On the logical structure of relativity theories*

by

H. Andréka J. X. Madarász I. Németi

July 5, 2002

With contributions from A. Andai, G. Sági, I. Sain, and Cs. Tőke.

Preliminary version; comments are welcome

*Research supported by CCSOM of Univ. Amsterdam and by the Hungarian National Foundation for scientific research grants No's T35192, T30314, T23234, F17452, T16448. Madarász's research was also supported by Bolyai College Budapest. This research was also supported by grant No 049.011.014 of the Netherlands Organization for Scientific Research.

Contents

1 Introduction

1.1 Broad introduction

This work grew out from a lecture notes for a course given March 1 – April 30, 1998 in CCSOM of the University of Amsterdam. The course description included below also serves as part of the present introduction.

Course description (modified to serve as a "broad introduction" for the present work)

This course is designed mainly for people familiar with logic who are interested in using logic for gaining a deeper understanding of (at least part of) reality. This includes e.g. people studying the methodology or philosophy of sciences, or people interested in the axiomatic [or logic based] approaches to relativity theory, or people interested in the relationship between the "Reichenbach-Grünbaum version" of relativity and the "standard version", or people familiar with modern logic who are interested in an *easily comprehensible*, "intuitively helpful" and at the same time *precise* introduction into the exciting areas of relativity and cosmology.

(I) Historical perspective and some of our goals. Tarski formalized geometry as a theory of first-order logic. The point here is to use *only* first-order logic; no external "devices" or tacit assumptions are allowed to enter the picture. Motivated by Tarski, P. Suppes [243] raised the problem of formalizing the theory of special relativity as a theory purely in first-order logic.[1] This problem was studied by Ax, Goldblatt and others. In the present work we want to work on a "programme" which is related to the just quoted one (e.g. insist on using only first-order logic)

[1] There are certain methodological reasons why we want to stick with (possibly many-sorted and perhaps modal) first-order logic (FOL) as opposed to using higher-order logic with its standard semantics. These reasons are connected with the fact that higher-order logic is not absolute in the set theoretic sense cf. e.g. Barwise-Feferman [43], e.g. p.33, below item 2.1.1, and §XVII.2.1., and therefore no effective complete proof system can exist for higher-order logic. Putting it more bluntly: There is *no completeness* theorem for higher-order logic, moreover it is *impossible* to obtain a completeness theorem for higher-order logic (this follows e.g. from Gödel's incompleteness theorem). The above mentioned reasons for sticking with FOL were presented at various logic conferences in Amsterdam (during the period 1994-1998) and can be (partially) recovered from Sain [232], cf. also Johan van Benthem's papers quoted in item (II) "Logical core" e.g. [267]. We collected and explained some of these reasons in the Appendix.

but is slightly more general (than the quoted one) in various respects, e.g. in the following. A possible approach to axiomatizing special relativity in first-order logic (FOL) would be the following: Axiomatize, first, Minkowskian geometry in FOL and then try to build a relativity theory on top of that. Here, we want to develop a different, more ambitious approach than this, namely primarily (or firstly) we want to write up a natural and convincing axiomatization, call it *Specrel*, of special relativity in FOL, and then we want to study and develop this first order theory *Specrel*, such that studying *Specrel* would lead us to "deriving" something like Minkowskian geometry as a "theoretical construct" (i.e. Minkowskian geometry will show up as a "theoretical consequence" of our "primary" theory *Specrel*). Some of our reasons for this preference are summarized in items (i)–(iii) below. (i) We want to start out with such axioms about the subject matter of special relativity (i.e. motion etc.) which are *self-evident* (in some sense). In other words, we would like to *derive* (in some sense) relativity theory from easily comprehensible, natural axioms which are convincing (and acceptable) even for the outsider (who does not know anything about relativity) and who would be reluctant to accept that the world works as described by e.g. Minkowskian geometry if he/she was *simply* told by some *authority* that this is the way "the world is". (Of course everybody accepts this but some people might accept it *only* as a *fact* and not as an explanation.) In other words, in the present work we seek *understanding* and insight as opposed to mere *knowledge*. (ii) We also want to push ahead in the direction of general relativity. (iii) Further essential reasons for our preference (for not starting with Minkowskian geometry) will be indicated in items (IV), (V) below.

