\right)$, we get

$$a = \frac{e^2}{16\pi} \tag{9.113}$$

for the constant. By expanding $\frac{1}{\gamma}$ in powers of ϵ up to the second order, Poincaré noted that equation (9.112) reduced to

$$\phi\left(1 - \frac{\epsilon^2}{2}\right) = a\left(1 + \frac{\epsilon^2}{3}\right).$$

However, the correct result would have been

$$\phi\left(1 - \frac{\epsilon^2}{2}\right) = 2a\left(1 + \frac{\epsilon^2}{3}\right). \tag{9.114}$$

Poincaré did rarely keep track of constants. The reason for the discrepancy is probably that the value of a did not matter to him. Since I would like to compare Poincaré's result with Abraham's, I will continue as if Poincaré had used the correct result.

By differentiating (9.114) with respect to ϵ he obtained

$$-\epsilon\phi'\left(1 - \frac{\epsilon^2}{2}\right) = \frac{4}{3}\epsilon a. \tag{9.115}$$

Letting ϵ approach zero resulted in the equations

$$\phi(1) = 2a, \quad \phi'(1) = -\frac{4}{3}a, \quad \frac{\phi'(1)}{\phi(1)} = -\frac{2}{3}. \tag{9.116}$$

By comparing this result with (9.111), it finally followed that $m = -\frac{2}{3}$ in agreement with the Langevin-Bucherer electron model. The latter was thus the only deformable electron model with $\frac{\theta}{\gamma} = 1$ in dynamical equilibrium. For all other electron models with $\frac{\theta}{\gamma} = 1$ such as Lorentz's, it was necessary to add non-electromagnetic adhesive forces to ensure their stability.

Since Lorentz's electron model was the only possible candidate for a viable explanation of the relativity of motion, Poincaré's next step was to determine "what additional forces, other than the binding forces, are necessary to intervene to account for the law of Lorentz."[41] The simplest way to derive such forces was to add a complementary potential L_c depending on the axes $(r, \theta r, \theta r)$ of the electron to the Lagrangian of its self-field:

$$J = \int L_e + L_c \, \mathrm{d}t, \tag{9.117}$$

where L_e denoted the purely electromagnetic part of the new total Lagrangian $L = L_e + L_c$. Due to this addition, the equilibrium condition (9.104) had to be replaced by

$$\frac{\partial L_e}{\partial \theta} + \frac{\partial L_c}{\partial \theta} = 0, \tag{9.118}$$

$$\frac{\partial L_e}{\partial r} + \frac{\partial L_c}{\partial r} = 0. \tag{9.119}$$

From (9.103) it followed that

$$\frac{\partial L_e}{\partial \theta} = \frac{\phi'}{\gamma^3 r}, \qquad \frac{\partial L_e}{\partial r} = -\frac{\phi}{\gamma^2 r^2}.$$

Combining these results with the two conditions and taking into account $\frac{\theta}{\gamma} = 1$, $r = R\theta^m$ and (9.116), Poincaré obtained

$$\frac{\partial L_c}{\partial \theta} = \frac{4}{3}\frac{a}{R\theta^{m+3}}, \qquad \frac{\partial L_c}{\partial r} = \frac{2a}{R^2\theta^{2m+2}}. \tag{9.120}$$

He noted that these equations were satisfied by a term of the form

$$L_c = Ar^\alpha \theta^\beta, \tag{9.121}$$

where A, α and β were constants. Substituting this equation into (9.120) and performing the respective differentations then resulted in

$$A\beta R^\alpha \theta^{m\alpha+\beta-1} = \frac{4}{3}\frac{a}{R\theta^{m+3}}, \quad A\alpha R^{\alpha-1}\theta^{m\alpha-m+\beta} = \frac{2a}{R^2\theta^{2m+2}}.$$

By introducing

$$\kappa = -\frac{m+2}{3m+2},$$

[41] Poincaré 2021, p. 157.

a comparison with (9.120) then showed that

$$\alpha = 3\kappa, \quad \beta = 2\kappa, \quad A = \frac{2a}{\alpha R^{\alpha+1}}. \tag{9.122}$$

Since the volume of the ellipsoidal electron was proportional to $r^3\theta^2$, Poincaré was able to conclude that the additional potential L_c was proportional to the power κ of the electron volume. To recognize this, we simply have to insert $\alpha = 3\kappa$ and $\beta = 2\kappa$ in (9.121):

$$L_c = A\left(r^3\theta^2\right)^\kappa. \tag{9.123}$$

In the case of Lorentz's electron model $m = -1$ and therefore $\kappa = 1$. It followed that the additional potential could be written as

$$L_c = P_c V_e, \tag{9.124}$$

where V_e was the volume of the moving electron and P_c was a constant. As we will see later, the addition of L_c increased the electron's energy in such a way that the standard relations (9.78) and (9.79) between Hamiltonian, Lagrangian and generalized momentum were satisfied. Adding the term L_c to the Lagrangian therefore solved the discrepancy in Lorentz's electron model.

9.9 RELATIVITY OF MOTION

Before Poincaré addressed this discrepancy in more detail, he demonstrated the inability to detect absolute motion using the Lorentz invariance of the complementary principle of least action J (9.117). To this end, he assumed that the electromagnetic field in S was generated by a finite number of electrons. He had already obtained the Lorentz invariance of the purely electromagnetic principle of least action J_e in §3, so that he only needed to consider the complementary term J_c generated by the Lorentz electrons. Since the latter fulfilled the condition (9.124), he could write the complementary Lagrangian L_c as

$$L_c = \int P_c \, \mathrm{d}\tau,$$

where the integral was extended to the entire space by setting the pressure P_c outside the electrons equal to zero. It followed that the complementary term J_c was equal to

$$J_c = \int P_c \, \mathrm{d}\tau \mathrm{d}t.$$

Introducing its corresponding term in S'

$$J'_c = \int P'_c \, \mathrm{d}\tau' \mathrm{d}t',$$

he obtained $P'_c = P_c$ by arguing that "if a point belong[s] to an electron, the corresponding point after the Lorentz transformation still belongs to the same electron."[42] The substitution of $l = 1$ into (9.66) resulted in

$$\mathrm{d}\tau' \mathrm{d}t' = \mathrm{d}\tau \mathrm{d}t.$$

Consequently,

$$J_c = \int P_c \, \mathrm{d}\tau \mathrm{d}t = \int P'_c \, \mathrm{d}\tau' \mathrm{d}t' = J'_c$$

and

$$J = J'.$$

Poincaré therefore felt justified in concluding:

> The theorem is thus general, it gives us at the same time a solution of the question we posed at the end of §1: finding the complementary forces which are unaltered by the Lorentz transformation. The additional potential L_c satisfies this condition.
>
> So we can generalize the result announced at the end of §1 and write:
>
> *If the inertia of electrons is exclusively of electromagnetic origin, if they are only subject to forces of electromagnetic origin, or to forces generated by the additional potential (L_c), no experiment can demonstrate absolute motion.*[43]

As explained above, Poincaré began his article by showing that a deformable electron could not be in equilibrium under the action of a purely electromagnetic Lagrangian. He therefore set himself the task of determining the additional potential necessary to 1) maintain equilibrium and 2) have the desired transformation properties. While he obtained the potential using the equilibrium condition in §6, he showed in §8 that it had the desired properties under Lorentz transformations. Poincaré had thus achieved his goal.

[42] Poincaré 2021, p. 164.
[43] Ibid., p. 165. I have changed Poincaré's notation for the additonal potential to L_c.

Before we address the Sorbonne lectures, we still need to see that the addition of the complementary potential to the Lagrangian of the electron solved the discrepancy in Lorentz's electron model (where $m = -1$ and $\frac{\theta}{\gamma} = 1$). Furthermore, we need an explanation for Poincaré's claim that "the inertia of electrons is exclusively of electromagnetic origin." From (9.100), (9.113) and (9.122) it follows that the constant A is equal to

$$A = \frac{e^2}{24\pi R^4}. \tag{9.125}$$

Using this result and (9.123) we obtain

$$L_c = Ar^3\theta^2 = \frac{e^2}{24\pi\gamma R}, \tag{9.126}$$

while (9.103) and (9.116) give us

$$L_e = \frac{\phi(1)}{\gamma R} = \frac{e^2}{8\pi\gamma R}. \tag{9.127}$$

We can also express L_e by the electromagnetic energy of the electron at rest E'_e. For $\epsilon = 0$ equation (9.89) yields

$$A' = \frac{1}{3}E'_e, \tag{9.128}$$

which in combination with (9.90) results in

$$L_e = \frac{E'_e}{\gamma}. \tag{9.129}$$

If we combine (9.126) and (9.127), we find that the total Lagrangian L can be written as

$$L = L_e + L_c = \frac{4}{3}L_e. \tag{9.130}$$

Calculating the momentum using L therefore results in

$$G = \frac{\partial L}{\partial \epsilon} = \frac{4}{3}\frac{\partial L_e}{\partial \epsilon} = -\frac{4}{3}\gamma\epsilon E'_e. \tag{9.131}$$

On the other hand, we can also obtain the momentum using its electromagnetic definition via (9.91):

$$G = -4\gamma\epsilon A' = -\frac{4}{3}\gamma\epsilon E'_e. \tag{9.132}$$

Since the two calculations lead to the same result, we come to the conclusion that the discrepancy first observed by Abraham has been resolved.

Let us also compare Poincaré's and Abraham's results for the complementary energy E_c. The latter corresponds to L_c in Poincaré's model:

$$E_c = L_c = \frac{1}{3}L_e = \frac{1}{3}\frac{E_e'}{\gamma}. \tag{9.133}$$

For the electron at rest we get

$$E_c' = \frac{1}{3}E_e'. \tag{9.134}$$

If we insert this result into Abraham's expression (9.18) for E_c, we obtain

$$E_c = {E_c}' + \frac{1}{3}\left(\frac{E_e'}{\gamma} - E_e'\right) = \frac{1}{3}\frac{E_e'}{\gamma}.$$

This means that Poincaré's corrections to Lorentz's electron model met Abraham's requirements and determined the value of E_c'. The Frenchman also had an answer to Abraham's question about the nature of the forces responsible for the additional potential:

> So what are these forces that create the potential L_c? They can obviously be compared to a pressure which would reign inside the electron; all occurs as if each electron were a hollow capacity subjected to a constant internal pressure (volume independent); the work of this pressure would be obviously proportional to the volume changes. In any case, I must observe that this pressure is negative.[44]

To see this, we recall that L_c could be written as

$$L_c = P_c V_e, \tag{9.135}$$

where V_e was the volume of the compressible electron and P_c a constant referred to as Poincaré stress after its originator. Its task was to prevent the surface charge of the deformable electron from flying apart by canceling the stress caused by the electron's self-field and increasing the electron's energy so that the standard relations (9.78) and (9.79) between Hamiltonian, Lagrangian and generalized momentum were satisfied. To calculate

[44] Poincaré 2021, p. 165.

the value of this constant stress, we equate (9.135) and (9.123) for $\kappa = 1$ and isolate P_c:

$$P_c = \frac{Ar^3\theta^2}{V} = \frac{3A}{4\pi} = \frac{e^2}{32\pi^2 R^4},$$

where we have used (9.125) and

$$V_e = \frac{4}{3}\pi r^3 \theta^2.$$

In contrast to Poincaré's own claim, our calculation shows that on his formal definitions the constant P_c is positive.[45]

To continue, let us explain how Poincaré substantiated his claim that the mass of the electron was purely electromagnetic in origin:

> Now assessing the mass of the electron – I mean the "experimental mass", that is to say the mass for low velocities – we have (cf. §6):
>
> $$L_e = \frac{\phi\left(\frac{\theta}{\gamma}\right)}{\gamma^2 r}, \quad \theta = \gamma, \quad \phi = [2]a, \quad \theta r = b;$$
>
> hence
>
> $$L_e = \frac{[2]a}{b\gamma} = \frac{[2]a}{b}\sqrt{1-\epsilon^2},$$
>
> I can write for very small ϵ
>
> $$L_e = \frac{[2]a}{b}\left(1 - \frac{\epsilon^2}{2}\right),$$
>
> so that the mass, both longitudinal and transverse, will be $\frac{[2]a}{b}$.[46]

Let us express Poincaré's value for the longitudinal and transverse mass of the electron at low speeds

$$m_0 = \frac{2a}{b}$$

[45] On Poincaré's own definitions, the constant P_c comes out positive, since $L_c = P_c V_e$ was added to the Lagrangian as a stabilizing term. Yet in his prose he described this as a "negative pressure." The mismatch reflects a difference of conventions: in continuum mechanics, a cohesive stress resisting expansion is often called a negative pressure (tension), whereas in Poincaré's formalism the same stabilizing contribution appears with a positive sign in the Lagrangian.

[46] Poincaré 2021, p. 165.

using $b = R$ and (9.113):

$$m_0 = \frac{2a}{R} = \frac{e^2}{8\pi R}. \tag{9.136}$$

In contrast, substituting $c = 1$ into equation (8.47), Lorentz obtained

$$m_0 = \frac{e^2}{6\pi R}. \tag{9.137}$$

This leads to the question: How did Poincaré derive his incorrect result? To suggest a plausible answer, rewrite the formula for the momentum (9.131) as the sum of an electromagnetic and a complementary part:

$$G = \frac{\partial L}{\partial \epsilon} = \frac{\partial L_e}{\partial \epsilon} + \frac{\partial L_c}{\partial \epsilon} = G_e + G_c.$$

The calculation of the electromagnetic component at low speeds via (9.127) and a Taylor expansion of the second order then gives

$$G_e = \frac{\partial L_e}{\partial \epsilon} = \frac{e^2}{8\pi R}\gamma\epsilon \doteq \frac{e^2}{8\pi R}\epsilon.$$

If we insert this result into the definition for the longitudinal and transverse mass (8.46), we get Poincaré's value:

$$m_1 = \frac{\mathrm{d}G_e}{\mathrm{d}\epsilon} \doteq \frac{e^2}{8\pi R}, \qquad m_2 = \frac{G_e}{\epsilon} \doteq \frac{e^2}{8\pi R}.$$

However, we should have used the total Lagrangian $L = \frac{4}{3}L_e$ (9.130) to obtain the correct result

$$m_0 = \frac{4}{3}\frac{e^2}{8\pi R} = \frac{e^2}{6\pi R}.$$

Poincaré's conclusion that the mass of the electron was of purely electromagnetic origin was therefore wrong. Because of his error, he was also mistaken to think that his theory was compatible with both the electromagnetic worldview and the principle of relativity!

9.10 SORBONNE LECTURES

In his Sorbonne lectures of 1906 and in later seminars and papers, Poincaré returned to his operational understanding of local time by means of optical signaling and proved that it remained valid for all orders of $\frac{u}{c}$ by showing that simultaneity was transitive. A contrary result would have violated the principle of relativity because it would have allowed the detection of motion through the ether. Let us examine his argument in its first published version as part of the article "La Dynamique de l'Électron"(1908).[47]

Assuming Lorentz contraction, Poincaré considered two observers in common translational motion with respect to the ether. However, the two observers themselves should believe that they are at rest. If we denote the speed with respect to the ether u, we can imagine that the two observers belong to a frame S' in standard configuration with respect to the ether-based frame S. One could then ignore the z-axis by rotating the axes of the two frames in such a way that both observers were situated in the (x, y)-plane.

Poincaré further assumed that the first observer emitted a flash of light at the origo of S and S' at time 0. From the perspective of a third observer, who was at rest relative to the ether, at true time t the first and second observers had moved ut relative to the ether, while the flash of light was located on a sphere of radius ct centered at the origo of S. However, when this sphere was viewed from the perspective of the moving observers, it appeared as a light ellipsoid with half-axes

$$(a, b, c) = (\gamma ct, ct, ct) \tag{9.138}$$

due to the contraction of the moving rulers in the direction of motion (see figure 9.1).

Since the eccentricity of the ellipse in the (x, y)-plane was equal to

$$e = \sqrt{1 - \frac{b^2}{a^2}} = \frac{u}{c}, \tag{9.139}$$

the focal distance

$$f = |OF| = ea = \gamma ut \tag{9.140}$$

corresponded to the apparent distance traveled by the observers during true time t. It followed that the first observer was at the focus F of the ellipse at time t. Assuming that the second observer received the flash of light at true time t, his position M would be on the ellipsoid. Since the first and second observer were comoving with S', the apparent distance $|FM|$ between them did not depend on time.

[47] To trace later developments, consult the works of Scott Walter (2009; 2014), the leading expert on this matter.

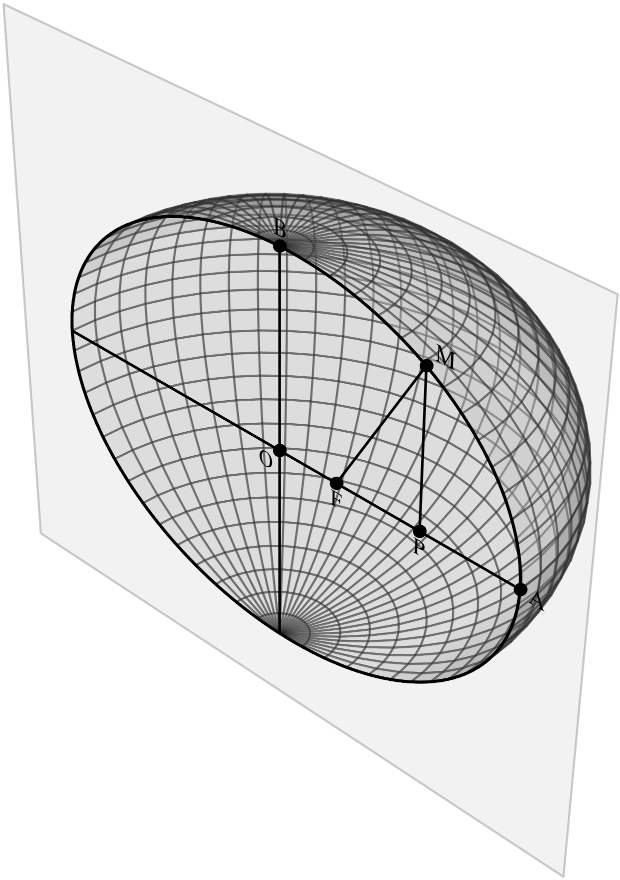


Fig. 9.1: Poincaré's light ellipsoid ($a = |OA|$, $b = |OB|$, $f = |OF|$).

According to a property of ellipses:

$$|FM| + e|FP| = \frac{b^2}{a}, \tag{9.141}$$

where P denoted the projection of M onto the x'-axis. Since the first observer was at the origin of S', the distance $|FP|$ corresponded to the x'-coordinate of the second observer. The combination of (9.138), (9.139), (9.140) and (9.141) then resulted in

$$t = \gamma \frac{|FM|}{c} + \gamma \frac{ux'}{c^2}. \tag{9.142}$$

Since the two observers were unaware of their common motion, on arrival of the light signal the second observer would set his clock so that the apparent transit time t' of the signal coincided with an apparent speed of c:

$$|FM| = ct'. \tag{9.143}$$

Substituting this into (9.142) and isolating for t', Poincaré obtained

$$t' = \gamma^{-1} t - \frac{ux'}{c^2} \tag{9.144}$$

in accordance with the transitivity condition. To see this more clearly, let clocks A and B be synchronized as described. The latter is the case if and only if

$$t'_B = t'_A.$$

If we rewrite this using (9.144), we get the relationship

$$t_A - t_B = \gamma u \frac{x'_A - x'_B}{c^2}. \tag{9.145}$$

between the corresponding true times t_A and t_B. If we assume that clock B is synchronized with clock C, we also get

$$t_B - t_C = \gamma u \frac{x'_B - x'_C}{c^2}.$$

It follows that clocks A and C are also synchronized:

$$t_A - t_C = t_A - t_B + t_B - t_C = \gamma u \frac{x'_A - x'_C}{c^2},$$

which had to be proven.

Poincaré went on to note that it was possible to derive the following equations from the above:

$$x' = \gamma\,(x - ut), \tag{9.146}$$

$$t' = \gamma\left(t - \frac{ux}{c^2}\right), \tag{9.147}$$

where x denoted the true x-coordinate of the second observer at true time t. If we remember that the light signal was emitted at the origo of S at $t = 0$, it follows that the true x-coordinate of the first observer is equal to ut. The apparent horizontal distance between the two observers, which corresponds to the x'-coordinate of the second observer, is therefore equal to (9.146). If we substitute (9.146) into (9.144), we get (9.147). The light ellipsoid argument thus provided the physical interpretation of the Lorentz-transformed coordinates as those measured by moving observers. In conjunction with the proof of the Lorentz invariance of the Lorentz-Maxwell equations in the 1906 memorandum, this meant that electrodynamical phenomena were consistent with the principle of relativity.

Poincaré also provided an argument for the transitivity of clock synchronization by means of optical cross-signaling. The operational advantage of this method over the first was that it did not require the determination of the apparent distance between two observers. The French mathematician had already calculated the time t_+ that a light signal would need to travel from the first observer to the second:

$$t_+ = \gamma \frac{|FM|}{c} + \gamma \frac{ux'}{c^2}.$$

To determine the travel time t_- of the reversed signal, he could have repeated the argument by reversing the roles of the two comoving observers. However, he preferred to observe that the x'-coordinate had a sign, so that the projection term in (9.141) changed sign. From this he concluded

$$t_- = \gamma \frac{|FM|}{c} - \gamma \frac{ux'}{c^2}.$$

While the second observer set his clock to 0 when the first observer's signal arrived at true time t_+, the first observer, who assumed that he was at rest relative to the ether and therefore believed that $t_+ = t_-$, set his clock to

$$\frac{\tau}{2} = \frac{t_+ + t_-}{2} = \gamma \frac{|FM|}{c} \tag{9.148}$$

at the arrival of the return signal. So from the stationary observer's point of view, the first observer made a mistake equal to

$$\frac{\tau}{2} - t_- = \frac{t_+ - t_-}{2} = \gamma \frac{ux'}{c^2}. \tag{9.149}$$

However, as Poincaré noted, it was not possible to detect this mistake because it led to a transitive condition for simultaneity. In particular, in agreement with the transitivity condition (9.148) for the first method, two comoving observers A and B would judge events to be simultaneous if and only if their true times t_A and t_B differed by

$$t_A - t_B = \gamma \frac{u(x'_A - x'_B)}{c^2}. \tag{9.150}$$

In his commentary on this argument, Scott Walter has identified a critical problem:

> Notice that Poincaré's thought experiment involves observers in motion exchanging light signals, but does not involve clocks in motion, since light time-of-flight is measured in true time, t, not apparent time, t'. Comoving observers have no access to true time, of course, which may be one reason why Poincaré never used this particular thought experiment again.[48]

I agree with Walter that Poincaré measured travel times in true times and did not aknowledge time dilation. But as far as I can see, he thereby did not intend that all clocks were at rest with respect to the ether. If we substitute $|FM| = ct'$ into (9.148) and place the origo of S' in M instead of F, we can rewrite (9.149) as

$$t' = \gamma^{-1} t - \frac{ux'}{c^2}.$$

That is to say, we have obtained equation (9.144) from the first method. It follows that the simultaneity condition (9.150) in terms of true time corresponds to the correct condition $t'_A = t'_B$ in terms of apparent time. Poincaré should have made this connection clear. Although it remains unsatisfactory that he did not do so, equation (9.149) should be seen as Poincaré's attempt to explain the transitivity of apparent simultaneity.