(II) Searching for the logical core. In a few papers, Johan van Benthem elaborates the idea of separating out the "logical cores" of certain logics. The idea here is separating out the really essential part (from the logical point of view) from the whole "burden of mathematical machinery attached to the subject in the course of time" cf. e.g. Benthem [267] and also Sain [232]. A part of the present work can be considered as carrying out this "Benthemian" programme for (parts of) relativity theory.

(III) Searching for insight. All discussions will be in terms of simple concepts. When formalizing our (language and) axioms we will confine ourselves to a very plain language, using such easily comprehensible concepts as "bodies" or "observers". Whenever we need more complex concepts like "energy", "entropy" or "curvature of space-time", then we will first define these, as a logician would do, in terms of our plain language. This way we hope to gain real insight into *why* certain exotic predictions of relativity theory are "predicted". It also allows us to make the axioms with which we started subject to debate: both because of the plain language in which they are expressed and because of the purely logical nature of our reasoning.

(IV) Not "only the heart". Sometimes physical theories are formalized in the following style: Only the "heart" (in some sense) of the theory is formalized; and then the so obtained formal theory comes together with a *non-formalized*, natural language explanation of how to use the formal theory. This natural language text is often called the "interpretation" of the formal theory. An example for such an "only the heart" approach would be formalizing e.g. Minkowskian geometry in first-order logic and then writing an explanation in natural language on how to use Minkowskian geometry for solving problems in special relativity.

In the present work we intend to formalize the whole theory and *not* only the heart. In particular, we want to obtain a formalized theory which contains its own "interpretation" (where the word "interpretation" is used in the above sense).

(V) Conceptual analysis.[2] We also study variants of special relativity in the

[2] By *conceptual analysis* of a theory, like relativity, we understand the following activity. First one identifies the key concepts of the theory, and then using these concepts one formulates the key "principles" or axioms of the theory like e.g. the "principle of the speed of light". Then one *refines* these principles to sub-principles e.g. the principle, that "the speed of light is independent of the velocity of its source" cf. principle (P1) in Friedman [90, p.159]. (Eventually, these sub-principles need to be formalized as concrete formulas in the logical language of our theory.) Then one investigates the *logical relationship* between the so obtained sub-principles (i.e. refined axioms), and also the logical relationship between (the various combinations of) these sub-principles and the distinguished *theorems* (i.e. "predictions") of the theory in question. E.g. a variant, say (P1*), of (P1) says that for each observer the speed of light does not depend on its direction. (This is a kind of isotropy principle). As an example for conceptual analysis we will prove in §3.4, that, of the "speed of light" sub-principles, (P1*) is sufficient for proving the nonexistence of faster than light observers. Another example of conceptual analysis is a result Judit Madarász obtained by following suggestions and encouragement from Gyula Dávid (Dept. Gen. Physics ELTE Univ.) summarized as follows. She derived almost all of special relativity from natural, simple and convincing axioms *not involving* the speed of light or anything connected to electrodnamics. Chapter 5 (pp.704-769) of this work is devoted to elaborating this result and discussing its consequences. As a piece of motivation we note that in classical electrodynamics (P1) and (P1*) are valid but (P2) (saying that the speed of light is the same for all observers) is not necessarily valid, cf. e.g. Friedman [90]. (So, the quoted piece of conceptual analysis implies that the nonexistence of faster than light observers already follows from classical electrodynamics [while certain other parts of special relativity do not].)

Further examples of conceptual analysis, taken from pure mathematics, are the independence of the Continuum Hypothesis in set theory, or the theorem saying that the Banach-Tarski "ball decomposition paradox" is equivalent with the axiom of choice, or that Birkhoff's theorem (characterizing varieties of universal algebra) depends on the Axiom of Foundation. In general, what is known as "reverse mathematics" belongs to conceptual analysis (of set theory or category theory, depending on our choice of foundation for mathematics).

form of "competing" axiom systems formalized in first-order logic. The reason for having several versions for the theory, i.e. several axiom systems, is that this way we can study the *consequences* of the various axioms, enabling us to find out which axiom is responsible for some interesting or "exotic" prediction of relativity theory. Among others, this enables us to refine the *conceptual analysis* of relativity theory in Friedman [90] and Rindler [224], or compare the Reichenbach-Grünbaum approach to relativity (cf. L. E. Szabó [244] or [90]) with the standard one.

Such a conceptual analysis leads to the desire of making the axioms weaker and weaker (and at the same time more and more "convincing and natural") but such that (almost) all important predictions of the theory in question remain provable. It is surprising, how few axioms (i.e. assumptions) remain needed in the end for proving almost all the outstanding predictions. In this connection we note that (as we mentioned in footnote 2) following suggetions from Gyula Dávid[3], in Chapter 5 (due to Madarász) she proves that in the framework to be presented here one can avoid all the axioms involving the speed of light (i.e. photons) and still have a nice, convincing, finite axiom system from which almost all important predictions of special relativity are provable. Madarász's axiom system is purely kinematic, it does not mention anything related to e.g. electricity.[4] In our opinion, the process of so weakening/refining the axiom system(s) can lead to (i) improving the conceptual analysis discussed in Friedman [90] and Rindler [224] and (ii) improving our understanding of the theory in question and why/how it works.

(VI) Fruits of "logicization". After Tarski and his followers formalized geometry purely in FOL (first-order logic), two useful things happened (among others of course), as follows: (i) They applied the full machinery of mathematical logic (including e.g. the theory of definability) to study the new theory. (ii) They arrived at a *"hierarchy" of axiom systems*, i.e. theories of geometry.[5] (Then they studied this hierarchy.) Both developments led to new insights and proved useful.

Analogously, after we have formalized (parts of) relativity in FOL, we will carry through (in the present work) items (i) and (ii) above for relativity in place of geometry. E.g. we will ask logical questions concerning the logical structure of the theory, the number of non-elementarily-equivalent models, classification of models, definability issues, connections with Gödel's incompleteness proofs, etc. Among other things, we will use logic to find out which axioms are responsible for certain

[3] Cf. [71].

[4] We are mentioning this because of the related work Blészer-Gnädig-Varga [49] deriving relativity from pre-Maxwellian, simple laws of electrodynamics.

[5] It belongs to the very spirit of the "axiomatic method" to study instead of a single theory a *hierarchy* of its variants, alternatives, weaker versions etc. A typical example is the theory of arithmetic as presented e.g. in Hájek and Pudlák [119].

surprising predictions of relativity theory like e.g. "no observer can move faster than the speed of light", "the twin paradox" or issues concerning the possibility of time travel.

As was already indicated in footnote 1 on p.6, we have some methodological motivation for sticking with (possibly many-sorted) first-order logic. An outline of the relevant methodological/metalogical results and considerations is included as an appendix.

A further use of "logicising" relativity will be the following. We will be able to give qualitative answers to qualitative questions in a precise mathematical manner. No more will "quantitative" be a synonym for "precise" or "mathematical" (cf. e.g. Rohlich [226]). In analogy with the structuralist chapters (like e.g. category theory or algebra) of mathematics[6], one can develop a structuralist approach to relativity which can serve as a complement for the more traditional approach.

(VII) Temporal logic. Temporal and modal logics of relativity theory have been around in the literature cf. e.g. Goldblatt [106], Benthem [265], Belnap [46]. Our first-order theory of relativity can be used as a foundation for such temporal-modal logics. This is analogous with the situation in temporal logics of programs and actions. Namely, in Sain [231] a many-sorted first-order theory (analogous to our present first-order theory of relativity) was elaborated about the "world of actions and programs" and later it was used as a foundation for temporal logics of programs and actions in Andréka et al. [14] and in many other works e.g. by Pasztor, Csirmaz, Richter, Gergely and other researchers.

(VIII) Connections with Gödel's incompleteness theorems. Hawking, Weinberg and others suggested the possibility of a final Theory of Everything. In the literature it is often argued that Gödel's incompleteness theorem renders such theory impossible. This is a challenge for the logician. In more detail: Recently there has been an extensive debate, in the literature of relativity theory and related areas, concerning the connections between relativity (and its possible variants) and Gödel's incompleteness theorems.[7] These debates were triggered by the programme called searching for a "final theory" (or sometimes T.O.E.) proposed by Hawking, Weinberg and others. Cf. e.g. Hájek [118, p.291], Stöltzner [242], Dyson [76, p.53], Regge [221, p.296] for the critiques using Gödel's incompleteness theorem as a "weapon" against the "final theory". We investigate the issue and answer some questions in §3.8 (pp.294-346), §6.6.9 herein, in [16], in [17], in Chapter 7 of [19]. Cf. also the "laws of nature" part of Chapter 6.