9.11 CONCLUSION

In 1905 Abraham demonstrated the inconsistency of Lorentz's electron model within the framework of Lagrangian mechanics. Despite this dilemma, Poincaré did not reject Lorentz's 1904 theory because it allowed him to rigorously fulfill the principle of relativity by proving the exact invariance of electrodynamic phenomena with respect to uniform translational motion. He achieved this result in a series of successive steps. Assuming the principle of relativity between a frame at rest and a frame in

[48] Walter 2014, p. 135.

uniform rectilinear motion relative to the ether, he proved the exact invariance of the Lorentz-Maxwell equations under Lorentz transformations with an undetermined scale factor l. In §2 of his paper, he then formulated Lorentz's theory within the framework of Lagrangian mechanics in accordance with his conception of mechanical explanation. This formulation enabled him to demonstrate the invariance of the entire theory under Lorentz transformations with undetermined scale factor solely by the invariance of the principle of least action (§3). On the basis of a group-theoretical proof, Poincaré argued that the scale factor must be equal to 1 in order to satisfy the principle of relativity. Nevertheless, in §6 he analyzed different electron models that were compatible with Lorentz transformations with an undetermined scale factor. Generalizing the calculations of Abraham, Poincaré proved that only Langevin-Bucherer's model agrees with a Lagrangian defined only by its self-field, while Lorentz's model alone fulfilled the principle of relativity. Poincaré thus found himself in a dilemma in the sense that the result of §4 was incompatible with §6. In order to find a solution, he supplemented the Lagrangian of Lorentz's model with an additional potential that made the electron dynamically stable and still complied with the principle of relativity. Whilst Poincaré left the physical interpretation of Lorentz's theory unclear in his 1906 paper, the light ellipsoid argument shows that he interpreted the Lorentz-transformed coordinates and fields as corresponding to apparent measurements by moving observers. In particular, he showed that apparent time measurements fulfill the principle of relativity by proving the transitivity of apparent simultaneity.

CHAPTER 10

Einstein's 1905 Relativity Paper

We now turn to Einstein's seminal paper "Zur Elektrodynamik bewegter Körper" (1905). In this paper, he attempted to resolve the apparent contradiction between the laws of mechanics and electrodynamics by assuming the principle of relativity in conjunction with the principle of the constancy of the speed of light.[1] Einstein's preliminary formulation of the two principles was given in the opening remarks. While the principle of relativity required that "the same laws [...] will be valid for all frames of reference for which the equations of mechanics hold good," the principle of the constancy of the speed of light prescribed that "light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body."[2] The principle of relativity was established as an integral part of Newtonian mechanics. In contrast, the principle of the constancy of the speed of light was traditionally only accepted as a consequence of the wave theory of light. According to this theory, the principle only applied to frames in which the luminiferous ether oscillates around a stationary position. However, unlike Lorentz and Poincaré, Einstein rejected the introduction of such a medium to explain the propagation of light and therefore could not presuppose the existence of an ether-based frame of reference. Instead, he referred the light principle to an arbitrarily chosen inertial frame, whose coordinate system he designated as the resting or stationary system. By also assuming the principle of relativity, he was able to show that the coordinates of two inertial frames can be related by a Lorentz transformation without any reference to the ether. By implementing the principle of relativity at the level of kinematics, Einstein thus transformed it from an epistemological principle into an ontological principle. After deriving the Lorentz transformations, Einstein concluded the kinematical part of his paper by establishing both length contraction and time dilation as consequences of his theory (§3). In the electrodynamic part, he proved the Lorentz invariance of the homogeneous and inhomogeneous Maxwell equations (§6 and §9). In §§7–8 he then demonstrated the power of his approach by applying the transformation equations for coordinates and fields to account for the optical Doppler effect, stellar aberration and the pressure of light on a reflecting surface. In the following, we will limit

[1] Hereafter also referred to as the light principle.
[2] Einstein 1952, pp. 37–8.

ourselves to an analysis of the kinematical part.

10.1 THE APPARENT CONTRADICTION

Galilean transformations were used in Newtonian mechanics to relate inertial frames. It followed that velocities referred to one inertial frame transformed into other inertial frames according to the Galilean rule for the transformation of velocities. This contradicted Einstein's conjunction of relativity principle and light principle. To see this, let us assume that an inertial frame S' moves at constant velocity $\mathbf{u}$ relative to an inertial frame S. Secondly, we assume that the light principle refers to S (i.e. we choose S as the frame associated with Einstein's stationary system). Then the speed of a light pulse emitted in the direction given by $\mathbf{u}$ is c in S and $c - u$ in S'. However, the principle of relativity states that the light principle, in order to express a law of nature, must apply to all inertial frames. A direct consequence of the combination of the principle of relativity and the light principle is therefore the proposition that in all inertial frames light has, regardless of the state of motion of its source, the same constant speed c in all directions. The apparent dilemma was resolved by Einstein's insight that the Newtonian concept of inertial frame was too vague because it did not specify how simultaneity was distributed between distant points in an inertial frame. In particular, Einstein showed how "simultaneity at distant points" could be defined by a method similar to Poincaré's, which did not lead to contradictions between the principle of relativity and the light principle.

10.2 EINSTEIN'S CONSTRUCTION OF NATURAL COORDINATES

Let us take a look at Einstein's account of how to define time throughout an inertial frame. At the beginning of §1 he chose a coordinate system "in which the equations of Newtonian mechanics hold good" and called it the resting system for simplicity's sake. He also noted that the stationary system corresponded to a Cartesian coordinate system for a rigid frame:

> If a material point is at rest relatively to this system of co-ordinates, its position can be defined relatively thereto by the employment of rigid standards of measurement [i.e. rigid measuring rods] and the methods of Euclidean geometry, and can be expressed in Cartesian co-ordinates.[3]

[3] Ibid., p. 38.

Despite Einstein's apparent statement to the contrary, the remaining sections of the 1905 paper showed that neither Newton's second nor third law of motion were really valid in the frame in question. As Torretti has pointed out, to avoid an explicit inconsistency, we must conclude that Einstein only assumed that the first law of motion was valid in the stationary frame.[4] This law postulated the existence of at least one frame (the stationary frame) relative to which all force-free (henceforth *free*) bodies moved in straight lines at uniform translational speeds. While Einstein relied on rigid rods as standard measuring devices for spatial distances, his time measurements were based on the concept of a natural clock that kept true time (as Newton put it).

To mathematically express this condition, we assume that we have a reference frame S to which we want to apply coordinates (x, y, z, t). It follows that the three spatial coordinates of S should encompass a 3-dimensional Euclidean space. If we also assume that the frame is inertial (in Einstein's sense), the trajectory of each free particle must satisfy the equation of motion

$$\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = 0, \tag{10.1}$$

where $\mathbf{r} = (x, y, z)$ denotes the spatial coordinates of the particle and t its time coordinate. The condition that Newton's first law of motion holds in the stationary frame thus restricts the conventionality of the time coordinate function. However, the condition does not define it uniquely. To realize this, we notice that the equation of motion for free particles is invariant under linear transformations of the form

$$\begin{aligned} x^\dagger &= x, & y^\dagger &= y, \\ z^\dagger &= z, & t^\dagger &= t + \mathbf{k}\cdot\mathbf{r}, \end{aligned} \tag{10.2}$$

where $\mathbf{k} = (a, b, c)$ represents an arbitrary constant real-valued 3-dimensional vector.[5] These transformations amount to a redistribution of the relation of simultaneity between events. In particular, the locus of events for which $t = 0$ corresponds to $t^\dagger = \mathbf{k}\cdot\mathbf{r}$. For example, the coordinates of events E_1 and E_2 initially labeled $(1, 1, 1, 0)$ and $(a, b, c, 0)$ transform into $(1, 1, 1, a+b+c)$ and $(a, b, c, a^2+b^2+c^2)$ respectively. As Harvey Brown puts it, in the world of "*free* particles moving in Euclidean 3-space, there

[4] Torretti 1983, p. 51.
[5] Brown 2005, p. 20.

is no privileged notion of simultaneity."[6] Brown therefore concludes that "Newtonian simultaneity is a by-product of the introduction of forces into the [Newtonian] theory."[7] This can be demonstrated by noting that Newton's three laws of motion are invariant only under the Galilean group of all mappings from (x, y, z, t) to $(x^\dagger, y^\dagger, z^\dagger, t^\dagger)$:

$$t^\dagger = t + t_0, \qquad \mathbf{r}^\dagger = A\mathbf{r} + \mathbf{u}t + \mathbf{r}_0, \tag{10.3}$$

where $\mathbf{r}^\dagger = (x^\dagger, y^\dagger, z^\dagger)$, A is an orthogonal 3×3 matrix, $\mathbf{u}$ and $\mathbf{r}_0$ are constant 3-dimensional vectors, and t_0 is a constant.[8] Provided that all three laws of motion are valid, the time coordinate function is therefore uniquely determined in every inertial frame up to time translation, from which it follows that Newtonian simultaneity is absolute. This shows that Einstein, who adhered to Newton's first law of motion but rejected the second and third laws, had the latitude to distribute time differently in his stationary frame.

The particular choice Einstein made was to distribute time throughout the stationary frame by synchronizing clocks at distant positions P_1 and P_2 by sending a light signal back and forth between the two positions and synchronizing the two clocks according to the prescription

$$t_2 = \frac{t_1 + t_3}{2},$$

where t_1 and t_3 were the departure time and the arrival time at P_1 and t_2 was the arrival time at P_2. This means that the clock at P_2 would be synchronous with the clock at P_1 if and only if the light signal was reflected at P_2 when the clock at P_2 read the time $t_2 = \frac{t_1+t_3}{2}$. This definition of time amounted to the stipulation that the speed of the selected signal was constant in opposite directions in a given reference frame. Although Einstein presented this synchronization method as a matter of definition, he had to implicitly assume, as Torretti has pointed out, that it was compatible with (10.1), because otherwise an inertial frame would not satisfy Newton's first law of motion, as was prescribed at the beginning.[9] Since Einstein assumed that space is homogeneous and isotropic, Torretti goes on to explain that it

[6] Ibid., p. 20.

[7] Ibid., p. 20.

[8] For a geometrical interpretation of Galilean group of transformations, see Friedman 1983, pp. 360–1.

[9] Torretti 1983, p. 53.

is possible to reformulate Einstein's definition of time in an inertial frame S into an equivalent stipulation, making the affinity between his definition and the classical rendition of the law of inertia clearer:

> *Light Hypothesis:* A light pulse transmitted through empty space in any direction from a source at rest in F [S] travels in a straight line traversing equal distances in equal times.[10]

10.3 THE PRINCIPLES OF RELATIVITY

After clarifying Einstein's concept of an inertial frame, Torretti goes on to express Einstein's rendition of the principle of relativity and the principle of the constancy of the speed of light in a more consistent and unambiguous way. Einstein's preliminary formulation of the relativity principle required that "the same laws [. . .] will be valid for all frames of reference for which the equations of mechanics hold good."[11] His final presentation of this principle in the 1905 paper, however, stated that

> [t]he laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to one or the other of two systems of co-ordinates in uniform translatory motion.[12]

According to Torretti, the two formulations of the principle of relativity are not equivalent. In the following, I will explain how the Italian attempted to resolve this ambiguity.

Let us assume that we have a reference frame S, which is inertial in Einstein's sense, so that 1) free particles satisfy equation (10.1) in S and 2) light pulses travel in S according to the light hypothesis. If we let (x, y, z, t) denote an inertial coordinate system adapted to S, we can rewrite the light hypothesis as the requirement that any two events $E_1 = (x_1, y_1, z_1, t_1)$ and $E_2 = (x_2, y_2, z_2, t_2)$ related by a light ray satisfy

$$c^2(t_2 - t_1)^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2, \qquad (10.4)$$

[10] Torretti 1983, p. 54.
[11] Einstein 1952, pp. 37–8.
[12] Ibid., p. 41.

where c denotes the speed of light in S. Let (x', y', z', t') be a coordinate system obtained by transforming (x, y, z, t) using the Lorentz transformation

$$\begin{aligned} t' &= \gamma\left(t - \frac{ux}{c^2}\right), \\ x' &= \gamma(x - ut), \\ y' &= y, \\ z' &= z, \end{aligned} \tag{10.5}$$

where $\gamma = (1 - \frac{u^2}{c^2})^{-1/2}$. At this point Torretti observes that (x', y', z', t') is adapted to a frame S' moving at constant velocity $\mathbf{u}$ relative to S and that the equations (10.1) and (10.4) are invariant under the transformations (10.5). It follows (unsurprisingly) that (x', y', z', t') is an inertial coordinate system and S' is an inertial frame in Einstein's sense. Next, Torretti constructs a coordinate system $(x^\dagger, y^\dagger, z^\dagger, t^\dagger)$ that is adapted to the spatial part of S' by transforming (x', y', z', t') by a linear transformation of the form (10.2). As just mentioned, equation (10.1) is invariant under such transformations. Despite this fact, however, $(x^\dagger, y^\dagger, z^\dagger, t^\dagger)$ is not inertial, since the transformation of (x', y', z', t') into $(x^\dagger, y^\dagger, z^\dagger, t^\dagger)$ redistributes the simultaneity relation between events in such a way that equation (10.4) does not hold relative to $(x^\dagger, y^\dagger, z^\dagger, t^\dagger)$. Therefore, $(x^\dagger, y^\dagger, z^\dagger, t^\dagger)$ and the frame $S^\dagger$ to which it is adapted are not inertial. This means that Torretti has constructed a non-inertial frame $S^\dagger$ in uniform translational motion with respect to the inertial frame S.

So, on the one hand, Einstein's final version of the relativity principle applies between the frames S and $S^\dagger$. But, on the other hand, the principle does not apply in Einstein's preliminary version. Moreover, this circumstance contradicts the physical equivalence of the two frames assumed in the final version of the principle, which means that Einstein's final version was not quite as restrictive as he intended. As can be seen from the preliminary version, he wanted to formulate the principle in such a way that inertial frames were physically equivalent or indistinguishable. To eliminate this inconsistency, Torretti slightly modifies Einstein's final version of the principle of relativity and the light principle:

> *Relativity Principle:* The laws by which the states of physical systems undergo change are not affected, whether these changes of state refer to one or the other of two Lorentz charts [(Einsteinian) inertial coordinate systems]. [...]

Light Principle: Any ray of light moves in the stationary frame with constant velocity c *in vacuo*, whether the ray is emitted by a stationary or a moving body.[13]

10.4 EINSTEIN'S DERIVATION OF THE LORENTZ TRANSFORMATIONS

To derive the transformation equations between two of these inertial coordinate systems, Einstein introduced an inertial frame S at rest and an inertial frame S' moving with velocity $\mathbf{u}$ relative to S. Adapting inertial coordinate systems (x, y, z, t) and (x', y', z', t') to S and S' so that the y-axis and the z-axis were both perpendicular to the x-axis and the the x'-axis was paralllel with the x-axis and moved along it with speed u, his task was to find the equations for the transformation between (x, y, z, t) and (x', y', z', t'). These transformations had to be linear, he noted right at the beginning, "on account of the properties of homogeneity that we attribute to space and time."[14]

He then introduced an auxiliary coordinate system $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ expressed by (x, y, z, t):

$$\begin{aligned} \overline{x} &= x - ut, & \overline{y} &= y, \\ \overline{z} &= z, & \overline{t} &= t. \end{aligned} \tag{10.6}$$

Since this coordinate system moved with a velocity $\mathbf{u}$ along the x-axis in S, it followed that it was adapted to S'. However, for the sake of clarity, we will introduce $\overline{S}$ and distinguish between (x', y', z', t') in S' and $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ in $\overline{S}$.

Einstein went on to consider the synchronization of two clocks situated on the x'-axis in S'. While the first clock C_0 was placed at a position P_0 corresponding to the origin in S', the second clock C_1 was located at P_1 at a distance x' from P_0. To synchronize clock C_1 using clock C_0 as reference, a light beam was emitted at P_0 at time t'_0 as read on C_0, received at P_1, reflected back to P_0 and arriving at time t'_2 as read on C_0. On arrival at P_1 clock C_1 was set to

$$t'_1 = \frac{t'_0 + t'_2}{2}. \tag{10.7}$$

Without further ado, he went on to express t'_0, t'_1 and t'_2 as functions of

[13] Torretti 1983, p. 54.

[14] Einstein 1952, p. 44. For a proof of Einstein's proposition see Rindler 1982, p. 14.

$(\overline{x}, \overline{y}, \overline{z}, \overline{t})$:

$$t'\left(\overline{x}, 0, 0, \overline{t} + \frac{\overline{x}}{c-u}\right) = \frac{t'\left(0, 0, 0, \overline{t}\right) + t'\left(0, 0, 0, \overline{t} + \frac{\overline{x}}{c-u} + \frac{\overline{x}}{c+u}\right)}{2}. \tag{10.8}$$

Let us fill in some details. Without loss of generality, we assume that the spatial origins of S and S' coincide at $t = t' = 0$. Using (10.6) it follows that S' and $\overline{S}$ have the same spatial origin.

We will therefore prove (10.8) for $t'_0 = t = \overline{t}$. The light beam leaves P_0, the spatial origin of S' (and $\overline{S}$), at $t'_0 = 0$ in S' corresponding to $t = 0$ in S. This event E_0 has coordinates $(0, 0, 0, 0)$ in $\overline{S}$, so that

$$t'_0 = t'(0, 0, 0, 0). \tag{10.9}$$

The light beam arrives at P_1 a distance x' apart from P_0 at t'_1 in S'. To the x'-abscissa in S' corresponds a $\overline{x}$-abscissa in $\overline{S}$. P_1 has therefore spatial coordinates $(\overline{x}, 0, 0)$ in $\overline{S}$. The arrival time t_1 as registered in S can be calculated by means of the light principle:

$$x = ct_1. \tag{10.10}$$

Applying (10.6) it follows that

$$\overline{x} = ct_1 - ut_1 = (c-u)t_1.$$

If we solve for t_1, we obtain the arrival time as measured in S

$$t_1 = \frac{\overline{x}}{c-u},$$

and thus t'_1 as a function of $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ is

$$t'_1 = t'\left(\overline{x}, 0, 0, \frac{\overline{x}}{c-u}\right). \tag{10.11}$$

The event E_2, which corresponds to the reception of the return signal from P_1, takes place at the origin in S' and $\overline{S}$. To calculate the travel time t_2 of the return signal as registered in S, we note that event E_1 has coordinates $(\overline{x}, 0, 0, t_1)$ in $\overline{S}$ and $(x, 0, 0, t_1)$ in S, while event E_2 has the coordinates

$(0, 0, 0, t_1 + t_2)$ in $\overline{S}$ and $(u\,(t_1 + t_2), 0, 0, t_1 + t_2)$ in S. Consequently, the signal must cover a distance $x - u(t_1 + t_2)$ in S. According to the light principle in S, we therefore obtain

$$ct_2 = x - u(t_1 + t_2). \tag{10.12}$$

If we substitute $x = \overline{x} + ut_1$ into this equation and solve for t_2, the result is

$$t_2 = \frac{\overline{x}}{c + u}. \tag{10.13}$$

t_2' as a function of $(\overline{x}, \overline{y}, \overline{z}, \overline{t})$ is therefore

$$t_2' = t'\left(0, 0, 0, \frac{\overline{x}}{c - u} + \frac{\overline{x}}{c + u}\right). \tag{10.14}$$

If we substitute (10.9), (10.11) and (10.14) into (10.7), we finally get (10.8) for $t = \overline{t} = 0$:

$$t'\left(\overline{x}, 0, 0, \frac{\overline{x}}{c - u}\right) = \frac{t'(0, 0, 0, 0) + t'\left(0, 0, 0, \frac{\overline{x}}{c-u} + \frac{\overline{x}}{c+u}\right)}{2}. \tag{10.15}$$

Einstein went on to differentiate both sides of this equation with respect to $\overline{x}$:

$$\begin{aligned} &\frac{\partial t'}{\partial \overline{x}} + \frac{1}{c - u}\frac{\partial t'}{\partial \overline{t}} = \frac{1}{2}\left(\frac{1}{c - u} + \frac{1}{c + u}\right)\frac{\partial t'}{\partial \overline{t}} \\ \Rightarrow \quad &\frac{\partial t'}{\partial \overline{x}} + \frac{u}{c^2 - u^2}\frac{\partial t'}{\partial \overline{t}} = 0. \end{aligned} \tag{10.16}$$

This differential equation provided the functional dependence of t' on $\overline{x}$ and $\overline{t}$. By imagining light beams traveling back and forth along the y'-axis or z'-axis, Einstein also obtained the differential equations for the functional dependence of t' on $\overline{y}$ and $\overline{z}$

$$\frac{\partial t'}{\partial \overline{y}} = 0, \quad \frac{\partial t'}{\partial \overline{z}} = 0, \tag{10.17}$$

implying that t' was independent of $\overline{y}$ and $\overline{z}$. Since t' should be a linear function, this meant that

$$t' = a(u)\left(\overline{t} - \frac{u}{c^2 - u^2}\overline{x}\right), \tag{10.18}$$

where $a(u)$ was an unknown function (to be determined later).

Einstein's next step was to ensure that the light principle was fulfilled in S'. For a light beam emitted at $t' = 0$ in the direction of increasing x', this implied

$$x' = ct'.$$

Substituting (10.18) into this equation, he obtained

$$x' = a(u)c\left(\bar{t} - \frac{u}{c^2 - u^2}\bar{x}\right).$$

Since the light beam also had to move at speed c in S,

$$x = ct,$$

from $\bar{x} = x - ut$ and $t = \bar{t}$ it followed that

$$\bar{t} = \frac{\bar{x}}{c - u}.$$

Substituting this equation into the above expression for x' then resulted in

$$x' = a(u)\frac{c^2}{c^2 - u^2}\bar{x}. \tag{10.19}$$

By requiring that light beams along the y'-axis also fulfill the light principle, he obtained

$$\bar{t} = \frac{y}{\sqrt{c^2 - u^2}}.$$

By inserting this and $\bar{x} = 0$ into

$$y' = ct' = a(u)c\left(\bar{t} - \frac{u}{c^2 - u^2}\bar{x}\right),$$

he could write y' as

$$y' = a(u)\gamma y. \tag{10.20}$$

Similar considerations applied to z', so that

$$z' = a(u)\gamma z. \tag{10.21}$$

By substituting $\overline{x} = x - ut$ into (10.18) and introducing $\phi(u) = a(u)\gamma$, he collected his set of transformation equations (10.18)-(10.21) thus far:

$$\begin{aligned} x' &= \phi(u)\gamma(x - ut), \\ y' &= \phi(u)y, \\ z' &= \phi(u)z, \\ t' &= \phi(u)\gamma\Big(t - \frac{u}{c^2}x\Big). \end{aligned} \tag{10.22}$$

As Torretti points out, Einstein only proved this set of transformations for special events, but he attributed general validity to it. The reason why (10.22) actually has general validity, Torretti continues, arises from the fact that it s a linear transformation of $\mathbb{R}^4$, "which is therefore fully determined by its value at four linear independent points."[15] Since Einstein's synchronization events satisfy this condition, generality follows.[16]

After Einstein had proved the validity of the light principle in S', he demonstrated its consistency with the principle of relativity. To illustrate this, he suggested that at time $t = t' = 0$ a ray of light was emitted at the common origin of S and S'. Taking into account a point (x, y, z) in S that the wave reached at time t, it followed from the light principle in S that

$$x^2 + y^2 + z^2 = c^2t^2. \tag{10.23}$$

Transforming this using (10.22), he obtained

$$x'^2 + y'^2 + z'^2 = c^2t'^2. \tag{10.24}$$

Consequently, the corresponding wave in S' was also spherical and traveled at speed c. From this, Einstein concluded that the transformations (10.22) were consistent with both principles.