---

[6] Cf. e.g. Michael Makkai [180, 181].

[7] There are several of these varying in strength. (Therefore there are several Gödel incompleteness properties of theories.)

(IX) On our choice of language (a return to the subject of "Searching for insight"). Let us return to item (II) entitled "Searching for insight". Our saying there that we will try to use "simpler" concepts whenever possible instead of "more complex" ones was a perhaps simplified way of referring to the following more subtle considerations:

The concepts potentially usable in scientific theories (such as e.g. relativity) have been partially ordered in the literature as being more observable (and less "theoretical") or less observable and more theoretical. (Here "observable" also means primary.) This **observable/theoretical** distinction, or rather hierarchy, is recalled from the literature (of relativity theory) in e.g. Friedman [90, pp.4–5]. This observable/theoretical hierarchy is not perfectly well defined and is known to be problematic, but as Friedman puts it, it is still better than nothing. E.g. the movements of the oceans called tides are more observable (or closer to be observable) than the pull of gravity of the Moon which, we think, is causing them. That is, the gravitational force field of a mass-point (like the moon) is a more theoretical concept than the movement of a body (e.g. ocean's shore). (Actually the gravitational force field might turn out to be a "wrong concept" and we may have to replace it with something else like the curvature of space-time. Probably the movement of the ocean's shore-line will be less questionable as a "something" which one can talk about.)

In this work we will formalize theories of relativity in (many-sorted) first-order logic. When formalizing a theory (in first-order logic), an important step is choosing the language (i.e. vocabulary) of the theory. Here, we will try to choose the basic concepts of our language as observational as possible; and will introduce more theoretical concepts (as definitions) at later stages, when development of the theory justifies them (c.f. e.g. §6.2, §6.9). Eventually, this process will lead to the introduction of a new theory with new basic concepts (new language). We find this a natural way of theory development / theory "understanding"/ theory analyzing.

As Friedman [90, p.4] points out, the observational/theoretical distinction is not an absolute one. E.g. what is an observational concept at a certain stage of theory development might turn out to be a theoretical one later. But to our minds, this seems to be in harmony with the modern approach of logic where theories are considered as dynamic objects (as opposed to the more classical "eternally frozen" idea of theories), cf. e.g. van Benthem [268] (cf. also Andréka-van Benthem-Németi [31], Andréka-Németi-Sain [28], Gärdenfors [97], Németi [203] and references in [268] for approaches to the dynamic trend in mathematical logic). Therefore, our theory of relativity will come in stages. In the first stage, we choose our vocabulary to be observational (relative to the state of being in the first stage). In a later stage, we will have enough results for revising of what we consider observational (in that stage) and accordingly we will revise our language (i.e. vocabulary). Then, the process continues in the same spirit. (There are of course even later stages and later revisions of language. The point here is that we will try to keep all such revisions well motivated by results obtained in earlier stages.) Cf. e.g. §6.2, §6.9.

(X) Addressing the "why"-type questions. In P. Davies & J. Gribbin [72] on pp.94–95 Paul Davies writes the following about his early studies of relativity: "As my education continued, I came to learn of the various predictions of relativity ... All these results I took to be true, but what they actually *meant* remained a puzzle to me. ... I had learned how to manipulate the formulae ... I could work out what would actually happen, but I had no understanding of *why* this should be so."

The above quotation illustrates that beside the authors of the present work there are other people (e.g. P.D. and J.G.) who believe that certain why-type questions are legitimate (we will try to explain below what kind of why-type questions we have in mind). It often happens that a student of relativity asks e.g. "why is the twin paradox true", or "why do moving spaceships shrink", or "why do clocks closer to a black hole run slower than clocks distant from the black hole".

A possible, *mathematically precise* answer is the following. The teacher declares that space-time is a manifold with certain properties, then he writes up Einstein's equations, then he declares that they are true (for the physical world) and then he derives the effect in question (e.g. the twin paradox) from these equations. Having received and *digested* this answer, several students feel that now they can calculate what will happen but that they still have not been told "*why* this should be so" (cf. the quotation from P.D. above). Putting it more boldly, the student feels that his *why-type question has not been addressed*.