To determine $\phi(u)$, Einstein introduced a third inertial system $S^\dagger$ that moved at velocity $-\mathbf{u} = (-u, 0, 0)$ in S'. Assuming that the origins in $S^\dagger$, S' and S coincided at time $t^\dagger = t' = t = 0$, he was able to relate the three inertial systems to each other by a two-step application of (10.22). The

[15] Torretti 1983, p. 60.
[16] For details, see ibid., p. 60.

transformations from S' to $S^\dagger$ resulted in

$$\begin{aligned} x^\dagger &= \phi(-u)\gamma(x' + ut'), \\ y^\dagger &= \phi(-u)y', \\ z^\dagger &= \phi(-u)z', \\ t^\dagger &= \phi(-u)\gamma\left(t' + \frac{u}{c^2}x'\right). \end{aligned} \tag{10.25}$$

Rewriting (10.22) as

$$\begin{aligned} x' + ut' &= \phi(u)\gamma^{-1}x, \\ y' &= \phi(u)y, \\ z' &= \phi(u)z, \\ t' + \frac{u}{c^2}x' &= \phi(u)\gamma^{-1}t, \end{aligned} \tag{10.26}$$

and inserting the result into (10.25) gave transformations from S to $S^\dagger$:

$$\begin{aligned} x^\dagger &= \phi(u)\phi(-u)x, \\ y^\dagger &= \phi(u)\phi(-u)y, \\ z^\dagger &= \phi(u)\phi(-u)z, \\ t^\dagger &= \phi(u)\phi(-u)t. \end{aligned} \tag{10.27}$$

As the relationships between the spatial coordinates were not dependent on time, he concluded that S and $S^\dagger$ must be identical, implying that

$$\phi(u)\phi(-u) = 1. \tag{10.28}$$

Einstein then determined $\phi(u)$ by considering a rod of length l oriented along the y'-axis as measured by an observer in S'. Since its end points in S' had coordinates $y_1' = 0$ and $y_2' = l$, it followed from (10.22) that the length of the rod as measured by an observer in S corresponded to

$$y_2 - y_1 = \frac{l}{\phi(u)}. \tag{10.29}$$

For reasons of symmetry, the length of the rod as measured in S could only depend on the magnitude of the relative velocity and not on the direction. Consequently, $\frac{l}{\phi(u)} = \frac{l}{\phi(-u)}$ or

$$\phi(u) = \phi(-u). \tag{10.30}$$

Inserting this into (10.28) gave

$$\phi(u)^2 = 1. \tag{10.31}$$

Since the negative solution was not physically feasible, it followed that

$$\phi(u) = 1. \tag{10.32}$$

By substituting this result into (10.22), he finally obtained the Lorentz transformations:

$$\begin{aligned} x' &= \gamma(x - ut), \\ y' &= y, \\ z' &= z, \\ t' &= \gamma\left(t - \frac{u}{c^2}x\right). \end{aligned} \tag{10.33}$$

10.5 LENGTH CONTRACTION AND TIME DILATION

Einstein went on to derive the so-called Lorentz contraction as a consequence of his theory by imagining a spherical body of radius R at rest in S' and with its center at the origin:

$$x'^2 + y'^2 + z'^2 = R^2. \tag{10.34}$$

To evaluate the shape of this body as observed in S, we must take an instantaneous slice in S, i.e. we impose $t = 0$. By substituting (10.33) for $t = 0$ into (10.34), Einstein obtained

$$(\gamma x)^2 + y^2 + z^2 = R^2.$$

It followed that the moving body had the shape of an ellipsoid with axes $(\frac{R}{\gamma}, R, R)$ in S. While the dimensions of the body perpendicular to the direction of motion did not change when measured by the stationary observer, the dimension parallel to the motion appeared to be contracted by the factor $\frac{1}{\gamma}$.

Next, Einstein imagined a clock attached to the origin of S'. This clock would move in S according to the equation $x = vt$. Substituting $x = vt$ into (10.33), the result was

$$t' = \frac{t}{\gamma}.$$

The stationary observer would conclude from this that the moving clock runs slower than his stationary clocks.

10.6 CONCLUSION

In 1905, Einstein proved how to obtain the Lorentz transformations using only the principle of relativity and the light principle. In particular, he did not introduce a luminiferous ether and accordingly did not invoke an ether-based frame of reference. Moreover, Einstein, like Poincaré, obtained the exact invariance of Maxwell's equations under Lorentz transformations in accordance with the principle of relativity. Although the Lorentz invariance implied that it was impossible to empirically distinguish the ether-based frame from any other inertial frame, Poincaré retained the ether as the carrier of the electric and magnetic fields and distinguished between real and apparent motion on this basis. Einstein, on the other hand, was able to interpret the principle of relativity ontologically by eliminating any reference to the ether. In particular, he explained length contraction not as a dynamical influence of the ether on matter, but as a kinematical consequence of his revolutionary concepts of space and time. Furthermore, in contrast to Poincaré, Einstein realized that the physical interpretation of the Lorentz transformations implied that an observer at rest in an inertial frame S would find that clocks moving relative to S run slower than clocks at rest in S.

CHAPTER 11

Relativity Theory and Space-Time Structure

This chapter is devoted to an alternative derivation of the Lorentz transformations, which goes back to the work of a contemporary of Einstein, the Russian physicist Vladimir Sergeyevitch Ignatowski.[1] Since its introduction, this approach has been repeated and refined in numerous articles such as Jean-Marc Lévy-Leblond's "One More Derivation of the Lorentz Transformation" (1976) or the paper "Lorentz Transformations from the First Postulate" (1975), which was co-authored by A. R. Lee and T. M. Kalotas. Because special relativity provides a common space-time framework for all other theories, the starting point of this approach is the insight that the formulation of special relativity should not be dependent on principles derived from more specialized theories. As a result, electrodynamics, and particularly the constancy of the speed of light, should not play a central role in the derivation of the Lorentz transformations.

11.1 THE EXCLUSION OF THEORY-SPECIFC POSTULATES

In the introduction to "Einige allgemeine Bemerkungen zum Relativitätsprinzip" (1910), Ignatowski expressed a preliminary version of the principle of relativity by imagining frames in uniform translational motion relative to each other:

> No experiment that we carry out in one of the frames can determine the latter's absolute motion, but only its relative motion with respect to other frames. We thus have to perceive this frame as stationary. All frames are equivalent.[2]

Following Einstein, Ignatowski thus interpreted the relativity of motion ontologically in the sense that the empirical indistinguishability of inertial frames implied their ontological equivalence. In contrast to Einstein, however, he claimed that – apart from some very general postulates about space, time and motion such as the homogeneity and isotropy of space – the principle of relativity was sufficient to single out so-called "inertial transformations" up to a certain universal limiting speed associated with the

[1] Ignatowski 1910; Ignatowski 1911.

[2] Ignatowski 1910, p. 1–2. My translation. Literally, Ignatowski referred to worlds ("Welten") instead of frames in this passage. However, later on, he used "worlds" and "frames of reference" ("Bezugssysteme") interchangeably.

structure of space-time itself and whose value was then (contended to be) empirically ascertainable. The existence of such a universal constant was accordingly a consequence rather than a prerequisite of the theory:

> It is possible to prove the existence of a universal space constant solely by means of the principle of relativity. This approach is contrary to A. Einstein, who, next to the principle of relativity, at the outset presupposes the speed of light as a universal constant.[3]

In addition to avoiding a theory-specific "second postulate," Ignatowski argued that his approach also had the advantage of clarifying the meaning of the principle of relativity itself:

> By means of the transformation, we obtain a new mathematical expression for the law. However, because of the principle of relativity, the law must preserve its form. If this is not the case, we have to conclude that the law itself is invalid. The law must have precisely such a form that the latter does not change under the transformation. The principle of relativity plays, so to say, the role of a controlling authority during the construction of the mathematical form of physical laws.[4]

By singling out inertial transformations, the principle made it possible to transform any physical law from one inertial frame to another. It followed that we should regard the principle of relativity as a controlling authority ("Kontrollinstanz") in the sense of a meta-principle that prescribed the form of every empirical law.

11.2 CLOCK SYNCHRONIZATION

Ignatowski went on to discuss length and time measurements. While the former required rigid rods as standards of length, the latter required synchronized clocks at different locations in the reference frame. In contrast to Poincaré and Einstein, Ignatowski did not limit himself to clock synchronization by means of light signals, as he tried not to give optics a special status:

> Based on what mechanism do we determine the synchronism? It is straightforward; by means of an arbitrary physical method. Let it be

[3] Ibid., p. 1. My translation.
[4] Ibid., p. 2. My translation.

> mechanical or optical. Because if a physical measurement determines that two clocks at rest with respect to us run synchronously, they do indeed run synchronously. As a result of this, we have determined the synchronism unambiguously, because it can always be tested by a new experiment.[5]

Given its importance for Ignatowski's approach, it would have been helpful to have at least one example of a mechanical method for clock synchronization. Even if he did not give any examples himself, it is not difficult to find some in the literature. Wolfgang Rindler, for example, has argued that it would be possible to synchronize clocks by utilizing sound signals:

> [I]t will be well to point out that identical coordinates can be assigned *without* the use of light signals – though perhaps less conveniently. For example, [...] clocks could be synchronized by a *sound* signal from the origin if the frame were filled with still air, or by rifle bullets of known velocity shot from the origin in all directions at time t_0. It is clear that *if* there exists a time in terms of which the physics in the frame is isotropic, our above methods have determined it, since they are based on the isotropic propagation of light, sound, rifle bullets, or whatever.[6]

Following Ignatowski, Rindler thus emphasized that clock synchronization did not attribute any special status to the propagation of light, but merely presupposed the isotropy of space.

11.3 DERIVING THE INERTIAL TRANSFORMATIONS

We now proceed to the derivation of the inertial transformations, i.e. the transformation equations between inertial frames. We will take Lévy-Leblond's derivation as our starting point, as Ignatowski left some details unclear. The Frenchman began his reasoning by repeating Ignatowski's understanding of special relativity:

> By establishing special relativity on a property of the speed of light, one seems to link this theory to a restricted class of phenomena, namely, electromagnetic radiations. However, [...] [w]e believe that

[5] Ignatowski 1910, p. 2. My translation.
[6] Rindler 1982, p. 11.

> special relativity at the present time stands as a universal theory describing the structure of a common space-time arena in which all fundamental processes take place. All the laws of physics are constrained by special relativity acting as a sort of "super law,"[4] and electromagnetic interactions here have no privilege other than a historical or anthropocentric one.[7]

Starting from the principle of relativity in terms of an infinite, continuous class of equivalent reference frames related to each other by a class of inertial transformations, he then went on to determine the members of this class by successively imposing four general hypotheses: homogeneity of space-time, isotropy of space, inertial transformations constitute a group action on the set of equivalent frames, causality condition. As we shall see below, it turned out that the combination of the principle of relativity and these four hypotheses was sufficient "to single out the Lorentz transformations and their degenerate Galilean limit as the only possible inertial transformations."[8]

Instead of examining the general case, Lévy-Leblond restricted his derivation to the nature of inertial transformations in the two-dimensional case. More precisely, he determined the inertial transformation of the spatio-temporal coordinates (x, t) of an arbitrary event in an inertial frame S to its coordinates (x', t') in another inertial frame S'. The first thing he noticed was that "since we have assumed the existence of an infinite continuous class of inertial frames, the relationship between any two of them depends upon a certain number of parameters $\{a_1, \ldots, a_N\}$, the values of which characterize any special inertial transformation:"[9]

$$\begin{aligned} x' &= f(x, t; a_1, \ldots, a_N), \\ t' &= g(x, t; a_1, \ldots a_N). \end{aligned} \tag{11.1}$$

He then argued that the space and time translations

$$x' = x + \xi, \qquad t' = t + \tau \tag{11.2}$$

had to be members of the class of inertial transformations because we may

[7] Lévy-Leblond 1976, p. 271.
[8] Ibid., p. 276.
[9] Ibid., p. 272.

choose space and time origins arbitrarily.[10] It followed that he could restrict his investigation to the class of inertial transformations with common space-time origin, which implied that he could reduce the number of parameters in his transformation formulas (11.1) from N to $n = N - 2$:

$$\begin{aligned} x' &= F(x, t; a_1, \dots, a_n), \\ t' &= G(x, t; a_1, \dots, a_n), \end{aligned} \tag{11.3}$$

with

$$\begin{aligned} 0 &= F(0, 0; a_1, \dots, a_n), \\ 0 &= G(0, 0; a_1, \dots, a_n). \end{aligned} \tag{11.4}$$

Next, Lévy-Leblond argued that inertial transformations between inertial frames with common origin depended on only one parameter, which reduced (11.3) to

$$x' = F(x, t; a), \qquad t' = G(x, t; a), \tag{11.5}$$

and the conditions (11.4) to

$$0 = F(0, 0; a), \qquad 0 = G(0, 0; a). \tag{11.6}$$

To see this, we were told to consider a moving object with the equation of motion $x = \varphi(t)$ and the initial condition $\varphi(0) = 0$ in S. According to (11.1), the corresponding equation of motion and its initial condition in S' could then be expressed by

$$x' = \varphi'(t') = F(\varphi(t), t; a_1, \dots, a_n), \tag{11.7}$$

$$0 = \varphi'(0) = F(0, 0; a_1, \dots, a_n). \tag{11.8}$$

Writing $\varphi(t)$ in terms of its Taylor expansion at $t = 0$ and using the chain rule, it followed that any derivative of x' with respect to t' in S' "will depend upon the speed, acceleration, etc., of the object at the origin in the first inertial frame [S] *and* upon the n parameters."[11] From this he concluded that $n = 1$:

[10] Lévy-Leblond did not bother to write down the freedom to choose the origins of space and time as an independent hypothesis because it turns out to be a special case of the fundamental postulate of the homogeneity of space-time, which we will introduce in a moment.

[11] Lévy-Leblond 1976, p. 272.

> We know from simple physical experience that speed, indeed, is only relative and can be varied from one inertial frame to the other; this is the empirical basis of the principle of relativity. We know, though, that the same is not true for acceleration, which is associated with physical effects differentiating various frames. It follows that $n = 1$.[12]

Let us go over this point in more depth. By varying $a_1, \ldots, a_n$ we can find an inertial frame S' in which the derivatives of φ' up to nth order have arbitrarily preassigned values due to the dependence of φ' and its derivatives (with respect to t') on the values of the derivatives of φ at $t = 0$ and $a_1, \ldots, a_n$. However, physical experience has taught us that we can only identify inertial frames in which the velocity $\varphi^1(0)$ is arbitrary. From this we can conclude that there is only one degree of freedom.

11.3.1 Homogeneity of Space-Time

We now turn to the assumption of the homogeneity of space-time, which meant that it had the same properties at all locations and at all times. According to Lévy-Leblond, this assumption could be formulated more precisely as the hypothesis that "the transformation properties of a spatio-temporal interval $(\Delta x, \Delta t)$ depend only on that interval and not on the location of its endpoints (in the considered reference frame)."[13] Since $\mathrm{d}x'$ and $\mathrm{d}t'$ could be expressed as

$$\begin{aligned} \mathrm{d}x' &= \frac{\partial F}{\partial x}\mathrm{d}x + \frac{\partial F}{\partial t}\mathrm{d}t, \\ \mathrm{d}t' &= \frac{\partial G}{\partial x}\mathrm{d}x + \frac{\partial G}{\partial t}\mathrm{d}t, \end{aligned} \tag{11.9}$$

it followed that the coefficients $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial t}$, $\frac{\partial G}{\partial x}$ and $\frac{\partial G}{\partial t}$ had to be independent of x and t. On the other hand, the calculation of the differentials of F and G resulted in

$$\begin{aligned} \mathrm{d}F &= \frac{\partial F}{\partial x}\mathrm{d}x + \frac{\partial F}{\partial t}\mathrm{d}t, \\ \mathrm{d}G &= \frac{\partial G}{\partial x}\mathrm{d}x + \frac{\partial G}{\partial t}\mathrm{d}t. \end{aligned} \tag{11.10}$$

[12] Ibid., p. 272.

[13] Ibid., p. 272.

Together with (11.9) this implied that F and G had to be linear functions of x and t, so that Lévy-Leblond could write

$$x' = H(a)x - K(a)t, \tag{11.11a}$$
$$t' = L(a)t - M(a)x, \tag{11.11b}$$

using that $x'(0,0) = t'(0,0) = 0$ and choosing signs for later convenience.

He then defined *inertial motions* as those obtained by applying an inertial transformation to the coordinates of an object at rest in an inertial frame. According to (11.11a), inertial motions therefore fulfilled the equation of motion

$$H(a)x - K(a)t = C, \tag{11.12}$$

where C was a constant that corresponded to the spatial position x' of an object at rest in the inertial frame S'. Stated otherwise, inertial motions were uniform motions at speed $v = \frac{K(a)}{H(a)}$. It followed that Lévy-Leblond could use v as parameter instead of the pseudo parameter a. More precisely, defining functions

$$\gamma(v) = H(a), \tag{11.13a}$$
$$\lambda(v) = \frac{L(a)}{H(a)}, \tag{11.13b}$$
$$\mu(v) = \frac{M(a)}{H(a)}, \tag{11.13c}$$

it was possible to rewrite the general transformation formulas in the form

$$\begin{aligned} x' &= \gamma(v)(x - vt), \\ t' &= \gamma(v)(\lambda(v)t - \mu(v)x). \end{aligned} \tag{11.14}$$

11.3.2 Isotropy of Space

Lévy-Leblond went on to introduce the hypothesis of the isotropy of space, which stated the physical equivalence of all possible orientations. In a one-dimensional space, for example, the postulated physical equivalence reduced to the two possible orientations of the spatial axis. He then examined an arbitrary event with coordinates (x, t) in S and (x', t') in S', so that (x', t') was related to (x, t) by an inertial transformation (11.14) with parameter v. Next, he defined the frames R and R' by requiring the corresponding coordinates (X, T) and (X', T') to be related to the coordinates (x, t) in S and (x', t') in S' by

$$\begin{aligned} X &= -x, & T &= t, \\ X' &= -x', & T' &= t'. \end{aligned} \tag{11.15}$$

Assuming that S and S' were inertial frames, the isotropy of space implied that R and R' were also inertial frames. Consequently, an inertial transformation of the form (11.14) with some parameter u existed between (X, T) and (X', T'):

$$\begin{aligned} X' &= \gamma(u)(X - uT), \\ T' &= \gamma(u)(\lambda(u)T - \mu(u)X). \end{aligned} \tag{11.16}$$

Replacing X, T, X', and T' by their corresponding expressions in x, t, x' and t', it followed that

$$\begin{aligned} -x' &= \gamma(u)(-x - ut), \\ t' &= \gamma(u)(\lambda(u)t + \mu(u)x). \end{aligned} \tag{11.17}$$

A comparison between (11.14) and (11.17) then led to

$$\gamma(u) = \gamma(v), \tag{11.18a}$$

$$u\gamma(u) = -v\gamma(v), \tag{11.18b}$$

$$\lambda(u)\gamma(u) = \lambda(v)\gamma(v), \tag{11.18c}$$

$$\mu(u)\gamma(u) = \mu(v)\gamma(v). \tag{11.18d}$$

From (11.18a) and (11.18b) Lévy-Leblond was able to deduce that

$$u = -v, \tag{11.19}$$

which was the result one would expect. Substituting $u = -v$ into (11.18c) – (11.18d) and using (11.18a), he obtained the following parity properties for γ, λ, and μ:

$$\gamma(-v) = \gamma(v), \tag{11.20a}$$

$$\lambda(-v) = \lambda(v), \tag{11.20b}$$

$$\mu(-v) = -\mu(v). \tag{11.20c}$$

11.3.3 Inertial Transformations as Group Action

Lévy-Leblond continued his derivation of inertial transformations by introducing the hypothesis that the physical equivalence of inertial frames entailed that the set of inertial transformations constituted a group action on the set of inertial frames. This implied that the following three conditions had to be satisfied:

(a) Identity transformation. There had to exist a neutral element in the group of inertial transformations mapping (x, t) onto itself. The identity transformation clearly corresponded to the inertial transformation with $v = 0$. It followed that

$$\gamma(0) = 1, \tag{11.21a}$$

$$\lambda(0) = 1, \tag{11.21b}$$

$$\mu(0) = 0. \tag{11.21c}$$

(b) Inverse transformation. Assuming that (x', t') was obtained from (x, t) by an inertial transformation with parameter v, there had to exist an inverse transformation from (x', t') to (x, t) of the form (11.14) with some parameter w. This required

$$\begin{aligned} x &= \gamma(w)(x' - wt'), \\ t &= \gamma(w)(\lambda(w)t' - \mu(w)x'). \end{aligned} \tag{11.22}$$

In addition, the inversion of (11.14) resulted in

$$\begin{aligned} x &= \frac{1}{\gamma(v)}\left(1 - \frac{v\mu(v)}{\lambda(v)}\right)^{-1}\left(x' + \frac{v}{\lambda(v)}t'\right), \\ t &= \frac{1}{\gamma(v)}\left(1 - \frac{v\mu(v)}{\lambda(v)}\right)^{-1}\left(\frac{1}{\lambda(v)}t' + \frac{\mu(v)}{\lambda(v)}x'\right). \end{aligned} \tag{11.23}$$

By comparing (11.23) and (11.22), Lévy-Leblond obtained equations relating the unknown w to v and constraining γ, λ, and μ:

$$w = -\frac{v}{\lambda(v)}, \tag{11.24a}$$

$$\lambda(w) = \frac{1}{\lambda(v)}, \tag{11.24b}$$

$$\mu(w) = -\frac{\mu(v)}{\lambda(v)}, \tag{11.24c}$$

$$\gamma(w) = \frac{1}{\gamma(v)}\left(1 - \frac{v\mu(v)}{\lambda(v)}\right)^{-1}. \tag{11.24d}$$

The comparison of (11.24a) and (11.24b) then led to a functional equation for λ:

$$\lambda(\tfrac{-v}{\lambda(v)}) = \tfrac{1}{\lambda(v)}. \tag{11.25}$$

Defining an associated function ξ by

$$\xi(v) \stackrel{\text{def}}{=} \frac{v}{\lambda(v)}, \tag{11.26}$$

Lévy-Leblond claimed that $\lambda(0) = 1$ (by means of (11.21b)) implied

$$\left.\frac{\mathrm{d}\xi}{\mathrm{d}v}\right|_{v=0} = 1. \tag{11.27}$$

To see this, we differentiate ξ with respect to v:

$$\frac{\mathrm{d}\xi}{\mathrm{d}v} = \frac{\lambda\frac{\mathrm{d}v}{\mathrm{d}v} - v\frac{\mathrm{d}\lambda}{\mathrm{d}v}}{\lambda^2} = \frac{\lambda - v\frac{\mathrm{d}\lambda}{\mathrm{d}v}}{\lambda^2}. \tag{11.28}$$

On the other hand, (11.20b) constrained λ to be an even function, so that (11.25) resulted in

$$\lambda(\tfrac{v}{\lambda(v)}) = \tfrac{1}{\lambda(v)}. \tag{11.29}$$

Substituting $\xi(v)$ for $v/\lambda(v)$ in (11.29), Lévy-Leblond obtained

$$\xi(\xi(v)) = v, \tag{11.30}$$

or

$$\xi^{-1}(v) = \xi(v). \tag{11.31}$$

This meant that, as figure 11.1 vividly illustrates, the graph of $\xi(v)$ had to be symmetrical with respect to the line $\xi = v$. And since $\frac{\mathrm{d}\xi}{\mathrm{d}v}(0) = 1$, the

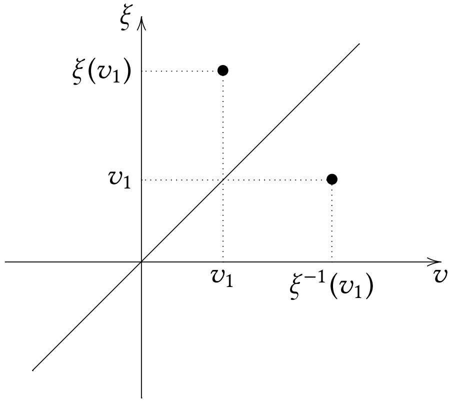


Fig. 11.1: The graph of $\xi(v)$ is symmetrical with respect to $\xi = v$ due to the condition $\xi(v) = \xi^{-1}(v)$.

graph of $\xi(v)$ had to be tangent to the line $\xi = v$. However, since λ and ξ were continuous functions,[14] it followed that $\xi(v) = v$, which implied

$$\lambda(v) = 1. \tag{11.32}$$

By substituting this result into (11.24a), Lévy-Leblond found that the parameter w of the transformation (11.22) was given by

$$w = -v, \tag{11.33}$$

in accordance with our physical intuitions. Substituting this result into equation (11.24d) and replacing $\gamma(-v)$ with $\gamma(v)$ in accordance with (11.20a), he derived the following relationship between the functions γ and μ:

$$\gamma(v)^2(1 - v\mu(v)) = 1. \tag{11.34}$$

(c) Composition law. By performing two inertial transformations of the form (11.14) successively from (x, t) to (x_1, t_1) and then from (x_1, t_1) to (x_2, t_2),

[14] Here I follow Lévy-Leblond, who did not include a rigorous proof because of the high degree of difficulty of this result.

taking into account that $\lambda(v) = 1$, he obtained

$$\begin{aligned} x_1 &= \gamma(v_1)(x - v_1 t), \\ t_1 &= \gamma(v_1)(t - \mu(v_1)x), \end{aligned} \tag{11.35}$$

$$\begin{aligned} x_2 &= \gamma(v_2)(x_1 - v_2 t_1), \\ t_2 &= \gamma(v_2)(t_1 - \mu(v_2)x_1). \end{aligned} \tag{11.36}$$

It followed that the direct transformation from (x, t) to (x_2, t_2) was

$$x_2 = \gamma(v_1)\gamma(v_2)(1 + \mu(v_1)v_2)\left(x - \frac{v_1 + v_2}{1 + \mu(v_1)v_2}t\right), \tag{11.37a}$$

$$t_2 = \gamma(v_1)\gamma(v_2)(1 + v_1\mu(v_2))\left(t - \frac{v_1 + v_2}{1 + v_1\mu(v_2)}x\right). \tag{11.37b}$$

The composition law then required that (11.37) could be identified with a transformation of the form (11.14) with some parameter V:

$$\begin{aligned} x_2 &= \gamma(V)(x - Vt), \\ t_2 &= \gamma(V)(t - \mu(V)x). \end{aligned} \tag{11.38}$$

Identifying $\gamma(V)$ in (11.37a) and (11.37b) implied

$$\gamma(V) = \gamma(v_1)\gamma(v_2)\big(1 + \mu(v_1)\, v_2\big) = \gamma(v_1)\gamma(v_2)\big(1 + v_1\, \mu(v_2)\big). \tag{11.39}$$

This yielded

$$\mu(v_1)v_2 = v_1\mu(v_2), \tag{11.40}$$

so that Lévy-Leblond could conclude

$$\mu(v) = \alpha v \tag{11.41}$$

for some proportionality constant α. Using (11.34) with $\gamma(v) = \alpha v$, he obtained

$$\gamma(v) = \frac{1}{\sqrt{1 - \alpha v^2}}, \tag{11.42}$$

where the sign of the square root was determined using (11.21a). The so-called law of addition of velocities then followed by identifying V in (11.37):

$$V = \frac{v_1 + v_2}{1 + \alpha v_1 v_2}. \tag{11.43}$$

In conclusion, depending on whether $\alpha < 0, \alpha = 0$, or $\alpha > 0$, three different cases arose:

11.3.4 Three Cases Depending on α

(i) $\alpha < 0$. By defining $\alpha = -\kappa^{-2}$, where κ had dimensions corresponding to a speed, Lévy-Leblond obtained the transformation laws

$$x' = \frac{x - vt}{\sqrt{1 + \frac{v^2}{\kappa^2}}},$$
$$t' = \frac{t + \frac{vx}{\kappa^2}}{\sqrt{1 + \frac{v^2}{\kappa^2}}}, \tag{11.44}$$

and accordingly recognized that $v \in \mathbb{R}$ was allowed. The corresponding law of velocity addition was then

$$V = \frac{v_1 + v_2}{1 - \frac{v_1 v_2}{\kappa^2}}. \tag{11.45}$$

(ii) $\alpha = 0$. This case led to the Galilean transformations

$$x' = x - vt,$$
$$t' = t. \tag{11.46}$$

and the Galilean law of velocity addition

$$V = v_1 + v_2. \tag{11.47}$$

(iii) $\alpha > 0$. The definition of $\alpha = \sigma^{-2}$, where σ was a constant whose dimensions corresponded to a speed, resulted in

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{\sigma^2}}},$$
$$t' = \frac{t - vx/\sigma^2}{\sqrt{1 - \frac{v^2}{\sigma^2}}}, \tag{11.48}$$

and

$$V = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{\sigma^2}}. \tag{11.49}$$

This was the "Lorentz case,"[15] restricting v to the interval $[-\sigma, \sigma]$.

[15] Lévy-Leblond 1976, p. 276.

11.3.5 *Causality*

In case (i) and (iii), it followed from the transformation formulas (11.44) and (11.48) respectively that the time interval between two events depended on the reference frame. Therefore, argued Lévy-Leblond, "in order to maintain *some* order in the universe, we would like to require the existence of at least a class of spatio-temporal events such that the *sign* of the time interval, that is, the nature of a possible causal relationship, is not changed under inertial transformations."[16] Stated otherwise, in cases where this condition was not met, causal explanations would not be possible. For if one assumed that an event E_1 was the cause of another event E_2, by hypothesis there would exist an inertial frame S' in which $t'(E_2) < t'(E_1)$.

Lévy-Leblond therefore investigated which of the three cases met this condition. The Galilean transformations (11.46) obviously fulfilled it for all time intervals. In the Lorentzian case (iii), it was met for those space-time intervals in which $|\Delta x/\Delta t| \leqq \sigma$, taking into account the restricted range of values for v. In case (i), however, the causality condition was not fulfilled. To recognize this, he noted that

$$\Delta t' = \frac{\Delta t + \frac{v\Delta x}{\kappa^2}}{\sqrt{1 + \frac{v^2}{\kappa^2}}} \tag{11.50}$$

for any given spatio-temporal interval $(\Delta t, \Delta x)$. It followed that one could always find a velocity $v \in \mathbb{R}$ so that $\Delta t'$ and Δt had opposite signs.

11.4 EXPERIMENTAL DETERMINATION OF THE LIMITING VELOCITY

Lévy-Leblond ended his analysis by concluding that σ was identical to the speed of light c:

> The Lorentz case is characterized by a parameter with the dimensions of a velocity, which is a universal constant associated with the very structure of space-time. A further analysis of the possible objects moving in such a space-time shows that this constant *turns out* to be the (invariant) velocity of zero-mass objects.[17]

Lévy-Leblond did not really go into how the identity between σ and c could be confirmed experimentally. To clarify this matter, it is more instructive to turn to the paper of Lee and Kalotas:

[16] Ibid., p. 276.
[17] Ibid., p. 276.

> The possibility that $\sigma^2 = \infty$ could not be ruled out on the basis of theory alone, but there is an abundance of experimental evidence which points to the fact that σ^2 is finite. Among these are experiments on the relativistic increase in mass of elementary particles with velocity (e.g., electrons)[,] the limiting electron speeds observed in linear accelerators, and the increase in mean life time of unstable high energy particles (such as mesons) in flight in accordance with the time-dilation formula which is easily derivable from the Lorentz transformation. The fact that the mass of a particle increases with its speed at all is an indication that the Galilean transformations ($\sigma^2 = \infty$) are invalid. Experiments actually show that [. . .] the limiting velocity σ is indistinguishable from the speed of light c within the present experimental limits of accuracy.[18]

To identify σ with a physical speed, one can point to empirical laws that predict a frame-independent propagation speed. Maxwell's equations do exactly that: in vacuum they imply waves of speed c. If, in accordance with the relativity principle, Maxwell's equations have the same form in all inertial frames, then the invariant speed of the kinematics must equal the wave speed of that law, i.e. $\sigma = c$. Alternatively — closer in spirit to Ignatowski and Lévy-Leblond — one can avoid invoking Maxwell: derive the one-parameter σ-family kinematically and then determine σ by measuring a high-speed phenomenon. Empirically one finds that σ is finite and numerically equal to c. The first clear evidence that fast cosmic muons — then called "mesotrons" — live longer in flight than at rest was reported by Bruno Rossi and David B. Hall in their 1941 paper, "Variation of the Rate of Decay of Mesotrons with Momentum." In 1963, the famous "mountain vs. sea level" experiment by David Frisch and James Smith showed that many more muons reach sea level than Galilean physics would allow, in quantitative agreement with Lorentz time dilation.[19]

11.5 CONCLUSION

In this chapter, we examined an alternative derivation of the Lorentz transformations. It was proposed by the Russian physicist Ignatowski and further developed by other physicists such as Lévy-Leblond. The approach

---

[18] Lee and Kalotas 1975, p. 436.
[19] Frisch and Smith 1963.

emphasized the importance of formulating special relativity without relying on principles from specific theories such as the constancy of the speed of light. Instead, the derivation of the Lorentz transformations was based on more general principles about space, time and motion common to both Newtonian and special relativistic physics. Accordingly, Ignatowski argued that clock synchronization could be achieved by using any physical method, be it mechanical or optical. This was in contrast to Poincaré and Einstein, who limited themselves to clock synchronization by means of light signaling. The demonstration of the Lorentz transformations is explained starting from Lévy-Leblond's work. It involved imposing the following general principles about space, time and motion: the homogeneity and isotropy of space, the homogeneity of time, the principle of relativity, the existence of inertial transformations as a group action on the set of inertial frames, and a causality condition. The demonstration itself consisted of two parts. First, Lévy-Leblond derived the Lorentz transformations and their degenerate Galilean limit using the general principles mentioned. Second, the Lorentz transformations were singled out experimentally by equating the limiting speed with the speed of light.

Part III

THE COMPREHENSIBILITY OF NATURE

CHAPTER 12

Comprehending Relativity

Since my historical analysis of the development of special relativity reveals central inadequacies in Friedman's interpretation of the transition from Newtonian to special relativistic physics, this third and final part of the dissertation attempts to pave the way for a better understanding of the rationality involved in this transition by means of so-called comprehensibility conditions. These are based on Helmholtz's conception of the comprehensibility of nature and were developed in recent writings of Olivier Darrigol. This chapter therefore proceeds as follows: First, I introduce Darrigol's concept of comprehensibility conditions. Second, we examine how Darrigol articulates similarities and differences between his comprehensibility conditions and Friedman's constitutive principles. Finally, I show how comprehensibility conditions illuminate important parts of my historical analysis.

12.1 HISTORICAL INADEQUACIES IN FRIEDMAN'S *DYNAMICS OF REASON*

My historical analysis of the development from Lorentz's ether-based theory of electrodynamics to Ignatowski's alternative derivation of special relativity reveals historical inadequacies in Friedman's interpretation of the transition from the Newtonian to the special relativistic conception of space-time. First of all, both Lorentz and Poincaré abandoned Newton's second and third laws of motion in their efforts to reconcile electrodynamics with the null result of ether drift experiments. Thus, they did not constitute Newtonian space-time by utilizing Newton's laws of motion. Instead, they hold on to Newtonian space-time by elevating measurements conducted at rest relative to the alleged ether as real. Lorentz's 1904 paper did not prove the invariance of Maxwell's equations under Lorentz transformations, and for that matter, his theory of electrodynamics was not empirically equivalent to Einstein's. Moreover, Lorentz had no physical interpretation of his concept of local time in terms of dilated clocks or the like. Instead, local time was a purely mathematical construct in the theory of corresponding states that allowed him to compare the theoretical predictions of an experiment conducted at rest and in uniform translational motion relative to the ether. In agreement with Friedman's interpretation, Poincaré provided a Lagrangian formulation of electrodynamics in 1905 that was empirically equivalent to Einstein's and published in two articles both entitled "Sur la Dynamique de l'Électron" (1905, 1906). However, the aforementioned Lagrangian did not lead to Newtonian mechanics, but to special relativistic mechanics. Consequently, in 1905 Poincaré was ready to reject Newtonian mechanics despite the hierarchy of science in *La Science et l'Hypothèse* (1902). The strategy of his 1905-6 papers, as the Frenchman emphasized in his book *Science et Méthode* (1908), was to reverse the logic of Lorentz's approach. He did thus not introduce Lorentz's mechanisms as primary hypotheses to prove the principle of relativity. On the contrary, by assuming that Lorentz's 1904 ether-based version of Maxwell's equations had to meet the principle of relativity, Poincaré derived how space and time coordinates had to transform under boosts of a laboratory frame from rest to uniform rectilinear motion relative to the ether. Rather than assuming Lorentz's controversial mechanisms, Poincaré emphasized that this strategy allowed him to derive them as consequences of a hypothesis that nobody had been able to challenge: the principle of relativity. In contrast to Friedman's uniqueness claim, Ignatowski encountered an alternative way of defining inertial frames that was suitable for both Newtonian and special relativistic physics.

12.2 COMPREHENSIBILITY CONDITIONS

While the article "Constitutive Principles versus Comprehensibility Conditions in Post-Kantian Physics" (2020) gives a very accessible introduction to Darrigol's concept of comprehensibility conditions, the Frenchman has formulated the most comprehensive account of this concept in his book *Physics and Necessity* (2014). The aim of the book, as Darrigol explains at the beginning, is to convey "the beauty and fertility of a moderate rationalism in which the necessity of some of our theories is derived from the contingent possibility of certain ways of understanding the physical world."[1] The paradigm of this kind of necessity, Darrigol explains, was formulated by Helmholtz in response to Kant's transcendental necessity:

> [T]he best necessity arguments all have to do with the comprehensibility of nature in a Helmholtzian sense [...] The Helmholtzian type of necessity shares the Kantian focus on the comprehensibility of nature. However, for a Helmholtzian the various conditions of comprehensibility, no matter how natural they may seem, are not to be regarded as a priori certain. Only experience can tell us to what

[1] Darrigol 2014, p. xii.

> extent these conditions are met. Thus for a Helmholtzian, when a certain theory or certain laws are argued to be necessary, it is only in the limited sense that they derive from natural but fallible preconditions for the comprehensibility of nature.[2]

Darrigol's comprehensibility conditions thus have essential features in common with Friedman's constitutive principles. On the one hand, both concepts resemble Kant's a priori categories of the understanding by agreeing that the mind imposes conditions on our experience of the world. On the other hand, both concepts deviate from Kant's categories by acknowledging that there is no uniquely determined choice of these conditions.

According to Darrigol, however, there are also important differences. One difference is how the constitutive principles are chosen in a particular period of the history of physics:

> Whereas in relativized transcendentalism this choice is either free or determined by contingent cultural conditions, comprehensibility conditions derive from the success of very broad empirical assumptions in a given domain of experience.[3]

Another difference concerns the purpose of constitutive principles:

> In relativized transcendentalism, these principles never completely determine a physical theory; they only define a frame that needs to be filled with empirical data and laws. In contrast, some comprehensibility arguments purport to completely determine a theory.[4]

Put differently, Darrigol's comprehensibility conditions do not purport to define frames of reference empirically, in constrast to Friedman's constitutive principles. Instead, they are characterized by two paradoxical qualities:

> In brief, comprehensibility principles have a theory-generating power and an empirical immediacy or naturalness that the principles of the relativized a priori do not have.[5]

Because of this conjunction of qualities, comprehensibility conditions facilitate rational inferences that "seriously improve our understanding of

[2] Ibid., pp. 340–341.
[3] Ibid., p. vii.
[4] Ibid., p. vii.
[5] Ibid., p. 345.

the structure and contents of the theory they imply."[6] For example, Darrigol continues, "we seek deeper justifications of theories we have earlier reached in a more grouping, experimentally guided manner."[7] An excellent example of this case (which we will discuss in more detail below) is Poincaré's effort to derive Lorentz's hypotheses as consequences of the relativity principle. Or Ignatowski's effort to free special relativity from its historical but ultimately contingent ties to electrodynamics. A second case is the role that comprehensibility conditions play in certain thought experiments. Darrigol himself refers to Poincaré's electromagnetic radiators from 1900, in which "the thought-experiment serves to judge the compatibility of the theory with a general principle of physics, here the equality of action and reaction."[8] As we have already seen, the reason why Poincaré valued this principle so highly was based on another thought experiment, which showed that its negation would imply the incomprehensible possibility of perpetual motion. In the remainder of this chapter, I want to highlight further where comprehensibility arguments (i.e. rational deductions including comprehensibility conditions as premises) have played a fundamental role in my historical analysis of the transition from Newtonian to special relativistic physics.

12.3 POINCARÉ

In addition to the impossibility of perpetual motion, Poincaré employed the following comprehensibility conditions on Darrigol's (incomplete) list: the principle of least action, the homogeneity and isotropy of space and the principle of relativity. The homogeneity and isotropy of space constitute an integral part of Poincaré's foundations of physical geometry. They express the natural requirement that the geometrical properties of objects (such as length, surface area and volume) remain unchanged under spatial translations and rotations. Although Poincaré agreed with Helmholtz that these conditions implied the constant curvature of space, he defended geometrical conventionalism against the German by arguing that it would be impossible to determine this curvature empirically.

When one thinks of the principle of least action, naturalness and empirical immediacy are perhaps not the first qualities that come to mind. How-

[6] Darrigol 2014, p. viii.
[7] Ibid., p. vii.
[8] Ibid., p. 361.

ever, since this principle enabled Poincaré to give mechanical explanations for physical systems without having to specify any mechanisms, Darrigol emphasizes that Poincaré regarded it as an essential part of "a global evolution of physics from naive mechanism to a 'physics of principles' in which the mathematical structure of phenomena was directly revealed, without the subterfuge of hidden mechanisms."[9] Put differently, Poincaré understood the principle of least action as a phenomenological turning point in the history of physics.

From Poincaré's point of view, the principle of relativity was based on the null results of ether drift experiments. But it went beyond these results by requiring the impossibility of detecting uniform rectilinear motion relative to the ether for any future experiment. Unlike Lorentz's mechanisms, such as the contraction of the electron, this general principle was easy to comprehend and difficult to dismiss because of the many failed attempts to falsify it. Poincaré's strategy was therefore to reverse the logic of Lorentz's approach. More precisely, he used the relativity principle as one of the premises in a comprehensibility argument to derive Lorentz's more controversial mechanisms (besides the Lorentz transformations) as consequences.

Although Poincaré's argument was far superior to Lorentz's, Darrigol reveals why it had its own weaknesses due to the circumstance that it employed the validity of Lorentz's ether-based theory of electrodynamics as one of its other premises:

> A first weakness of this version of the theory of relativity is that it makes the Lorentz transformations depend on the electromagnetic nature of all physical processes [...] Another weakness is that despite Poincarés early criticism of the ether, his approach still maintains the ether as the carrier of electromagnetic energy and as the reference frame in which true space and time are measured.[10]

While Darrigol acknowledges that "[t]he choice of an ether frame is an arbitrary convention that enables us to preserve the traditional concepts of space and time as those defined in this frame,"[11] this choice implies an unnecessary entanglement between the Lorentz transformations and electromagnetism:

---

[9] Ibid., p. 79.
[10] Ibid., p. 138.
[11] Ibid., p. 138.

> Although this is a logically possible option, the surviving ether makes it more difficult to disconnect the Lorentz transformations from electromagnetism and to perceive their kinematic necessity.[12]

A decisive step in this direction was taken by Einstein.

12.4 EINSTEIN

Einstein avoided any reference to the ether by simply assuming the existence of a reference frame in which light always propagates at a definite speed *c*, regardless of the state of motion of the emitting body. He then introduced the principle of relativity stating that the "laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to one or the other of two systems of coordinates in uniform translatory motion."[13] Darrigol notes that while the first principle "is a trivial consequence of the ether-based theory of light, as long as the given system of coordinates refers to the ether,"[14] the relativity principle was well known to every student of mechanics. Nevertheless, together with the natural assumptions of the homogeneity of time and the homogeneity and isotropy of space, these well-established principles enabled Einstein to derive revolutionary concepts of space and time. As we have seen, Poincaré avoided such conclusions by attributing effects such as the Lorentz contraction to the dynamical impact of the ether on matter. Accordingly, the Lorentz transformations expressed the relationships between true space and time measures in the ether frame and their apparent measures in a frame in uniform rectilinear motion relative to the ether. In Einstein's new kinematics, on the other hand, Poincaré's distinction was rejected by acknowledging that measures – such as length, time and mass – were frame-dependent. But even though Einstein's derivation and interpretation of the Lorentz transformations differed significantly from Poincaré's, a similarity remained. In contrast to the Frenchman, Einstein did not base the Lorentz transformations on the validity of Maxwell's equations. But since the light principle served as a premise in his derivation, the transformations retained their special ties to electromagnetism.

12 Darrigol 2014, p. 138.
13 Einstein 1952, p. 41.
14 Darrigol 2014, p. 138.

12.5 IGNATOWSKI

The first attempt to derive the necessity of special relativity without theory-specific presuppositions was undertaken by Ignatowski in 1910. Above all, Ignatowski had to find a way to define inertial frames without reference to the light principle. Since he chose to structure his derivation by imposing certain requirements in the process, it easily escapes one's attention that he ended up defining inertial frames by requiring the homogeneity and isotropy of space together with the homogeneity of time. It turns out, for example, that the period of a Foucault pendulum depends on the latitude of its location. Consequently, a frame adapted to earth is not inertial. Ignatowski went on to introduce so-called inertial transformations as the transformations that related any two such inertial frames. The relativity principle then required that the class of inertial transformations constituted a group action on the set of inertial frames. In conjunction with the other conditions, it followed that the group of transformations depended on only one universal constant. By adding a causality condition to the list of assumptions, this constant could be interpreted as the limiting velocity of zero-mass objects. In the case of a finite limit, the derivation implied that the group of transformations corresponded to the Lorentz group. In the infinite case, the group was identified with the Galilean group. Which case applied was finally a question of empirical investigation.

This brief overview shows that comprehensibility arguments played a significant role in the theorizing of Poincaré, Einstein and Ignatowski. In addition to apparent differences between their strategies, the summary shows that all three used homogeneity and isotropy requirements in their derivations. Most surprisingly, Ignatowski chose to define inertial frames by means of these requirements. While this possibility is problematic for Friedman's interpretation of the necessity of constitutive principles in terms of their uniqueness, it also suggests a necessary connection between homogeneity and isotropy requirements on the one hand and mechanics on the other.

12.6 LANDAU AND LIFSHITZ

The most prominent investigation into this connection, which we shall see in the next chapter, was carried out by the Russian physicists Lev Landau and Evgeny Lifshitz as part of their seminal course on the foundations of theoretical physics. Although physicists acknowledge this series of textbooks as one of the most important and comprehensive foundational studies in the history of their subject, it has gone largely unnoticed among philosophers of physics. Two parts of their study are of particular interest to us: First, by applying the principle of least action in conjunction with Ignatowski's definition of inertial frames, they proved the conservation of linear and angular momentum. Second, by requiring that interactions propagate at either infinite or finite speed, they derived the remainder of Newtonian mechanics for the infinite case and the remainder of special relativistic mechanics for the finite case. The second derivation is important for us because it corresponds to a comprehensibility or necessity argument for Newtonian and special relativistic mechanics respectively. In particular, the proof explains how different sets of principles can define the same group of inertial frames. The first derivation, on the other hand, is fundamental because it allows us to formulate mathematical blueprints for the construction and synchronization of ideal clocks, which are shared across Newtonian physics, special relativity and ether-based electrodynamics.