Now, many of the science-popularizing books do address these why-type questions. But then, they rely on analogies, metaphors and their language is not mathematically precise. Therefore some of the above mentioned students still feel that their why-type question has not been answered with a sufficient *precision of logic*.

In the present work, we try (among others) to please this logically minded student. Namely, we will use the mathematically precise language of first-order logic, and in this framework we will try to address the above circumscribed why-type questions.[8] For this purpose, (i) we will keep (in the first part of this work) our vocabulary on the "observational" side of the observational/theoretical distinction discussed in item (IX) above, and (ii) *when proving a prediction* of relativity like e.g. the slowing down of moving clocks (i.e. time dilation), we will try to keep our axioms as few as possible, as simple and convincing as possible, and as weak as

[8]We did not define what precisely we mean by this particular kind of why-type questions, but we hope that the quotation from Davies & Gribbin [72] together with our discussion following it provides an "implicit definition" of the kind of why-type questions we are discussing.

possible. So, efforts will be made to keep the axioms both observational and simple; and to maintain a standard of discussing, analyzing and refining their intuitive meanings. At the same time theoretical concepts etc. will also be studied, but they will be postponed to the point where we will feel that we can tell the above outlined imaginary student *why* we introduce them.

We feel that a logical analysis of the theory in question (i.e. relativity), conducted purely in first-order logic (and conducted in the above outlined spirit), might be the right kind of framework for a logically precise and logically satisfying[9] approach to answering the why-type questions.

(XI) Prerequisites. Familiarity with first-order logic (FOL) is the most important prerequisite: some model theory of FOL, many-sorted FOL, e.g. Monk [197] or Enderton [82] provide all the background needed from logic. But cf. e.g. Ehrig-Mahr [79] for many-sorted universal algebra (and some many-sorted model theory). Some knowledge of linear algebra and fields would be useful, but not indispensable. Familiarity with the notions and terminology of universal algebra is also desirable. For universal algebra see e.g. [54], [129, Chapter 0], [192], [112].

[9] By "logically satisfying" we mean a style when one does not have to say things like "... well, this is not quite true, but if the laboratory is small enough and if we disregard this and that, then it will become almost true".

1.2 More specific introduction

About our motivation and aims:

As we already indicated in section 1.1 (I), some works of Tarski and his followers [131, 227, 237, 245, 246, 247, 251, 254] are devoted to the formalization of geometry as a theory of first-order logic. The point is to use *only* first-order logic, and no "external" devices or tacit assumptions are allowed. In analogy with the quoted works of Tarski et al, here we try (among others) to formalize relativity theory purely in first-order logic[10] (using nothing external, no tacit assumptions etc.). For further motivation and related work we refer to e.g. Ax [35], Buseman [56], Friedman [90], Goldblatt [108], Matolcsi [190], Mundy [199], Reichenbach [223], Schutz [236], Suppes [243]. For further, related logic-oriented approaches to axiomatizing relativity we refer to the references in the introduction of Schutz [236]. Reichenbach [223] is a rather important reference in this direction. In the list of references we include further related work. After formalizing the theory we also develop it to some extent and then use the formalized version to analyze the logical structure of the theory. First we concentrate on special relativity, then we move to accelerated observers, and then explore the possibilities of moving in the direction of general relativity.

An additional motivation is the following. First we quote from the book Matolcsi [190, p.11].

> "Mathematics reached a crisis at the end of the last century when a number of paradoxes came to light. Mathematicians surmounted the difficulties by revealing the origin of the troubles: the obscure notations, the inexact definitions; then the modern mathematical exactness was created and all the earlier notions and results were reappraised. After this great work nowadays mathematics is firmly based upon its exactness.
>
> Theoretical physics — in quantum field theory — reached its own crisis in the last decades. The reason of the troubles is the same. Earlier physics has treated common, visible and palpable phenomena, everything has been obvious."

[10]For some of the reasons why we want to stick with first-order logic, in addition to footnote 1 on p.6, we also refer to Ax [35], Mundy [199], da Costa et al. [66] and Sain [232], cf. also Barwise & Feferman [43] p.33 lines 15-26, i.e. immediately below Def.2.1.1, and the note on "absolute logics" on p.597 therein, and also §XVII.2 on p.609. As a simple example for these reasons we also note that there cannot exist a complete inference system (hence anything like Gödel's completeness theorem) for higher-order logics. Cf. also the Appendix.