12.7 INTERPRETIVE SCHEMES

Before we continue, let me explain how Darrigol understands these blueprints and why they are an indispensable extension of Friedman's account of the application of physical theories to experience. Darrigol calls such mathematical blueprints for interpretive schemes:

> A first central feature [of the definition of physical theories] is the inclusion of mathematical blueprints of experimental devices, which I call interpretive schemes.[15]

He introduces them in connection with his definition of physical theories:

> The theories of modern physics share the following characteristics:
>
> (a) The definition of a symbolic universe in which systems, states, transformations, and evolutions are defined by means of various magnitudes based on powers of $\mathbb{R}$ (or $\mathbb{C}$) and on derived functional spaces and algebras.
>
> (b) The postulation of theoretical laws that restrict the behavior of systems in the symbolic universe.

[15] Darrigol 2014, p. 340.

(c) The description of interpretive schemes that relate the symbolic universe to idealized experiments.

(d) Methods of approximation that enable us to derive the consequences that the theoretical laws have on the interpretive schemes.[16]

After a few examples, Darrigol finally contents himself with the following characterization of interpretive schemes:

> [A]n interpretive scheme consists of a given system of the symbolic universe together with a list of characteristic quantities that satisfy the following three properties:
>
> 1. They are selected among or derived from the (symbolic) quantities that define the state of this system.
> 2. At least for some of them, ideal measuring procedures are known.
> 3. The laws of the symbolic universe imply relations of a functional or statistical nature among them.[17]

Apart from the inclusion of methods of approximation, this definition of physical theories and their application to experience does not seem to differ substantially from Friedman's. However, such an assessment neglects that Darrigol's conception of interpretive schemes does not coincide with Friedman's constitutive principles.

One way to illustrate the difference is to look at Newtonian mechanics from both perspectives. According to Friedman's interpretation, Newton's laws of motion are a priori constitutive principles and not empirical laws in the proper sense. According to Darrigol, on the other hand, Newtonian mechanics does fit the definition of a physical theory, whereby the laws of motion function as theoretical laws. Thus, even though Newton's laws of motion may suffice to define inertial frames theoretically, we still need both interpretive schemes and non-theoretical knowledge for their empirical implementation:

> The symbolic universe of a theory never applies directly to a concrete situation. The application is mediated through interpretive schemes

[16] Ibid., p. 347.
[17] Ibid., p. 350.

> that describe ideal devices [including ideal measurement procedures] and quantitative properties of these devices. In order to build a concrete counterpart of a scheme, we must know the correspondence between ideal device and real device, as well as concrete operations that yield the measured quantities.[18]

Some of the examples of interpretive schemes that we have looked at in detail are the design of the Michelson-Morley experiment as well as Poincaré's and Einstein's methods for synchronizing clocks by light signals. The latter exemplify so-called indirect measurements that imply the laws of the theory itself.[19] In general, measuring devices of a theory can be based on mathematical blueprints that imply laws of other theories. In such cases, Darrigol calls the latter theories modules of the former. For example, time measurement by so-called light clocks would imply a modular structure from Ignatowski's perspective on special relativity. Another example would be the geometric module of mechanics. In the latter case, geometry is both a defining module and a schematic module. The former because geometric quantities, such as the dimensions of objects, are part of the symbolic universe of the theory. The latter is the case because the interpretive schemes of mechanics include geometric measurement procedures. In the case of length measurement in astronomy using optical methods, optics would be a schematic module, but not (necessarily) a defining module.

In a passage directed against Friedman's constitutive principles, Darrigol formulates the interplay between the interpretive schemes and the modular structure of a theory as follows:

> The interpretive schemes, not any law or principle, are responsible for the coordination between formalism and concrete experiments. Unlike coordinating principles, the interpretive schemes are not an invariable component of the theory: they follow the history of its applications. The symbolic universe and its laws control their form without dictating it. Their definition and their deployment requires a modular structure which itself evolves in the history of applications of the theory.[20]

For example, even if we define inertial frames by means of Newton's laws

[18] Darrigol 2014, p. 354.
[19] Ibid., p. 355.
[20] Darrigol 2020, pp. 4610–4611.

of motion (or homogeneity and isotropy requirements), we still need an interpretive scheme or blueprint, such as Foucault's pendulum, to determine whether a particular laboratory frame is inertial or not. As mentioned earlier, the concrete realization of such schemes may then require components or devices that draw on independent sub-theories as well as other theories. Furthermore, and most importantly, the schemes and modular structure of a theory may evolve in complex and unpredictable ways over the course of history. One of the best examples of such complexity is exactly the history of electrodynamics and its development from Lorentz and Poincaré to Einstein and beyond.

So while comprehensibility conditions serve as necessary premises in the derivation of a theory, they do not tell us how to apply the constructed theory to experience. The latter depends on the empirical realization of interpretive schemes through the modular structure of the theory. Even if the application of a theory presupposes an interpretive scheme, the latter is "variable and partly contingent" and consequently does not define a "constitutive a priori component of a theory."[21]

12.8 THE TRANSITION BETWEEN CONSTITUTIVE FRAMEWORKS

After reconsidering theory construction and the application of theories through Darrigol's concepts of comprehensibility conditions and interpretive schemes, we finally need to account for the transition between successive theories that require a change of constitutive framework:

> Friedman [appeals] to a philosophical meta-framework guiding the transition toward new constitutive principles. [...] In the view adopted in this essay, the modular structure of physical theory brings much rational continuity in the evolution of physics. But it does not tell us how to change the comprehensibility conditions. Here is a hint: these conditions being known for a received theory, try to relax them to get a more general and more unified theory.[22]

In contrast to Friedman, rather than involving a philosophical meta-framework, such transitions generally proceed by easing (at least) one of the comprehensibility conditions of the received theory. In order to construct the successor theory, we then need to reconsider all derivations of

[21] Ibid., p. 4611.
[22] Ibid., p. 4613.

the received theory that rely on the unrelaxed condition. As a result, we may also need to adapt the interpretive schemes to the new situation. Poincaré's derivation of the Lorentz transformations by applying the relativity principle to Lorentz's ether-based electrodynamics, for example, implied the replacement of Newtonian mechanics. In his lectures at the Sorbonne, he then demonstrated that his synchronization method by light signaling was in compliance with the new transformations. Ignatowski, on the other hand, recognized that it was possible to define inertial frames by means of homogeneity and isotropy conditions. He therefore emphasized that his 1910 derivation of the Lorentz transformations did not presuppose any theory-specific principle such as the light principle. Although he did not substantiate his claim with detailed examples, he also stressed that it would be easy to construct multiple schemes for the synchronization of clocks.

My comparison between Poincaré and Ignatowski may give the impression that "the freedom in defining the schemes and the tacit knowledge involved in their concrete realization [...] leave plenty of room for protecting theories from refutation."[23] Darrigol claims, however, that "Duhemian holism, or unrestricted open-endedness, do not occur in the actual practice of physics,"[24] because "[t]he modular structure of theories gives them much more rigidity in their adaption to the empirical world than some historians would have it."[25] In particular, he argues that "the modular structure severely limits the protective strategies because it restricts the form of the schemes and because it tends to confine tacit knowledge in the application of well-understood modules."[26] For example, I may add, even though Poincaré thought he could protect Newtonian space-time through an ether-based distinction between real and apparent measures, the Frenchman was still aware that what he called "the new mechanics" necessitated a new synchronizing method for clocks and the reshaping of all of physics in Lorentz-invariant form.

12.9 CONCLUSION

Based on my historical analysis of the development from Lorentz's ether-based theory of electrodynamics to Ignatowski's alternative derivation of

[23] Darrigol 2020, p. 4595.
[24] Ibid., p. 4596.
[25] Ibid., p. 4596.
[26] Ibid., p. 4595.

the Lorentz transformations, I have reviewed historical inadequacies in Friedman's interpretation of the transition from Newtonian to special relativistic physics. Given these inadequacies, I have proposed to explain the development of mathematical physics by means of Darrigol's concepts of comprehensibility conditions and interpretive schemes. In contrast to Friedman's a priori constitutive principles, comprehensibility conditions do not purport to define inertial frames empirically. Instead, they possess both theory-generating power and empirical immediacy, enabling rational deductions of physical theories that improve our understanding of their structure and content. Interpretive schemes are mathematical blueprints for idealized experiments that mediate the application of a theory to concrete experience. For example, both Poincaré and Einstein comprehended electrodynamical phenomena theoretically by presupposing the principle of relativity, and they utilized light signaling as a scheme for synchronizing clocks. Finally, we have explained how comprehensibility conditions and interpretive schemes facilitate the transition between successive theories.

CHAPTER 13

Space-Time Symmetries and Conservation Laws

In chapter 11, we examined how Ignatowski's approach to relativity revealed a number of comprehensibility conditions, such as the homogeneity and isotropy of space-time, that inertial frames of both the Newtonian and special relativistic framework had to satisfy. However, the equivalence between this very general definition of inertial frames and its two preceding counterparts – one based on Newton's laws of motion and the other on Einstein's light principle – remained somewhat unclear. In order to close this gap, we will therefore refer to the work of Lev Davidovitch Landau and Evgeny Lifshitz as well as Wolfgang Rindler. First, we will see how to obtain both the law of inertia and the principle of relativity from the homogeneity and isotropy of space-time. Secondly, we derive the conservation of linear momentum and angular momentum. Finally, by specifying the speed of interactions, we derive the divergent parts of Newtonian and special relativistic mechanics. These results clarify the relationship between space-time symmetries and conservation laws in mechanics and accordingly pave the way for the discussion of Friedman's concept of constitutive a priori principles in the final chapter.

13.1 THE LAW OF INERTIA

As we have seen, Ignatowski's starting point was the existence of a group of reference frames that, among other things, fulfilled the requirements of homogeneity and isotropy. Landau and Lifshitz take the same approach in the first volume of their pioneering series of textbooks on theoretical physics:

> In order to consider mechanical phenomena it is necessary to choose a *frame of reference*. The laws of motion are in general different in form for different frames of reference. [...] If we were to choose an arbitrary frame of reference, space would be inhomogeneous and anisotropic. This means that, even if a body interacted with no other bodies, its various positions and its different orientations would not be mechanically equivalent. [...] It is found, however, that a frame of reference can always be chosen in which space is homogeneous and isotropic and time is homogenous. This is called an *inertial frame*.[1]

[1] Landau and Lifshitz 1969, pp. 4–5.

By combining this concept of an inertial frame with the principle of least action, Landau and Lifshitz derive the law of inertia as a direct consequence. They first establish that the homogeneity of space and time implies that the Lagrangian for a free particle cannot depend explicitly on either the position vector $\mathbf{r}$ or the time t. It follows that the Lagrangian L must be a function of the velocity $\mathbf{v}$ alone. Since space is also isotropic, the Lagrangian cannot depend on the direction of $\mathbf{v}$, but only on its magnitude $\mathbf{v}^2 = v^2$:

$$L = L(v^2).$$

Due to the independence of L from $\mathbf{r}$, they now obtain $\frac{\partial L}{\partial \mathbf{r}} = 0$, and accordingly Lagrange's equations in vector notation reduce to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \mathbf{v}}\right) = \mathbf{0}. \tag{13.1}$$

This equation in turn implies that $\frac{\partial L}{\partial \mathbf{v}}$ must be a constant vector. However, since $\frac{\partial L}{\partial \mathbf{v}}$ is solely a function of $\mathbf{v}$, it follows that the velocity $\mathbf{v}$ of a free particle must be a constant vector. Landau and Lifshitz therefore "conclude that, in an inertial frame, any free motion takes place with a velocity which is constant in both magnitude and direction."[2] Stated otherwise, they derived the law of inertia as a consequence of the homogeneity of space-time and the isotropy of space, in accordance with the Newtonian and the special relativistic concept of an inertial frame.

13.2 THE PRINCIPLE OF RELATIVITY

The principle of relativity is traditionally introduced after the class of inertial frames, i.e. the assumption that any frame that moves uniformly and rectilinearly with respect to a given inertial frame is itself inertial is followed by the additional postulate that the laws of physics are the same in all inertial frames. However, since we have defined inertial frames as spatially and temporally homogeneous as well as spatially isotropic, rather than assuming the principle of relativity as an additional requirement, Rindler has argued that it is possible to understand it as already implicit in the homogeneity and isotropy requirements for inertial frames.

His proof depends on the "lemma that 'between' any two inertial frames S and S' there exists an inertial frame S'' relative to which S and S' have equal and opposite velocities."[3] For if the members of this class are ordered

[2] Ibid., p. 5.

[3] Rindler 1982, p. 8.

according to their velocity relative to S, it follows from continuity that one of the frames must have the required property. Rindler now imagines two identical experiments E and E', which are carried out in S and S' respectively. It is then possible to "transform E', by a spatial translation and rotation and a temporal translation, in S', into a position where it differs from E only by a 180[-degree] rotation in S''."[4] For reference, let us call the transformed experiment for E''. Due to the homogeneity and isotropy of S', E' and E'' must give identical results in S'. Moreover, E and E'' differ only by a rotation of 180 degrees in S''. By assuming that E moves in a certain direction with speed v in S'', we can conclude from the choice of S'' that E'' must move with the same speed but in the opposite direction. The homogeneity and isotropy of S'' now imply that E and E'' must give identical results in S''. It follows that the result of E in S is identical to the result of E' in S', which proves the principle of relativity.

Now that Rindler has clarified the relationship between the principle of relativity and the homogeneity and isotropy of inertial frames, we return to Landau and Lifshitz's account of mechanics based on the required space-time symmetries.

13.3 THE CONSERVATION OF LINEAR MOMENTUM

In this section, we will see how to obtain the conservation of linear momentum as a consequence of the homogeneity of space.[5] For this purpose, we consider a closed system consisting of N particles positioned at $\mathbf{r}_\alpha$ ($\alpha = 1, \ldots, N$). It follows from the homogeneity of space that the physical properties of the system remain unchanged for any parallel displacement of the entire system. In particular, the displacement of each particle by the same infinitesimal virtual translation $\boldsymbol{\epsilon}$ will not change the Lagrangian L of the system:

$$\delta L = 0.$$

According to the rules of the calculus of variations, the change in L due to the same infinitesimal virtual translation $\delta\mathbf{r}_\alpha = \boldsymbol{\epsilon}$ is also given by

$$\delta L = \sum_\alpha \frac{\partial L}{\partial \mathbf{r}_\alpha} \cdot \delta\mathbf{r}_\alpha = \sum_\alpha \frac{\partial L}{\partial \mathbf{r}_\alpha} \cdot \boldsymbol{\epsilon} = \boldsymbol{\epsilon} \cdot \sum_\alpha \frac{\partial L}{\partial \mathbf{r}_\alpha}.$$

[4] Rindler 1982, p. 8.
[5] Landau and Lifshitz 1969, pp. 15–6.

The combination of the two equations for δL then results in

$$\boldsymbol{\epsilon} \cdot \sum_\alpha \frac{\partial L}{\partial \mathbf{r}_\alpha} = 0.$$

Since $\boldsymbol{\epsilon}$ can be chosen arbitrarily, this is only possible if

$$\sum_\alpha \frac{\partial L}{\partial \mathbf{r}_\alpha} = \mathbf{0}. \tag{13.2}$$

If we apply the summation rule for differentiation and substitute the requirement (13.2) into the Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \mathbf{v}_\alpha} = \frac{\partial L}{\partial \mathbf{r}_\alpha},$$

we get

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_\alpha \frac{\partial L}{\partial \mathbf{v}_\alpha} = \sum_\alpha \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \mathbf{v}_\alpha} = \sum_\alpha \frac{\partial L}{\partial \mathbf{r}_\alpha} = \mathbf{0}. \tag{13.3}$$

Recalling Hamilton's standard formula for the general momentum

$$\mathbf{p}_\alpha = \frac{\partial L}{\partial \mathbf{v}_\alpha},$$

we can finally interpret (13.3) to mean that the total linear momentum

$$\mathbf{P}_{\text{tot}} = \sum_\alpha \mathbf{p}_\alpha$$

does not change with respect to time:

$$\dot{\mathbf{P}}_{\text{tot}} = \frac{\mathrm{d}}{\mathrm{d}t} \sum_\alpha \mathbf{p}_\alpha = \mathbf{0}.$$

Put differently, the total linear momentum of the system is conserved.

13.4 THE CONSERVATION OF ANGULAR MOMENTUM

We now proceed to demonstrate the conservation of angular momentum as a consequence of the isotropy of space.[6] Due to isotropy, the physical properties of a closed system remain unchanged for every rotation of the

[6] Ibid., pp. 18–9.

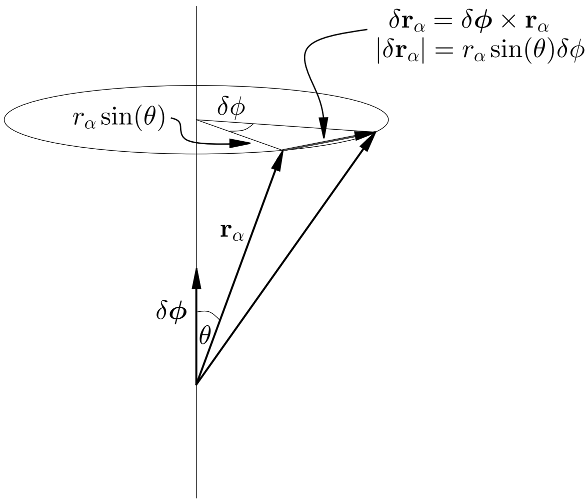


Fig. 13.1: The infinitesimal change in $\mathbf{r}_\alpha$ due to an infinitesimal rotation by $\delta\phi$.

entire system. In particular, a rotation of the system by an infinitesimal virtual angle $\delta\phi$ around some axis will not affect the Lagrangian:

$$\delta L = 0. \tag{13.4}$$

Due to the rotation, each particle in the system is displaced by a distance $\delta\mathbf{r}_\alpha$ and the direction of its velocity change by $\delta\mathbf{v}_\alpha$. It follows that the corresponding change in L can be calculated by

$$\delta L = \sum_\alpha \left(\frac{\partial L}{\partial \mathbf{r}_\alpha} \cdot \delta\mathbf{r}_\alpha + \frac{\partial L}{\partial \mathbf{v}_\alpha} \cdot \delta\mathbf{v}_\alpha \right).$$

Using Lagrange's equations, we can replace the derivative $\frac{\partial L}{\partial \mathbf{r}_\alpha}$ with $\dot{\mathbf{p}}_\alpha$:

$$\delta L = \sum_\alpha (\dot{\mathbf{p}}_\alpha \cdot \delta\mathbf{r}_\alpha + \mathbf{p}_\alpha \cdot \delta\mathbf{v}_\alpha).$$

Let us now use $\delta\boldsymbol{\phi}$ to denote the vector that is directed along the axis of rotation and whose magnitude corresponds to the angle of rotation $\delta\phi$. From figure 13.1, we obtain $|\delta\mathbf{r}_\alpha| = |\delta\boldsymbol{\phi} \times \mathbf{r}_\alpha|$. Since $\delta\mathbf{r}_\alpha$ is also perpendicular to both $\mathbf{r}_\alpha$ and $\delta\boldsymbol{\phi}$, we get $\delta\mathbf{r}_\alpha = \delta\boldsymbol{\phi} \times \mathbf{r}_\alpha$. A similar argument results in

$\delta \mathbf{v}_\alpha = \delta\boldsymbol{\phi} \times \mathbf{v}_\alpha$. Therefore,

$$\delta L = \sum_\alpha (\dot{\mathbf{p}}_\alpha \cdot \delta\boldsymbol{\phi} \times \mathbf{r}_\alpha + \mathbf{p}_\alpha \cdot \delta\boldsymbol{\phi} \times \mathbf{v}_\alpha). \tag{13.5}$$

If we equate (13.4) with (13.5) and interchange the factors, we get

$$0 = \sum_\alpha \delta\boldsymbol{\phi} \cdot (\mathbf{r}_\alpha \times \dot{\mathbf{p}}_\alpha + \dot{\mathbf{r}}_\alpha \times \mathbf{p}_\alpha) = \delta\boldsymbol{\phi} \cdot \sum_\alpha \frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{r}_\alpha \times \mathbf{p}_\alpha) = \delta\boldsymbol{\phi} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \sum_\alpha \mathbf{r}_\alpha \times \mathbf{p}_\alpha.$$

Since $\delta\boldsymbol{\phi}$ can be chosen arbitrarily, it follows that

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} \sum_\alpha \mathbf{r}_\alpha \times \mathbf{p}_\alpha = \frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t},$$

where

$$\mathbf{L} = \sum_\alpha \mathbf{r}_\alpha \times \mathbf{p}_\alpha$$

is called the total angular momentum of the system. This means that the conservation of total angular momentum is a consequence of the isotropy of space.

13.5 THE NEWTONIAN FRAMEWORK

Since all derivations have been based on common space-time symmetries, so far all results in this chapter apply to both the Newtonian and the special relativistic framework. In the remaining part of this chapter, following Landau and Lifshitz, we will see how to obtain separate formulas for the Lagrangian for a free particle and for a system of interacting particles in both the Newtonian and the special relativistic frameworks.

13.5.1 *The Lagrangian for a Free Particle*

We begin by determining the Newtonian Lagrangian for a free particle relative to an inertial frame.[7] In particular, let S' be an inertial frame moving with a uniform infinitesimal velocity $\boldsymbol{\epsilon}$ with respect to another inertial frame S. It follows from the Galilean law of velocity addition that a free particle moving with velocity $\mathbf{v}'$ in S' moves with velocity $\mathbf{v} = \mathbf{v}' + \boldsymbol{\epsilon}$ in S. Furthermore, in the Galilean case we have $t = t'$ and correspondingly

[7] Landau and Lifshitz 1969, pp. 6–7.

$\mathrm{d}t = \mathrm{d}t'$. Recalling that the Lagrangian for a free particle can only depend on the square of the velocity,

$$L = f(\mathbf{v}^2) = f(v^2), \tag{13.6}$$

we then apply the Galilean law of velocity addition and the condition of absolute time to the principle of least action relative to S:

$$\begin{aligned}\Gamma = \int_{t_1}^{t_2} L\,\mathrm{d}t &= \int_{t_1}^{t_2} f(v^2)\,\mathrm{d}t \\ &= \int_{t_1'}^{t_2'} f((\mathbf{v}' + \boldsymbol{\epsilon})^2)\,\mathrm{d}t' \\ &= \int_{t_1'}^{t_2'} f(v'^2 + 2\mathbf{v}'\cdot\boldsymbol{\epsilon} + \epsilon^2)\,\mathrm{d}t'.\end{aligned}$$

Taking $x = v'^2$ and $\delta = 2\mathbf{v}'\cdot\boldsymbol{\epsilon} + \epsilon^2$, we can expand $f(v'^2 + 2\mathbf{v}'\cdot\boldsymbol{\epsilon} + \epsilon^2)$ to first order in ϵ:

$$\begin{aligned}f(v'^2 + 2\mathbf{v}'\cdot\boldsymbol{\epsilon} + \epsilon^2) &= f(v'^2) + \frac{\partial f}{\partial v'^2}\left(2\mathbf{v}'\cdot\boldsymbol{\epsilon} + \epsilon^2\right) + O(\epsilon^2) \\ &= f(v'^2) + \frac{\partial f}{\partial v'^2}2\mathbf{v}'\cdot\boldsymbol{\epsilon} + \tilde{O}(\epsilon^2) \\ &\approx f(v'^2) + \frac{\partial f}{\partial v'^2}2\mathbf{v}'\cdot\boldsymbol{\epsilon},\end{aligned}$$

where $\tilde{O}(\epsilon^2)$ denotes all second-order terms in ϵ:

$$\tilde{O}(\epsilon^2) = O(\epsilon^2) + \frac{\partial f}{\partial v'^2}\epsilon^2.$$

Inserting the Taylor expansion to first order into the principle of least action,

$$\Gamma = \int_{t_1'}^{t_2'} f(v'^2) + \frac{\partial f}{\partial v'^2}2\mathbf{v}'\cdot\boldsymbol{\epsilon}\,\mathrm{d}t',$$

we notice that the integrand must correspond to the Lagrangian in S':

$$L'(\mathbf{v}') = f(v'^2) + \frac{\partial f}{\partial v'^2}2\mathbf{v}'\cdot\boldsymbol{\epsilon}. \tag{13.7}$$

Next, we can add a total time derivative of some function of coordinates and time to the Lagrangian without changing the equations of motion of the particle due to the so-called Gauge invariance of the principle of least action. To see this, let $\tilde{L}(\mathbf{r}, \mathbf{v}, t)$ and $L(\mathbf{r}, \mathbf{v}, t)$ be two functions that differ in the total time derivative of some function $G(\mathbf{r}, t)$:

$$\tilde{L}(\mathbf{r}, \mathbf{v}, t) = L(\mathbf{r}, \mathbf{v}, t) + \frac{\mathrm{d}}{\mathrm{d}t} G(\mathbf{r}, t).$$

The integration results in

$$\begin{aligned}\tilde{\Gamma} &= \int_{t_1}^{t_2} \tilde{L}(\mathbf{r}, \mathbf{v}, t)\, \mathrm{d}t \\ &= \int_{t_1}^{t_2} L(\mathbf{r}, \mathbf{v}, t)\, \mathrm{d}t + \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} G(\mathbf{r}, t) \mathrm{d}t \\ &= \Gamma + G(\mathrm{r}_2, t_2) - G(\mathrm{r}_1, t_1).\end{aligned}$$

Therefore, $\tilde{\Gamma}$ and Γ only differ by a quantity that does not contribute to the variation. This means that $\delta\tilde{\Gamma} = 0$ and $\delta\Gamma = 0$ lead to the same equations of motion as required.

If we combine the Gauge invariance of the principle of least action with the principle of relativity, it follows that the Lagrangian L' in S' must have the same form as the Lagrangian L in S up to a total time derivative of a function $G(\mathbf{r}', t')$:

$$L'(\mathbf{r}', \mathbf{v}', t') = f(v'^2) + \frac{\mathrm{d}}{\mathrm{d}t'} G(\mathbf{r}', t'). \tag{13.8}$$

If we now equate our two expressions for L' (13.7) and (13.8), we derive the condition

$$\frac{\mathrm{d}}{\mathrm{d}t'} G(\mathbf{r}', t') = \frac{\partial f}{\partial v'^2} 2\mathbf{v}' \cdot \boldsymbol{\epsilon}. \tag{13.9}$$

According to the chain rule, we can calculate the differential of $G(\mathbf{r}', t')$ as

$$\mathrm{d}G = \frac{\partial G}{\partial \mathbf{r}'} \cdot \mathrm{d}\mathbf{r}' + \frac{\partial G}{\partial t'}\, \mathrm{d}t' = \sum_{i=1}^{3} \frac{\partial G}{\partial r_i'}\, \mathrm{d}r_i' + \frac{\partial G}{\partial t'}\, \mathrm{d}t'. \tag{13.10}$$

By rewriting the right-hand side of (13.9) as well,

$$\mathrm{d}G = \frac{\partial f}{\partial v'^2} 2\mathbf{v}' \cdot \boldsymbol{\epsilon}\, \mathrm{d}t' = \frac{\partial f}{\partial v'^2} 2\mathrm{d}\mathbf{r}' \cdot \boldsymbol{\epsilon} = \sum_{i=1}^{3} \frac{\partial f}{\partial v'^2} 2\epsilon_i\, \mathrm{d}r_i', \tag{13.11}$$

we obtain the following equations by matching corresponding terms in (13.10) and (13.11):

$$\frac{\partial G}{\partial t'} = 0,$$
$$\frac{\partial G}{\partial r'_i} = 2\epsilon_i \frac{\partial f}{\partial v'^2}, \quad i \in \{1,2,3\}.$$

Put differently, G and its partial derivatives are only functions of the coordinates r'_i. To fulfill the condition (13.9), $\frac{\partial f}{\partial v'^2}$ can therefore only be a function of the coordinates r'_i. Since f is a function of v'^2, it follows that $\frac{\partial f}{\partial v'^2}$ must be a constant a. With hindsight we introduce $a = \frac{1}{2}m$ and call m the *mass* of the particle:

$$\frac{\partial f}{\partial v'^2} = a = \tfrac{1}{2}m. \tag{13.12}$$

If we integrate (13.12), we get

$$f(v'^2) = \tfrac{1}{2}mv'^2$$

by setting the integration constant to 0. This is legitimate because adding a constant to the Lagrangian does not change the equations of motion. Inserting this result into (13.6) finally gives the Lagrangian for a free particle in the Newtonian framework:

$$L(v) = \tfrac{1}{2}mv^2.$$

If we differentiate this formula with respect to $\mathbf{v}$, by virtue of the Lagrangian formula for the generalized momentum

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}},$$

we arrive at the classical formula for linear momentum:

$$\mathbf{p} = \frac{\partial}{\partial \mathbf{v}}\left(\tfrac{1}{2}m\mathbf{v}\cdot\mathbf{v}\right) = m\mathbf{v}.$$

13.5.2 The Lagrangian for a System of Particles

We now consider a closed system of N interacting particles. To represent this system, we add the Lagrangians for N non-interacting particles and subtract a function U of the coordinates describing the interaction:

$$L = \sum_\alpha \tfrac{1}{2}m_\alpha v_\alpha^2 - U(\mathbf{r}_1, \ldots, \mathbf{r}_N), \tag{13.13}$$

so that

$$\Gamma = \int_{t_1}^{t_2} L\,\mathrm{d}t = \int_{t_1}^{t_2} \sum_{\alpha} \tfrac{1}{2} m_\alpha v_\alpha^2 - U(\mathbf{r}_1, \ldots, \mathbf{r}_N)\,\mathrm{d}t.$$

The sum

$$T = \sum_{\alpha} \tfrac{1}{2} m_\alpha v_\alpha^2$$

corresponds to the total kinetic energy and U to the potential energy of the system. The circumstance that U depends only on the position of the particles at a given time, Landau and Lifshitz explain, is based on the premise that "a change in the position of any particle instantaneously affects all the other particles."[8] This in turn results from the conjunction of the relativity principle and absolute time:

> If the propagation of interactions were not instantaneous, but took place with a finite velocity, then that velocity would be different in different frames of reference in relative motion, since the absoluteness of time necessarily implies that the ordinary law of composition of velocities is applicable to all phenomena. The laws of motion for interacting bodies would then be different in different inertial frames, a result which would contradict the relativity principle.[9]

Stated differently, the relativity principle requires that interactions propagate according to Ignatowski's invariant limiting velocity in order for the relativity principle to be valid. Thus, interactions propagate instantaneously in Newtonian mechanics. In the special relativistic case, however, interactions must propagate at the speed of light.
Substituting the Lagrangian (13.13) into the equations of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \mathbf{v}_\alpha} = \frac{\partial L}{\partial \mathbf{r}_\alpha}, \tag{13.14}$$

results in

$$m_\alpha \frac{\mathrm{d}\mathbf{v}_\alpha}{\mathrm{d}t} = -\frac{\partial U}{\partial \mathbf{r}_\alpha}. \tag{13.15}$$

Introducing the force $\mathbf{F}_\alpha$ on the α-th particle

$$\mathbf{F}_\alpha = \frac{\partial L}{\partial \mathbf{r}_\alpha} = -\frac{\partial U}{\partial \mathbf{r}_\alpha},$$

[8] Landau and Lifshitz 1969, p. 8.
[9] Ibid., p. 8.

we then recognize formula (13.15) as Newton's second law of motion:

$$\mathbf{F}_\alpha = m_\alpha \ddot{\mathbf{x}}_\alpha.$$

To obtain Newton's third law, we rewrite the force in terms of the momentum:

$$\mathbf{F}_\alpha = \frac{\mathrm{d}(m_\alpha \mathbf{v}_\alpha)}{\mathrm{d}t} = \dot{\mathbf{p}}_\alpha.$$

This means that the force measures the change in linear momentum with respect to time. The conservation of total linear momentum for a closed system therefore implies that the total force sums to **0**:

$$\mathbf{F}_{\text{tot}} = \sum_\alpha \mathbf{F}_\alpha = \sum_\alpha \dot{\mathbf{p}}_\alpha = \dot{\mathbf{P}}_{\text{tot}} = \mathbf{0}.$$

In particular, if we consider a system with only two particles, we get

$$\mathbf{F}_1 + \mathbf{F}_2 = \mathbf{0},$$

which implies

$$\mathbf{F}_1 = -\mathbf{F}_2.$$

Stated otherwise, the force that the first particle exerts on the second is equal in magnitude, but opposite in direction, to the force that the second particle exerts on the first. This is Newton's law of reaction for a closed system of two particles. The general case then results from the principle of superposition.

13.6 THE SPECIAL RELATIVISTIC FRAMEWORK

We now turn to Ignatowski's second case of a finite limiting speed. In the following, we shall restrict ourselves to the case where this limit equals the speed of light c. The general account can then be obtained by regarding c as an unspecified finite positive constant. As in the Newtonian case, we proceed by examining the Lagrangian for a free particle and for a system of interacting particles.

13.6.1 Space-Time Interval and Proper Time

Let S' be the inertial rest frame of a free particle moving with velocity $\mathbf{v}$ relative to an inertial frame S. To derive the special relativistic Lagrangian for the particle in the next section, we notice that the so-called space-time interval $\mathrm{d}s^2$ between two infinitely close events (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) in S defined by

$$\begin{aligned}\mathrm{d}s^2 &= c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 \\ &= c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2\end{aligned}$$

is invariant under Lorentz transformations:

$$\begin{aligned}\mathrm{d}s'^2 &= c^2\mathrm{d}t'^2 - \mathrm{d}x'^2 - \mathrm{d}y'^2 - \mathrm{d}z'^2 \\ &= c^2\gamma^2\left(\mathrm{d}t - \frac{v\mathrm{d}x}{c^2}\right)^2 - \gamma^2\left(\mathrm{d}x - v\mathrm{d}t\right)^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \\ &= \gamma^2\left(c^2\mathrm{d}t^2 - 2v\mathrm{d}x\mathrm{d}t + \frac{v^2\mathrm{d}x^2}{c^2}\right) - \gamma^2(\mathrm{d}x^2 - 2v\mathrm{d}x\mathrm{d}t + v^2\mathrm{d}t^2) - \mathrm{d}y^2 - \mathrm{d}z^2 \\ &= \gamma^2(c^2\mathrm{d}t^2 - \mathrm{d}x^2)\left(1 - \frac{v^2}{c^2}\right) - \mathrm{d}y^2 - \mathrm{d}z^2 \\ &= c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \\ &= \mathrm{d}s^2.\end{aligned}$$

During an infinitesimal period of time $\mathrm{d}t$ the particle moves a distance $\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}$ in S. Since the particle is at rest in S', we obtain $\mathrm{d}x' = \mathrm{d}y' = \mathrm{d}z' = 0$ in S'. By applying the invariance of the space-time interval, we get

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 = c^2\mathrm{d}t'^2.$$

If we isolate $\mathrm{d}t'$ in this equation, the result is

$$\mathrm{d}t' = \mathrm{d}t\sqrt{1 - \frac{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}{c^2\mathrm{d}t^2}}. \tag{13.16}$$

Expressing the speed v of the particle in S by the infinitesimal space-time interval,

$$v^2 = \frac{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}{\mathrm{d}t^2},$$

then allows us to rewrite (13.16) as

$$\mathrm{d}t' = \mathrm{d}t\sqrt{1 - \frac{v^2}{c^2}}. \tag{13.17}$$

The time τ displayed by a clock comoving with a given object is called the *proper time* for the object. Equation (13.17) therefore expresses the (infinitesimal) proper time $\mathrm{d}\tau$ for the particle in terms of the time $\mathrm{d}t$ measured by clocks at rest in S.

13.6.2 The Lagrangian for a Free Particle

We know from the derivation of the law of inertia that the Lagrangian for a free particle is only a function of its speed. Since the rest frame S' of the free particle must be inertial, it follows that the Lagrangian has a constant value $a = L'(0)$ in S'. In hindsight, we now introduce m through

$$a = -mc^2$$

and call it the (rest) mass of the particle. Next, we rewrite the principle of least action for the particle using (13.17):

$$\Gamma' = \int_{\tau_1}^{\tau_2} -mc^2 \,\mathrm{d}\tau = \int_{t_1}^{t_2} -mc^2\sqrt{1-\frac{v^2}{c^2}}\,\mathrm{d}t.$$

Comparing the right-hand side with the expression for the action relative to S

$$\Gamma = \int_{t_1}^{t_2} L(v)\,\mathrm{d}t, \tag{13.18}$$

we obtain the expression for the Lagrangian in S:

$$L(v) = -mc^2\sqrt{1-\frac{v^2}{c^2}}. \tag{13.19}$$

If we differentiate this formula with respect to $\mathbf{v}$, we get the relativistic formula for the linear momentum by using the Lagrangian formula for the generalized momentum:

$$\mathbf{p} = \frac{\partial}{\partial \mathbf{v}}\left(-mc^2\sqrt{1-\frac{v^2}{c^2}}\right) = \frac{m\mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}} = \gamma m\mathbf{v}.$$

Expanding (13.19) and (13.6.2) to first order in $\frac{v}{c}$, we obtain

$$L = -mc^2 + \tfrac{1}{2}mv^2, \qquad \mathbf{p} = m\mathbf{v}.$$

Recalling that a constant does not change the equations of motion, we realize that we have regained the classical formulas.

13.6.3 The Interaction of Particles with Fields

We continue by considering the interaction of particles. As already mentioned, it follows from the principle of relativity that the propagation speeds of interactions must correspond to the invariant limiting speed. In the relativistic case, this limiting speed is identified as the speed of light in empty space. Due to the finite propagation speed, a direct interaction at a given time is only possible between neighboring points in space. To mediate the interaction between particles that are at a certain distance from each other, Landau and Lifshitz introduce the concept of a force field:

> [I]nstead of saying that one particle acts on another, we may say that the particle creates a field around itself; a certain force then acts on every other particle located in this field. [...] Interactions can occur at any one moment only between neighbouring points in space (contact interaction). Therefore we must speak of the interaction of the one particle with the field, and of the subsequent interaction of the field with the second particle.[10]

In the following, we first explain how to obtain the interaction of a particle with a field using the principle of least action. To describe this interaction, we subtract a potential U from the Lagrangian for a free particle. In contrast to the Newtonian case, which requires an instantaneous action at a distance, the relativistic potential is allowed to depend explicitly on position and velocity:

$$L = -mc^2\sqrt{1 - \frac{\mathbf{v}^2}{c^2}} - U(\mathbf{r}, \mathbf{v}).$$

Substituting this Lagrangian into the equations of motion,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \mathbf{v}} = \frac{\partial L}{\partial \mathbf{r}},$$

then results in

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\gamma(\mathbf{v})\mathbf{v}) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial U}{\partial \mathbf{v}} = -\frac{\partial U}{\partial \mathbf{r}}. \tag{13.20}$$

This means that the change in linear momentum of the system depends on the partial derivative of the potential with respect to position. It is therefore the relativistic counterpart to Newton's second law of motion. In contrast

[10] Landau and Lifshitz 1972, p. 151.

to the latter, the relativistic formula contains a second term on the left-hand side, which is due to the velocity dependence of the potential. Put differently, because the field also has momentum, we cannot interpret $-\frac{\partial U}{\partial \mathbf{r}}$ as the force on the particle. However, if we interpret the force $\mathbf{F}$ on the particle as its change in linear momentum with respect to time

$$\mathbf{F} = \dot{\mathbf{p}} = \frac{\mathrm{d}}{\mathrm{d}t}(m\gamma(\mathbf{v})\mathbf{v}),$$

we can obtain an expression for $\mathbf{F}$ in terms of the potential by means of equation (13.20):

$$\mathbf{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\gamma(\mathbf{v})\mathbf{v}) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial U}{\partial \mathbf{v}} - \frac{\partial U}{\partial \mathbf{r}}.$$

To extend the description to a closed system of N interacting particles, we sum the relativistic Lagrangian for N free particles and subtract a potential U that is a function of the coordinates and velocities:

$$L = -c^2 \sum_\alpha m_\alpha \sqrt{1 - \frac{\mathbf{v}_\alpha^2}{c^2}} - U(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N),$$

so that

$$\Gamma = \int_{t_1}^{t_2} L\,\mathrm{d}t = \int_{t_1}^{t_2} -c^2 \sum_\alpha m_\alpha \sqrt{1 - \frac{\mathbf{v}_\alpha^2}{c^2}} - U(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N)\,\mathrm{d}t.$$

It follows that the force $\mathbf{F}_\alpha$ on the αth particle

$$\mathbf{F}_\alpha = \dot{\mathbf{p}}_\alpha = \frac{\mathrm{d}}{\mathrm{d}t}(m_\alpha\gamma(\mathbf{v}_\alpha)\mathbf{v}_\alpha)$$

is given by

$$\mathbf{F}_\alpha = \frac{\mathrm{d}}{\mathrm{d}t}(m_\alpha\gamma(\mathbf{v}_\alpha)\mathbf{v}_\alpha) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial U}{\partial \mathbf{v}_\alpha} - \frac{\partial U}{\partial \mathbf{r}_\alpha}.$$

As previously stated, Newton's third law of motion is an expression of the infinite propagation speed of interactions and so does not hold in special relativity. Despite this circumstance, relativistic momentum is conserved for a closed system, as shown in section 13.3. We should therefore regard the constancy of the speed of light in all inertial frames as the relativistic counterpart of Newton's law of reaction.

13.7 CONCLUSION

In this chapter we have clarified the relationship between space-time symmetries and conservation laws in mechanics. First, we have shown how to derive the law of inertia and the principle of relativity from comprehensibility conditions of homogeneity and isotropy, which hold both in the Newtonian and in the special relativistic framework. Secondly, we have seen how to derive the conservation of angular momentum and linear momentum. Finally, by specifying the speed of interactions, we were able to derive the divergent parts of the two frameworks. The last derivations are crucial as they correspond to comprehensibility arguments for Newtonian and special relativistic mechanics respectively. Moreover, the proofs explain how different sets of principles can define the same group of inertial frames. The first derivation, on the other hand, is important because it allows us to formulate mathematical blueprints for the construction and synchronization of ideal clocks that are shared across Newtonian physics, special relativity and ether-based electrodynamics. How to design these blueprints is the topic of the next chapter.

CHAPTER 14

Newtonian and Relativistic Chrono-Geometry

Both Poincaré and Einstein envisioned the synchronization of given ideal clocks by light signals, while Ignatowski and Rindler pointed out that any wave could be utilized. Following Roberto Torretti's analysis of Newtonian chrono-geometry, I will develop a third alternative that works in both the Newtonian and the special relativistic framework. In particular, I will show how to base the concept of ideal clocks on the conservation of linear and angular momentum. Secondly, we will see how to synchronize clocks without waves as signals. The aim of this analysis is, first, to emphasize Ignatowski's point that electrodynamics in general and the speed of light in particular should not play any special role (except a historical one) in special relativity. Second, this analysis aims to show how Torretti's Newtonian clock behaves relativistically in order to provide a better understanding of the transition from Newtonian to special relativistic physics. In particular, the aforementioned blueprint for the construction and synchronization of clocks will in principle enable us to empirically detect the dilation of clocks. Following Poincaré, we can still hold on to Newtonian space-time by arguing that the dilation is caused by the precession of the moving clock due to its interaction with the ether. While this strategy can explain the dilation of the moving disk, I conclude that it cannot properly resolve the twin paradox.

14.1 TORRETTI'S ANALYSIS OF NEWTONIAN CHRONO-GEOMETRY

Torretti developed his analysis in his influential book *Relativity and Geometry* (1982). His starting point is that "Newtonian time, like Newtonian space, is accessible only with the assistance of Newton's dynamical principles."[1] Friedman takes essentially the same position, claiming that these laws implicitly define the classical spatio-temporal framework. In contrast to Friedman, however, Torretti gives a very eloquent analysis of the chrono-geometric significance of Newton's dynamical principles. The key point in Torretti's account is the introduction of an ideal portable clock to replace the well-known definition of Newtonian time by the first law of motion:

> According to the First Law, a body moving along a graduated ruler will – if both it and the ruler are free from the action of impressed

[1] Torretti 1983, p. 12.

> forces – traverse equal distances in equal times. The flow of Newtonian time can therefore be read directly from the distance marks the body passes by as it moves along the ruler. [...] A more portable kind of Newtonian clock can be conceived in the light of the Principle of the Conservation of Angular Momentum, a familiar consequence of the laws of motion.[11] By this Principle a freely rotating rigid sphere of constant mass will indefinitely conserve the same angular velocity. The magnitude of the latter can be measured with a dynamometer attached to the surface of the body at some distance from its axis.[2]

Equipped with a series of such spinning clocks rotating with the same angular velocity ω_c, Torretti suggests that we place them at regular intervals along the graduated ruler of the original Newtonian clock. If we assume that the moving body is another spinning clock that rotates at the angular velocity ω_c, then it is possible to calibrate the clocks along the ruler as the moving clock passes by at uniform speed. Hence, Torretti points out, "[i]n this way one could, it seems, 'diffuse' the same time over all the universe."[3]

However, the method described proves to be ambiguous. For example, we could decide to resynchronize the clocks along our ruler. To do this, we take two spinning clocks A and B with angular velocity ω_c, place them at the beginning of our ruler and synchronize their readings. Clock A is then sent along the ruler at the constant speed v_A as measured by the clocks on the ruler, while clock B moves along it with constant speed $v_B \neq v_A$. Imagine that we observe that A agrees with the clocks at rest relative to the ruler as it passes by them, while clock B appears to run slower. From this it follows that the angular velocity of B as measured with the clocks along the ruler, must be less than ω_c. The reason why "this does not clash with our assumption that it [the angular velocity of B] is exactly ω_c *if measured with a dynamometer affixed to B*," Torretti now emphasizes, is that "[t]he local processes that measure out the time parameter of Newton's Laws of Motion will unambiguously define a universal time only if they agree with each other transitively and stably – i.e. only if the synchronism of any such process A with a process B at one time and with a process C at another time ensures the synchronism of B with C at any time."[4] From Torretti's point of view, the discrepancy thus clearly contradicts the premises of Newton's

[2] Ibid., pp. 12–3.
[3] Ibid., p. 13.
[4] Ibid., p. 13.

laws of motion, but it is by no means absurd. To recognize the latter, we are told, it took the genius of Einstein. Not only did he grasp that the problem had its origins in Newtonian dynamics, but he also managed to work out a solution to it in terms of a theory of special relativity. Torretti then analyzed Einstein's 1905 article on special relativity. However, he did not return to his example of rotating disks as ideal clocks to extend his Newtonian analysis to the relativistic case. So let us continue where Torretti left off.

14.2 THE TROLLEY CLOCK

To extend Torretti's analysis to special relativity, we take the article "A Relativistic Trolley" (2016) by Vadim Matvejev, Oleg Matvejev and Øystein Grøn as a starting point. Instead of a rotating disk, they imagine a trolley rolling at a constant angular velocity Ω on a straight rail:

> Imagine that an observer on a trolley is spinning the drive wheel to a certain angular velocity Ω, and that this wheel is rolling along the rail without slipping, thus driving the trolley.[5]

It follows that the linear speed v of the rotating rim of the wheels as measured in S becomes

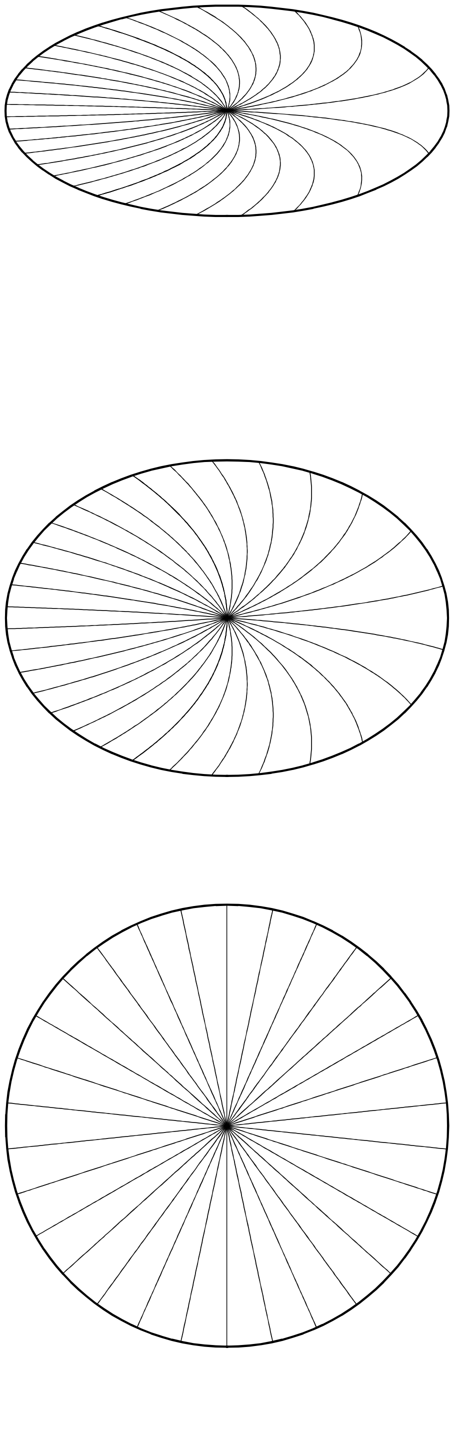
$$v = R\Omega,$$

where R denotes the radius of the rim of the rotating wheel. Before proceeding, although the authors do not emphasize this, we should clarify that 1) the wheels of the trolley are assumed to roll *purely* without slipping, and that 2) the rest frame S of the trolley is assumed to be inertial. That is to say, following Landau and Lifshitz, in it "space is homogeneous and isotropic and time is homogenous."[6] Due to the law of conservation of angular momentum in the inertial rest frame of the trolley, both the angular and linear velocity of the rim must remain unchanged in S. Since the wheels do not slip, it follows that the trolley rolls along the rail at the speed $v' = v$ with respect to the inertial rail frame S'.

Next, Matvejev and Grøn state that if we equip the trolley with a laser and a sensor, it can be understood as a ticking clock due to the constant angular velocity of the wheels:

[5] V. Matvejev, O. Matvejev, and Grøn 2016, p. 419.

[6] Landau and Lifshitz 1969, p. 5.



(a) $v' = 0$

(b) $v' = 0.7c$

(c) $v' = 0.9c$

Fig. 14.1: Lorentz contracted trolley wheel. In the first case, the wheel just spins without rolling. In the second and third case, it rolls along the railway to the left.

> Suppose now that on the trolley next to the wheel rim there is a generator of laser light and on the drive wheel there is a sensor. Each time the sensor passes the laser generator, it detects a signal. The laser and sensor act like a clock ticking with a period equal to the time it takes the sensor to pass the upper position of the wheel on two consecutive occasions.[7]

We now assume that we have positioned clocks along the railroad track. We can imagine these clocks as disks that rotate at the same angular velocity as our trolley clock. To synchronize the clocks in S' using the rolling trolley clock, we can imagine two different methods. Following the first method, we simply read the time t on the trolley clock as it passes by and set our railroad clock to the same time: $t' = t$. Following the second method, we first measure the initial distance l' between the trolley clock and the railroad clock. We also assume that the trolley begins rolling along the rails at $t = t' = 0$. When the trolley passes the railroad clock, we measure the speed v' of the trolley in S' using a photogate and our railroad clock and set the latter to $t' = \frac{l'}{v'}$.

Due to the absolute nature of time, both methods lead to the same synchronization in the Newtonian framework. It should come as no surprise that this is not the case in special relativity. What might be surprising is that the second method works in both frameworks. To see this and other unexpected results such as the Lorentz contraction of a trolley wheel rolling to the left at speeds close to c (see figure 14.1), we will turn to a relativistic analysis of the trolley clock in the next sections.

14.3 RELATIVISTIC ANALYSIS OF TROLLEY CLOCK

Let us consider the wheel of the trolley, which is equipped with a sensor. Following Matvejev and Grøn, we introduce an x-axis along the horizontal rail and a y-axis in the vertical direction passing through the center of the wheel. We then use P to denote the point at the origin of this coordinate system at time $t = 0$. It follows from this that the point moves along a circular path in the rest frame of the trolley given by

$$x = R\sin(\Omega t), \tag{14.1}$$

$$y = R(1 - \cos(\Omega t)). \tag{14.2}$$

[7] V. Matvejev, O. Matvejev, and Grøn 2016, p. 419.

To transform these equations to the rail frame S', we must apply the Lorentz transformations

$$t' = \gamma\left(t - \frac{v}{c^2}x\right),$$
$$x' = \gamma(x - vt), \tag{14.3}$$
$$y' = y, \tag{14.4}$$

where $v = R\Omega$, and the inverse time transformation

$$t = \gamma\left(t' + \frac{v}{c^2}x'\right). \tag{14.5}$$

In particular, inserting (14.1) and (14.5) into (14.3), the result is

$$x' = \gamma\left(R\sin\left(\gamma\Omega\left(t' + \frac{R\Omega}{c^2}x'\right)\right) - \Omega R\gamma\left(t' + \frac{\Omega R}{c^2}x'\right)\right)$$
$$\Rightarrow \quad x'\left(1 + \gamma^2\frac{\Omega^2R^2}{c^2}\right) = \gamma R\sin\left(\gamma\Omega\left(t' + \frac{R\Omega}{c^2}x'\right)\right) - \Omega R\gamma^2 t'$$
$$\Rightarrow \quad x' = R\left(\gamma^{-1}\sin\left(\gamma\Omega\left(t' + \frac{R\Omega}{c^2}x'\right)\right) - \Omega t'\right), \tag{14.6}$$

where we have made use of

$$1 + \gamma^2\frac{\Omega^2R^2}{c^2} = \gamma^2.$$

If we now insert (14.2) and (14.5) into (14.4), we get

$$y' = R(1 - \cos(\Omega t)) = R\left(1 - \cos\left(\gamma\Omega\left(t' + \frac{R\Omega}{c^2}x'\right)\right)\right). \tag{14.7}$$

Matvejev and Grøn go on to determine the distance l' between two consecutive points at which P touches the rail. Since the trolley moves in the negative direction of the x'-axis in the rail frame S', we impose the conditions $x' = -l'$ and $y' = 0$ at the next contact point. Substituting these into (14.7) gives

$$0 = R\left(1 - \cos\left(\gamma\Omega\left(t' + \frac{R\Omega}{c^2}x'\right)\right)\right)$$
$$= R\left(1 - \cos\left(\gamma\Omega\left(t' - \frac{R\Omega}{c^2}l'\right)\right)\right).$$

Hence,

$$1 = \cos\left(\gamma\Omega\left(t' - \frac{R\Omega}{c^2}l'\right)\right).$$

The cosine is equal to 1 whenever its argument is an integer multiple of 2π. For the first return to the rail, we take the smallest positive value:

$$\gamma\Omega\left(t' - \frac{R\Omega}{c^2}l'\right) = 2\pi. \tag{14.8}$$

Solving for t' yields

$$t' = \frac{2\pi}{\gamma\Omega} + \frac{R\Omega}{c^2}l'. \tag{14.9}$$

Next, we insert the conditions $x' = -l'$ and equation (14.8) into (14.6), which governs the x'-motion of the rim point:

$$-l' = R\left(\gamma^{-1}\sin(2\pi) - \Omega t'\right).$$

Since $\sin(2\pi) = 0$, the equation reduces to

$$l' = R\Omega t'. \tag{14.10}$$

Finally, we eliminate t' between equations (14.9) and (14.10). Substituting $t' = \frac{l'}{R\Omega}$ into (14.9) gives

$$\frac{l'}{R\Omega} = \frac{2\pi}{\gamma\Omega} + \frac{R\Omega}{c^2}l'.$$

Multiplying through by $R\Omega$ and rearranging, we obtain

$$l'\left(1 - \frac{R^2\Omega^2}{c^2}\right) = \frac{2\pi R}{\gamma}.$$

Recalling that $1 - \frac{R^2\Omega^2}{c^2} = \frac{1}{\gamma^2}$, this simplifies to

$$l' = 2\pi R\gamma. \tag{14.11}$$

Thus, in the rail frame the wheel rolls out a distance per revolution that exceeds its rest circumference by the Lorentz factor γ.

Before we continue, we note that by inserting $l' = 2\pi R\gamma$ into (14.9) we can determine the time dilation between the wheel's period of revolution as measured in S' and S:

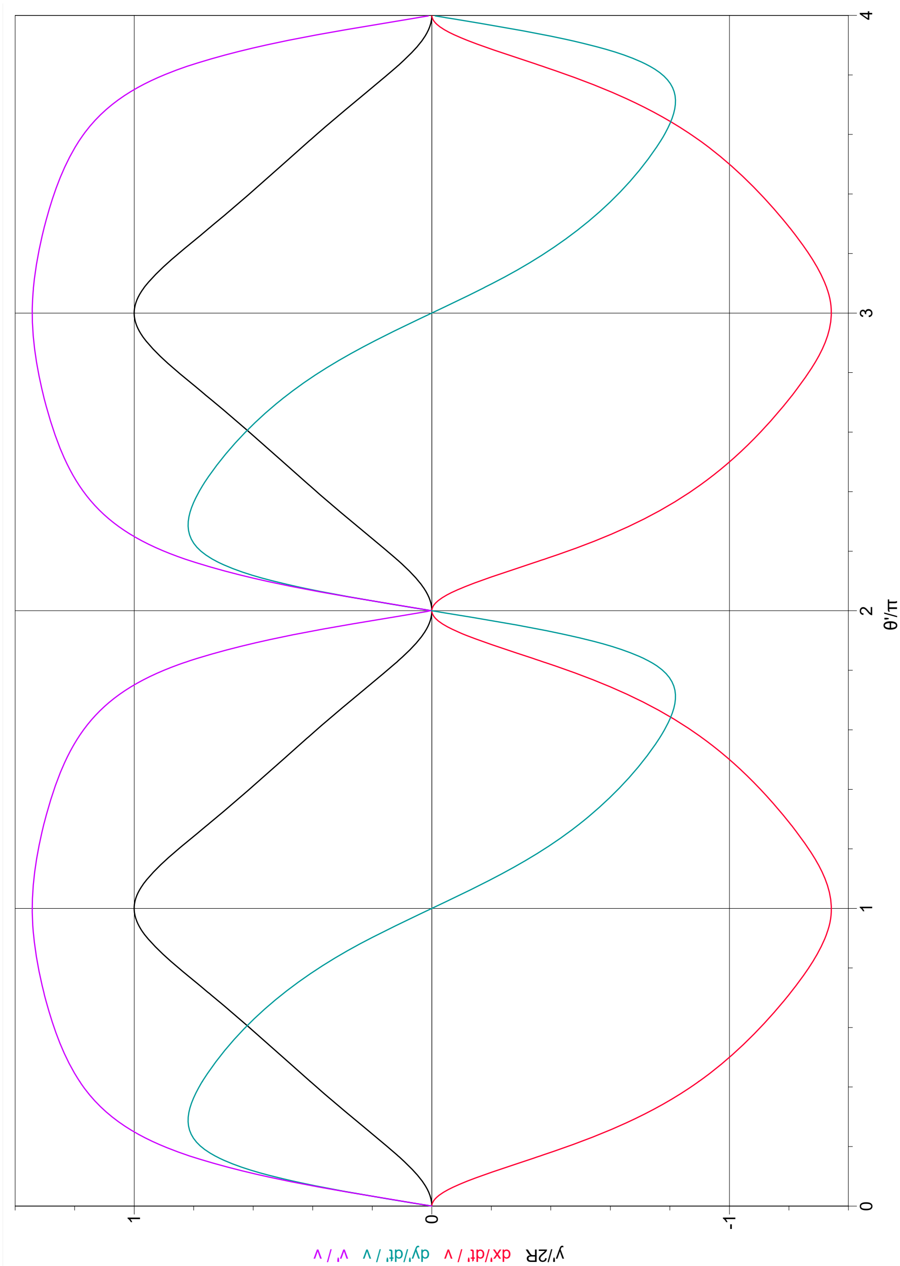


Fig. 14.2: Relativistic graphs for the motion of a point at the rim of the trolley wheel.

$$T' = \frac{2\pi}{\gamma\Omega} + \frac{R\Omega}{c^2} 2\pi R\gamma = \frac{2\pi}{\gamma\Omega}\left(1 + \frac{R^2\Omega^2}{c^2}\gamma^2\right) = \frac{2\pi}{\gamma\Omega}\gamma^2 = \gamma T.$$

If we divide l' by T', we obtain the linear speed v' of the trolley as required by the reciprocity of S and S':

$$v' = \frac{l'}{T'} = \frac{2\pi R\gamma}{T'} = \frac{2\pi R\gamma}{\gamma T} = \frac{2\pi R}{T} = v.$$

Let us now examine the formula (14.11). This relationship is remarkable due to the circumstance that the trolley wheel is Lorentz contracted in the rail frame S'. For it follows that the circumference C' of the wheel as measured in S' is shorter than its circumference C as measured in S:

$$C' < C = 2\pi R.$$

Nevertheless, the wheel "rolls out" a distance along the rail per turn that is longer than the circular circumference of the wheel C in S:

$$l' = \gamma C > C > C'.$$

Due to its puzzling nature, Matvejev and Grøn call this result the *trolley paradox*. To resolve it, they notice that every time the point P touches the rail, it is instantaneously at rest in the rail frame S'. Because if the wheel moved at non-zero speed, it would be slipping the rail. We can see that the equations of motion for the trolley wheel (14.6) and (14.7) actually fulfill this condition if we look at the black and purple graphs in figure 14.3. The common x-axis describes the angle θ' as measured in S' between the perpendicular through the center O of the wheel and the line through O and P. The black graph then outlines the height y' of P above the rail (divided by the wheel diameter) as a function of θ', while the purple graph shows the speed ν' of P as a function of θ'.[8] If we compare the two graphs, we find that ν' is 0 when P touches the rail ($y' = 0$), in accordance with the above condition. This in turn implies that the trolley wheel rolls out the rest length of its rim along the rail. As indicated above, in the trolley frame S the rim rotates with a linear or tangential velocity v, which is given by $v = \Omega R$. In the rail frame S', the same translational velocity $v' = v$ is observed, but the revolution period is dilated to $T' = \gamma T$. We can therefore compute the distance

[8] More precisely, the purple graph shows the speed ν' of P in S' divided by its (constant) speed v in S as a function of θ'.

l' rolled out along the rail in one revolution by integrating the translational motion over the time T':

$$l' = \int_0^{T'} v' \mathrm{d}t' = v'T' = v\frac{2\pi R}{v}\gamma = 2\pi R\gamma,$$

in agreement with equation (14.10). This shows that the longer rolling distance arises directly from the time dilation of the revolution period combined with the unchanged translational speed.

At the same time, we note that in the rail frame S' the circumference C' of the wheel, measured at a single instant t', is shorter than its rest-frame circumference $C = 2\pi R$. In fact,

$$C' < C = 2\pi R,$$

since the rim is deformed anisotropically by Lorentz contraction. Nevertheless, the wheel rolls out a distance per revolution that is longer than its own rest circumference C:

$$l' = 2\pi R\gamma > C > C'.$$

This apparent inconsistency is the essence of the so-called *trolley paradox.*

In order to understand how it is possible that the circumference of the rolling trolley wheel C' as measured in S' is shorter than its circumference C in S, we examine both the Lorentz-contracted trolley wheel (see figure 14.1) and the horizontal velocity $v'_x = \frac{\mathrm{d}x'}{\mathrm{d}t'}$ (divided by v) of the point P as a function of θ' (see figure 14.3). First, we recall that $v'_x \leq 0$ because the trolley wheel moves in the negative direction of the x'-axis. Secondly, we note that v'_x decreases in the interval $[0; \pi]$, reaches its minimum for $\theta' = \pi$ and increases in $[\pi; 2\pi]$. This implies that the distance between neighboring points $\mathrm{d}l'$ on the rim decreases with height: only the component along x' is Lorentz-contracted, by

$$\frac{1}{\gamma(\theta')} = \sqrt{1 - \frac{v'_x(\theta')^2}{c^2}},$$

so the local spacing depends on $v'_x(\theta')^2$ through the Lorentz factor. The resulting deformation of the trolley wheel is vividly illustrated in figure 14.1 as the changing gaps between spokes. This spatially nonuniform contraction explains why the wheel can cover, per revolution, a distance along the rail that exceeds its simultaneous circumference C' in S'. Near the contact region $v'_x = 0$, so there is essentially no contraction along x', and the rim "lays down" its largest segments onto the rail; aloft, where $|v'_x|$ is largest, segments are most contracted. Integrating once around the rim at fixed t' gives $C' < 2\pi R$, yet the no-slip rolling constraint ties the translational advance per turn to the material length deposited near contact, allowing the travelled distance per revolution $l' = 2\pi\gamma R$ to exceed C' without paradox. Hence the trolley paradox is resolved.

We then examine the synchronization of the clocks along the rail. Without loss of generality, we consider a rail clock that is one revolution l' away from the trolley at $t = 0$. According to the first method, we would simply read the trolley clock $t = T$ as it passes and set the railroad clock to

$$t' = t = T,$$

where T denotes the wheel's period of revolution as measured in S.[9] According to the second method, we would measure the speed v' of the trolley and divide the distance l' by it:

$$t' = \frac{l'}{v'} = \frac{2\pi R\gamma}{v} = \gamma T = T'.$$

In contrast to the first method, the second conforms to the special relativistic framework. Thus, the thought experiment regarding the trolley clock demonstrates how clock synchronization might reveal relativistic time dilation. However, by excluding a distinction between real and apparent measurements from the outset, this argument is based on an ontological interpretation of the relativity principle. Thus, Poincaré would arguably have tried to hold on to Newtonian space-time by replying that the time discrepancy is not real, but merely the result of the precession of the trolley clock due to its contraction caused by the interaction with the ether.

14.4 THE TWIN PARADOX

Even if such a reply would explain the dilation of spinning clocks, it is questionable how such a protective strategy would accommodate the twin paradox. To relate the latter to the trolley clock thought experiment, we

[9] Strictly speaking, we would have to set $T = 1$, as we have no other measures of time apart from the rolling wheel. However, for the sake of argument, we proceed without specifying the value of T.

just have to imagine one twin on the trolley, while the other stays at the beginning of the rail track. The twin on the trolley then travels back and forth on the track resulting in an age difference between the two twins given by the γ-factor. More precisely, while the twin on the trolley is going to age a time Δt, the twin that stayed at the beginning of the track will age a time $\Delta t' = \gamma \Delta t > \Delta t$. To calculate the latter, we recall that the rest frame S of the trolley is moving along the negative direction of the x'-axis of the rail frame S' at speed v. According to the Lorentz transformations, we thus have

$$\begin{aligned} t' &= \gamma\left(t - \frac{v}{c^2}x\right), \\ x' &= \gamma(x - vt). \end{aligned} \tag{14.12}$$

Since the twin on the trolley is at rest in S, we can place him at $x = 0$ without loss of generality. Starting at the common origin of S and S' at $t_0 = t'_0 = 0$, the trolley shifts its direction at $t_1 = \frac{1}{2}\Delta t$. While the turning point P has coordinates $\left(\frac{1}{2}\Delta t, 0\right)$ in S, its coordinates in S' can be calculated to $\left(\frac{1}{2}\gamma\Delta t, -\frac{1}{2}\gamma v\Delta t\right)$ by means of (14.12). Due to the shift of direction at P, measured in S' the trip back will be given by the equation

$$\begin{aligned} x' &= v(t' - t'_1) + x'_1 \\ &= v\left(t' - \frac{1}{2}\gamma\Delta t\right) - \frac{1}{2}\gamma v\Delta t \\ &= vt' - \gamma v\Delta t. \end{aligned} \tag{14.13}$$

It follows that the twin on the trolley will be back at the origin of S' ($x' = 0$) at time t'_2 given by

$$\begin{aligned} 0 &= vt'_2 - \gamma v\Delta t \\ \Rightarrow t'_2 &= \gamma\Delta t. \end{aligned}$$

So while the biological age of the twin on the trolley has increased by $t_2 = \Delta t$, the twin who remained at the beginning of the rail has aged by $t'_2 = \gamma\Delta t > \Delta t$. The only way Poincaré could have attempted to account for this result would be to allow the ether to dilate not only spinning clocks, but all temporal processes, including biological aging. However, since time dilations correspond only to apparent measurements, this would mean that we cannot evaluate the biological age of the younger twin as ontologically real. While such a strategy might be feasible from a conceptual point of view, as Lange points out, it "cries out to be eliminated – as Einstein does."[10]

14.5 CONCLUSION

In this chapter, we have shown how to imagine an interpretive scheme for the construction and synchronization of ideal clocks shared across the Newtonian and the special relativistic framework. More precisely, we imagined disks rotating at the same constant angular velocity as ideal clocks and placed them at a certain distance along a track. Next, we imagined synchronizing the stationary clocks by moving another clock along the track at a constant rectilinear speed. So far, practitioners of both frameworks should not suspect any problems. But repeating the procedure with a clock moving at a different speed should eventually reveal inconsistencies (in the form of special relativistic effects) to the Newtonian. To protect his spatiotemporal framework, he could follow Poincaré and attribute the dilation of the moving clock to its causal interaction with the ether. While this strategy might be feasible for spinning clocks, we concluded that it would not satisfactorily explain the twin paradox.

[10] Lange 2005, p. 709.

CHAPTER 15

Comprehending Friedman and the Dynamics of Reason

In this final chapter, I discuss the limitations of Friedman's concept of relativized a priori constitutive principles. First, I argue that the range of possibilities for safeguarding physical theories is narrower than Friedman acknowledges. Second, I question the uniqueness of Friedman's constitutive principles in the context of both Newtonian and special relativistic physics by arguing that different constitutions of the same spatio-temporal structure are possible. The Global Positioning System and the Gravity Probe B experiment are used as examples of the empirical realizability of two different schemes. Finally, I conclude that we should understand the role of the a priori in the context of mathematical physics in terms of Darrigol's comprehensibility conditions.

15.1 HOLISM

Following Darrigol's critique of holism, I would like to argue why the viability of protective strategies is narrower than Friedman suggests. According to the latter, Einstein accommodated the empirical discovery of the constancy of the speed of light in any inertial frame by elevating it to a constitutive principle:

> This same process of "elevation," in Einstein's hands, then makes it clear how an extension or continuation of Kant's original conception can also accommodate new and surprising empirical facts – in this case, the very surprising empirical discovery (to one or another degree of approximation) that light has the same constant velocity in every inertial frame. It now turns out, in particular, that we can not only impose already familiar and accepted mathematical frameworks (Euclidean geometry) on our rough and approximate perceptual experience, but, in appropriate circumstances, we can also impose entirely unfamiliar ones (the kinematical framework of special relativity).[1]

At the same time, Friedman emphasizes, Poincaré was justified in holding on to the Newtonian framework, including Newtonian mechanics, until the advent of general relativity:

[1] Friedman 2010, p. 708.

> Einstein's "elevation" of the principle of relativity and the light principle to the status of "presuppositions" or "postulates" definitive of a new geometrical-mechanical constitutive framework is thereby shown to be possible, but not necessarily obligatory. Thus, as I emphasized above, Poincaré, for example, had every right to resist this move and to continue to maintain the Newtonian framework – at least before the incorporation of gravitation into relativity around 1912.[2]

What bothers me most about Friedman's analysis of the transition from the Newtonian to the special relativistic framework is that it leaves more room for holism than a careful reading of this history provides. First of all, my historical investigation shows that by 1905 Poincaré had abandoned Newtonian mechanics to accommodate developments in electrodynamics. Although the Frenchman thought that he did not have to follow Einstein and discard Newtonian space-time in favor of its special relativistic successor, he was nevertheless aware that what he called "the new mechanics" implied the reshaping of all physics in Lorentz-invariant form. Second, my comparison between the trolley clock and the twin paradox in the previous chapter strongly suggests that Poincaré ultimately had to reject Newtonian space-time even without the advent of general relativity. Although he would be able to attribute the dilation of the moving clock to its precession caused by the ether, it is difficult to see how he could neglect the age difference between the twins as ontologically real. Before his death in 1912, of course, there was no experimental evidence for time dilation. So the point is not that it was irrational to hold on to Newtonian space-time at that time, but that the range of possibilities for safeguarding physical theories is narrower than Poincaré and Friedman admit.

15.2 TRANSCENDENTAL NECESSITY

A related aspect of Friedman's concept of constitutive principles is his understanding of their relativized transcendental character:

> I now reject the logical empiricists' notion of "coordinating principles" in favor of a truly transcendental constitutive a priori,[352] and I also insist, accordingly, on the transcendental necessity of the principles in question – relative to a given intellectual situation.[3]

[2] Friedman 2010, fn. 299.

[3] Ibid., p. 727.

Einstein's "postulates" of special relativity thereby serve as Friedman's primary example of this relativized transcendentalism:

> A central contention of Kant's original version of transcendental philosophy, as we know, is that the three Newtonian Laws of Motion are not mere empirical laws but a priori constitutive principles on the basis of which alone the Newtonian concepts of space, time, and motion can then have empirical application and meaning. What we have just seen is that Einstein's two fundamental "presuppositions" or "postulates" play a precisely parallel role in the context of special relativity.[4]

The two postulates refer to the relativity principle and the light principle respectively. Since these two principles enabled the first application of a non-Euclidean spatio-temporal structure to experience, they represent, according to Friedman, "the very first instantiation of a relativized and dynamical conception of the a priori."[5] I will now argue that my last two chapters call into question Friedman's concept of constitutive a priori principles as necessary in the sense of their uniqueness.

15.3 ALTERNATIVE SCHEMES

To argue against Friedman's uniqueness claim, we need only point to the possibility of alternative definitions in the works of Ignatowski and Landau and Lifshitz. As we have seen, Ignatowski offered a more general definition of inertial frames than Einstein, based on homogeneity and isotropy requirements. Landau and Lifshitz, for their part, derived the conservation of angular momentum and linear momentum by utilizing space-time symmetries together with the principle of least action. These developments are of crucial importance for our present discussion, as they allow for equivalent constitutions of the same spatio-temporal structure. Friedman is right, for example, when he says that the practitioners of Newtonian physics historically based the definition of inertial frames on Newton's three laws of motion. However, if we follow Ignatowski, we could also base Newtonian space-time on the homogeneity of space and time, the isotropy of space and the requirement that non-contact forces act instantaneously. A third constitution would be the conservation of angular momentum and linear

[4] Ibid., p. 708.
[5] Ibid., p. 708.

momentum together with the invariance of mass. Einstein based the space-time structure of special relativity on the relativity principle and the constancy of the speed of light. A second option would be the homogeneity of space and time, the isotropy of space and the requirement that non-contact forces propagate at the same finite speed. Alternatively, we could require the conservation of angular momentum and linear momentum together with the constancy of the speed of light. The possibility of applying the same spatio-temporal structure in different ways suggests that it is artificial to distinguish sharply between constitutive principles and proper empirical laws.

15.4 THE GRAVITY PROBE B EXPERIMENT

Friedman might follow Darrigol and object that the empirical realization of a reference frame by one of my alternative constitutions is less evident than its historical counterpart. In particular, while the arrival of the Global Positing System (GPS) has confirmed the realizability and effectiveness of clock synchronization via electromagnetic waves even in a general relativistic setting, an empirical counterpart to the trolley thought experiment seems more than questionable. Despite the immediate reasonableness of this assessment, the National Aeronautics and Space Administration (NASA) developed spherical gyroscopes similar to the trolley clock blueprint a decade ago as part of the realization of the Gravity Probe B experiment. The plan was to send spinning gyroscopes into orbit around the earth to measure general relativistic effects caused by the earth's rotation:

> Time and space, according to Einstein's theories of relativity, are woven together, forming a four-dimensional fabric called "space-time." The mass of Earth dimples this fabric, much like a heavy person sitting in the middle of a trampoline. Gravity, says Einstein, is simply the motion of objects following the curvaceous lines of the dimple. If Earth were stationary, that would be the end of the story. But Earth is not stationary. Our planet spins, and the spin should twist the dimple, slightly, pulling it around into a 4-dimensional swirl. This is what GP-B went to space in 2004 to check.
>
> The idea behind the experiment is simple: Put a spinning gyroscope into orbit around the Earth, with the spin axis pointed toward some distant star as a fixed reference point. Free from external forces, the gyroscope's axis should continue pointing at the star – forever. But

> if space is twisted, the direction of the gyroscope's axis should drift over time. By noting this change in direction relative to the star, the twists of space-time could be measured.[6]

The primary challenge was to produce gyroscopes that were spherical enough to keep their angular velocity in the absence of "outer" effects:

> The four gyroscopes in GP-B are the most perfect spheres ever made by humans. These ping pong-sized balls of fused quartz and silicon are 1.5 inches across and never vary from a perfect sphere by more than 40 atomic layers. If the gyroscopes weren't so spherical, their spin axes would wobble even without the effects of relativity.[7]

Next, the researchers had to develop technologies to measure the changes in the gyroscope's axis caused by the space-time vortex around the earth. The latter required them to eliminate disturbances caused by the transportation into orbit, the earth's magnetic field and the measurement of the gyroscope's angular velocity:

> [GP-B researchers] developed a "drag free" satellite that could brush against the outer layers of Earth's atmosphere without disturbing the gyros. They figured out how to keep Earth's magnetic field from penetrating the spacecraft. And they created a device to measure the spin of a gyro – without touching the gyro.[8]

Friedman might agree that both the trolley clock and the Gravity Probe B experiment are fascinating, but he will almost certainly counter that their retrospective nature means that they tell us little about the rational development of physics from classical mechanics to special relativity. I would reply that no one questions that the light principle played a crucial role in Einstein's solution to the prevailing crisis by means of a revolutionary theory of space and time. However, this does not imply that Einstein himself understood (or that we should understand) the conjunction of the light principle and the relativity principle as the unique set of constitutive principles that makes the new space-time framework empirically well-defined. As shown above, alternative derivations and mutual empirical realizations are possible. Consequently, the necessity does not depend on the uniqueness

[6] Phillips 2011.
[7] Ibid.
[8] Ibid.

of a set of constitutive principles. The point is rather that certain fundamental assumptions – so-called comprehensibility conditions – necessarily imply a unique space-time structure. Moreover, the principles do not define this space-time empirically in and of themselves. The latter, Darrigol emphasizes, requires both interpretive schemes and empirical know-how. Therefore, it is neither problematic for the status of these conditions that 1) other sets of conditions also imply the same spatio-temporal framework or 2) that multiple interpretive schemes can empirically realize this framework.

15.5 COMPREHENDING SPACE-TIME

Rather than endorsing Friedman's concept of constitutive principles, I conclude that we should follow Darrigol and understand the relativized a priori as the presupposition of certain natural assumptions or principles on which we base necessity arguments and derivations employed in theory construction. In particular, the isotropy of space and the homogeneity of space and time, together with the principle of least action, severely constrain what Friedman calls the space of empirical possibilities.

As I have argued in detail, both Poincaré and Einstein realized right away that the null result of ether drift experiments called into question not only the concurrent theory of electrodynamics, but also the foundations of physics itself. By assuming the principle of relativity between a frame at rest and a frame in uniform rectilinear motion with respect to the ether, Poincaré demonstrated the invariance of Maxwell's equations under Lorentz transformations. Einstein did not propose the existence of the ether, but arrived at the Lorentz transformations solely by applying homogeneity and isotropy requirements, the relativity principle and the constancy of the speed of light. Consequently, the transformations not only expressed the symmetries of (electro)dynamics, but also a revolutionary theory of space and time. In contrast to Einstein, Ignatowski derived the Lorentz transformations and their degenerate Galilean limit on the basis of homogeneity and isotropy requirements, the relativity principle and a few supplementary conditions of comprehensibility, such as causality. An essential part of this derivation was the proof of the existence of a limiting velocity for the propagation of any interaction. Since Newton's law of universal gravitation presupposes action-at-a-distance, it followed from the above that all non-contact forces propagate at infinite speed. However, no one doubted that electromagnetic waves propagate at a finite speed. Consequently, either some of the homogeneity and isotropy assumptions and their consequences are false, or all non-contact forces, including gravity, must propagate at the speed of light. In 1905, both Einstein and Poincaré opted for the latter solution. In particular, they started working on a Lorentz-invariant gravitational theory. From this perspective, it was clear to Einstein from the outset that there was nothing special about light propagation. So why did it play such a central role in his relativity paper? The short answer is that the constancy of the speed of light and the principle of relativity in and of themselves were considered uncontroversial. Nevertheless, Einstein argued that their conjunction, together with some additional assumptions, implied the comprehensibility of a revolutionary conception of space and time.

By combining the homogeneity and isotropy of space and the homogeneity of time with the principle of least action, Landau and Lifshitz showed how to derive the conservation of linear momentum and angular momentum. What makes this deduction remarkable is that it was based solely on principles common to both the Newtonian and the special relativistic framework. The divergent mechanical parts of the two frameworks were finally derived by specifying the speed of interactions. The first derivation is significant because it allows us to formulate mathematical blueprints for the construction and synchronization of ideal clocks that are shared across Newtonian physics and special relativity, while the second and third derivations are essential because they correspond to an argument for the comprehensibility of Newtonian and special relativistic mechanics, respectively. First and foremost, blueprints serve as thought experiments that allow us to compare different theories with each other. In our case, they allow us to comprehend the behavior of the same ideal clock under different geometrical-mechanical frameworks, thus reducing the problem of holism and incommensurability. Necessity arguments, for their part, help us to develop theories based on assumptions that are widely regarded as natural or unproblematic. Einstein, for example, based special relativity on homogeneity and isotropy assumptions that nobody questioned at the time. Therefore, although the theory had controversial implications, such as time dilation, to refute it one would have to expose the flaws in its derivation. As I argued in the previous chapter, for Poincaré to incorporate time dilation into his competitive theory would require that the ether had a causal influence on temporal processes. Even though it conflicted with the assumptions of homogeneity and isotropy, the somewhat similar idea that matter could influence space and time was developed only a few years later by Einstein in his efforts to reconcile gravitational theory with special relativity. But that's a story for another book.

Bibliography

Abraham, Max (1902). "Prinzipien der Dynamik des Elektrons". In: *Physikalische Zeitschrift* 4.1b, pp. 57–63.

— (1903). "Prinzipien der Dynamik des Elektrons". In: *Annalen der Physik* 10, pp. 105–179.

— (1905). *Theorie der Elektrizität: Elektronmagnetische Theorie der Strahlung*. Leipzig: Teubner.

Brown, Harvey R. (2005). *Physical Relativity: Space-Time Structure from a Dynamical Perspective*. Oxford: Clarendon Press.

Coffa, J. Alberto (1995). *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Ed. by Linda Wessels. Repr. First published 1991. Cambridge: Cambridge University Press.

Darrigol, Olivier (1995). "Henri Poincaré's Criticism of Fin de Siècle Electrodynamics". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 26, pp. 1–44.

— (2006). "The Genesis of the Theory of Relativity". In: *Poincaré Seminar 2005: Einstein 1905 - 2005*. Ed. by Damour, Thibault and Darrigol, Olivier, Duplantier, Bertrand, and Rivasseau, Vincent. Basel: Birkhäuser, pp. 1–32.

— (2014). *Physics and Necessity. Rationalist Pursuits from the Cartesian Past to the Quantum Present*. Oxford University Press.

— (2020). "Constitutive Principles versus Comprehensibility Conditions in Post-Kantian Physics". In: *Synthese* 197, pp. 4571–4616. URL: https://doi.org/10.1007/s11229-018-01948-2.

DiSalle, Robert (2006). *Understanding Space-Time*. Cambridge: Cambridge University Press.

— (2010). "Synthesis, the Synthetic A Priori, and The Origins of Modern Space-Time Theory". In: *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science*. Ed. by Michael Friedman, Mary Domski, and Michael Dickson. Open Court.

Einstein, Albert (1905). "Zur Elektrodynamik bewegter Körper". In: *Annalen der Physik* 322.10, pp. 891–921.

— (1952). "On the Electrodynamics of Moving Bodies". In: *The Principle of Relativity*. A collection of original memoirs on the special and general theory of relativity with notes by A. Sommerfeld transl. (from the German) by W. Perrett and G.B. Jeffery. New York: Dover Publications, pp. 35–65. Trans. of “Zur Elektrodynamik bewegter Körper”. In: *Annalen der Physik* 322.10 (1905), pp. 891–921.

Friedman, Michael (1983). *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science.* Princeton, N.J.: Princeton University Press.

— (2001). *Dynamics of Reason.* 1999 Kant Lectures at Stanford University. Stanford, California: CSLI Publications.

— (2010). “Synthetic History Reconsidered”. In: *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science.* Ed. by Michael Friedman, Mary Domski, and Michael Dickson. Open Court.

Friedman, Michael, Mary Domski, and Michael Dickson (2010). *Discourse on a New Method: Reinvigorating the Marriage of History and Philosophy of Science.* Open Court.

Frisch, David H. and James H. Smith (1963). “Measurement of the Relativistic Time Dilation Using Mu-Mesons”. In: *American Journal of Physics* 31.5, pp. 342–355.

Griffiths, David J. (1998). *Introduction to Electrodynamics.* 3rd ed. Benjamin Cummings.

Ignatowski, Vladimir (1910). “Einige allgemeine Bemerkungen zum Relativitätsprinzip”. In: *Physikalische Zeitschrift* 11, pp. 972–6.

— (1911). “Eine Bemerkung zu meiner Arbeit ‘Einige allgemeine Bemerkungen zum Relativitätsprinzip’”. In: *Physikalische Zeitschrift* 12, p. 799.

Janssen, Michael (1995). *A Comparison between Lorentz’s Ether Theory and Special Relativity in the Light of the Experiments of Trouton and Noble.* Ph.d. Thesis. University of Pittsburgh.

Johnson, Paul (2013). *Modern Times: A History of the World From the 1920s to the Year 2000.* English. Weidenfeld & Nicolson.

Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions.* 1st. Chicago: University of Chicago Press.

Landau, Lev D. and Evgeny Lifshitz (1969). *Mechanics.* 2nd ed. Pergamon Press.

— (1972). *Mechanics and Electrodynamics.* 1st ed. Pergamon Press.

Lange, Marc (2004). “Review Essay on *Dynamics of Reason* by Michael Friedman”. In: *Philosophy and Phenomenological Research* 68.3, pp. 702–712.

— (2005). “Theory and Reality: An Introduction to the Philosophy of Science”. In: *Australasian Journal of Philosophy* 83.1, pp. 122–124.

Lee, A. R. and T. M. Kalotas (1975). "Lorentz Transformations from the First Postulate". In: *American Journal of Physics* 43, pp. 434–437.

Lévy-Leblond, J.-M. (1976). "One More Derivation of the Lorentz Transformation". In: *American Journal of Physics* 44, pp. 271–277.

Lorentz, Hendrik Antoon (1892). "De Relatieve Beweging van de Aarde en den Aether". In: *Versl. Kon. Ak. Wet.* 1, pp. 74–79.

— (1904). "Maxwells elektromagnetische Theorie". In: *Encyclopädie der mathematischen Wissenschaften, mit Einschluss ihrer Anwendungen.* Ed. by Arnold Sommerfeld. Vol. 5. 2. Leipzig: Teubner. Chap. 13.

— (1916). *The Theory of Electrons and Its Application to the Phenomena of Light and Radiant Heat.* 2nd. Originally published 1909. Leipzig: Teubner.

— (1936). "La Théorie électromagnétique de Maxwell et son Application aux Corps mouvants". In: *Collected Papers.* Vol. II. Originally published 1892. The Hague: Martinus Nijhoff, pp. 164–343.

— (1937a). "Electromagnetic Phenomena in a System Moving With Any Velocity Smaller Than That of Light". In: *Collected Papers.* Vol. V. Originally published 1904. The Hague: Martinus Nijhoff, pp. 172–197.

— (1937b). "The Relative Motion of the Earth and the Ether". In: *Collected Papers.* Vol. IV. The Hague: Martinus Nijhoff, pp. 219–223. Trans. of "De Relatieve Beweging van de Aarde en den Aether". In: *Versl. Kon. Ak. Wet.* 1 (1892), pp. 74–79.

— (1937c). "Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern". In: *Collected Papers.* Vol. V. Originally published 1895. The Hague: Martinus Nijhoff, pp. 1–137.

Matvejev, Vadim, Oleg Matvejev, and Øystein Grøn (2016). "A Relativistic Trolley Paradox". In: *American Journal of Physics* 84.6, pp. 419–425. URL: https://doi.org/10.1119/1.4942168.

Miller, Arthur I. (1986). "A Study of Henri Poincaré's 'Sur la Dynamique de l'Électron'". In: *Frontiers of Physics: 1900–1911: Selected Essays.* Boston, MA: Birkhäuser Boston, pp. 29–150. URL: https://doi.org/10.1007/978-1-4684-0548-4_2.

Minkowski, Hermann (1909). "Raum und Zeit". In: *Physikalische Zeitschrift* 10. Lecture delivered at the 80th Assembly of German Natural Scientists and Physicians, Cologne, 21 September 1908, pp. 75–88.

Phillips, Tony (2011). *NASA Announces Results of Epic Space-Time Experiment.* Science @ NASA. URL: https://science.nasa.gov/science-news/science-at-nasa/2011/04may_epic.

Poincaré, Henri (1891). *Elektrizität und Optik.* Julius Springer.

— (1895a). "A Propos de la Théorie de M. Larmor". In: *L'Éclairage Électrique* 3. Part 1 & 2, pp. 5–13, 289–295.

— (1895b). "A Propos de la Théorie de M. Larmor". In: *L'Éclairage Électrique* 5. Part 3 & 4, pp. 5–14, 385–392.

— (1898). "La Mesure du Temps". In: *Revue de Métaphysique et de Morale* 6, pp. 1–13.

— (1900). "La Théorie de Lorentz et le Principe de la Réaction". In: *Archives Néerlandaises* 5, pp. 252–278.

— (1902). *La Science et l'Hypothèse.* Champs Sciences. Paris: Flammarion.

— (1905a). *La Valeur de la Science.* Paris: Flammarion.

— (1905b). "Sur la Dynamique de l'Électron". In: *Comptes rendus de l'Académie des Sciences* 140, pp. 1504–08.

— (1905c). "Sur la Dynamique de l'Électron". In: *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* 140, pp. 1504–1508.

— (1906). "Sur la Dynamique de l'Électron". In: *Rendiconti del Circolo matematico di Palermo* 21, pp. 129–75.

— (1908a). "La Dynamique de l'Électron". In: *Revue Générale des Sciences Pures et Appliquées* 19, pp. 386–402.

— (1908b). *Science et Méthode.* Paris: Flammarion.

— (1929). "The Value of Science". In: *The Foundations of Science: Science and Hypothesis, The Value of Science, Science and Method.* New York: The Science Press, pp. 199–355. Trans. of *La Valeur de la Science.* Paris: Flammarion, 1905.

— (1952). *Science and Hypothesis.* New York: Dover Publications. Trans. of *La Science et l'Hypothèse.* Champs Sciences. Paris: Flammarion, 1902.

— (2003). *Science and Method.* Mineola, New York: Dover Publications. Trans. of *Science et Méthode.* Paris: Flammarion, 1908.

— (2008). *The Theory of Lorentz and the Principle of Reaction.* Trans. by Steve Lawrence. URL: http://physicsinsights.org/poincare-1900.pdf. Trans. of "La Théorie de Lorentz et le Principe de la Réaction". In: *Archives Néerlandaises* 5 (1900), pp. 252–278.

— (2021). *On the Dynamics of the Electron.* URL: https://en.wikisource.org/wiki/Translation:On_the_Dynamics_of_the_Electron_(July). Trans. of

"Sur la Dynamique de l'Électron". In: *Rendiconti del Circolo matematico di Palermo* 21 (1906), pp. 129–75.

Psillos, Stathis (1996). "Poincaré's Concept of Mechanical Explanation". In: *Henri Poincaré: Science and Philosophy*. Ed. by Jean-Louis Greffe, Gerhard Heinzmann, and Kuno Lorenz. Berlin: Academie Verlag & Paris: Albert Blanchard, pp. 177–91.

Quine, W. V. (1951). "Two Dogmas of Empiricism". In: *Philosophical Review* 60, pp. 20–43.

— (1969). "Epistemology Naturalized". In: *Ontological Relativity and Other Essays*. Ed. by Willard van Orman Quine. Columbia University Press.

Reichenbach, Hans (1920). *Relativitätstheorie und Erkenntnis apriori*. Berlin.

— (1965). *The Theory of Relativity and A Priori Knowledge*. Berkeley. Trans. of *Relativitätstheorie und Erkenntnis apriori*. Berlin, 1920.

Rindler, Wolfgang (1982). *Introduction to Special Relativity*. Oxford: Clarendon Press.

— (2006). *Relativity: Special, General, and Cosmological*. 2nd ed. Oxford: Oxford University Press.

Rossi, Bruno and D. B. Hall (1941). "Variation of the Rate of Decay of Mesotrons with Momentum". In: *Physical Review* 59.3, pp. 223–228. DOI: 10.1103/PhysRev.59.223.

Ryckman, Thomas (2005). *The Reign of Relativity: Philosophy in Physics 1915-1925*. Oxford studies in philosophy of science. Oxford: Oxford Univ Press.

Shaffer, Michael J. and Michael L. Veber (2011). "Introduction". In: *What Place for the A Priori?* Ed. by Michael J. Shaffer and Michael L. Veber. Open Court, pp. 1–8.

Stump, David J. (2021). "The Relative A Priori". In: *The Routledge Handbook of Logical Empiricism*. Ed. by Thomas Uebel. Routledge, pp. 203–210.

Torretti, Roberto (1983). *Relativity and Geometry*. Oxford and New York: Pergamon.

Walter, Scott (2009). "Hypothesis and Convention in Poincaré's Defense of Galilei Space-Time". In: *The Significance of the Hypothetical in the Natural Sciences*. Ed. by Michael Heidelberger and Gregor Schiemann. De Gruyter, pp. 193–219.

— (2014). "Poincaré on Clocks in Motion". In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 47, pp. 131–141.

Zahar, Elie (1989). *Einstein's Revolution: A Study in Heuristic*. La Salle: Open Court.