...

> "It is quite evident, that we have to follow a way similar to that followed by mathematicians to create a firm theory based on mathematical exactness; having mathematical exactness as a guiding principle, we must reappraise physics, its most common, most visible and most palpable notions as well. Doing so we can hope we shall be able to overcome the difficulties."

Mathematics solved the above problem by using *logic*. Here we will experiment with doing the same in relativity theory, that is, build up (at least parts of) relativity theory in first-order logic.[11] Besides axiomatizing special relativity (and some of its distinguished variants), *purely* in first-order logic (and subjecting the so obtained first order theory to the usual logical and model-theoretic investigations), we do have more ambitious goals summarized in items (i)–(iii) below and in section 1.1 way above.

(i) We will also study the possibility of moving beyond special relativity in the direction of general relativity, and will look into having accelerated observers and having "partial" observers who do not "observe" all events that might be "observed" by other observers (cf. $\mathbf{Ax6_{00}}$ in section 3.3). Also,

(ii) We will look into possibilities for making our theory more general by making the axioms more flexible (i.e., weaker), and studying the number and structure of all complete theories extending our flexible theory.

(iii) Further, as was mentioned in section 1.1 (I), (III), (IV), (IX), to eliminate the need for an "interpretation" of the theory etc., our theory will have objects in it, which are not available in pure Minkowskian geometry. These are things like bodies, inertial bodies, photons, observers, etc. In this connection we note items (1–4) below: (1) If we identified Minkowskian geometry with special relativity, then this would yield an *uninterpreted* (in the physical sense) version of special relativity, while the first order theory which we develop here contains "its own interpretation", too. Cf. §1.1 (IV). (2) It is not clear to us how the conceptual analysis[12] suggested e.g. in [90] (or the Reichenbach-Grünbaum issues) could be squeezed into Minkowskian geometry. Cf. §1.1 (V). (3) Our formalized (special) relativity theory is undecidable, while the first order version of Minkowskian geometry in [108] is decidable, pointing in the direction that our theory is not reducible to pure Minkowskian geometry. (4) The observational/theoretical duality outlined in [90], cf. also section 1.1 (IX), motivates us to keep our vocabulary and axioms on the "observational" side (while

[11] The foundation of mathematics (i.e. axiomatic set theory) is also formalized in first-order logic.

[12] Which axiom is responsible for what, which axiom is intuitively more natural than the other, etc.

Minkowskian geometry remains more on the "theoretical" side).[13]

After having formalized relativity in first-order logic, one can use the well developed machinery of first-order logic for studying properties of the theory (e.g. the number of non-elementarily equivalent models, or its relationships with Gödel's incompleteness theorems, independence issues, definability questions etc).

As we said, we will have weaker versions and stronger versions of (formalized special) relativity. Then, we will see that *already* our weaker versions have the following interesting property. Let Th_0 be such a weaker version. Then Th_0 is *undecidable* and admits two natural, finitely axiomatizable extensions Th_1, Th_2 as follows.[14] Th_1 is hereditarily undecidable, moreover the conclusions of *both* of Gödel's *incompleteness* theorems hold for Th_1. As a contrast, Th_2 is decidable. These claims will be elaborated in Chapter 7.

***

So let's get started. We want to develop a kinematics.[15]

- What is kinematics?
- A theory of *motion*.
- What moves?
- Idealization: We assume that there are things called *bodies* (like "heavenly bodies") and they move.
- How do bodies move?
- Idealization: They change their (spatial) locations.
- What does change of location mean?
- At different *time instances* the same body has different *locations*.

OK, then there are time instances and locations involved (whatever they are). Let us fix that. Our paradigm says that time instances and locations are only relative to something which we will call *observers*.[16] So we assume that there are observers

[13] We feel that our basic concepts are more on the "observational" side and less on the "theoretical" side than those of, say, Minkowskian geometry. Agreement with this opinion is spelled out explicitly in Friedman [90, p.32, lines 15–19]. Accordingly, we will discuss Minkowskian geometry but it will be derived later, as a "theoretical property" of our more "observational" theory.

[14] To be more precise, Th_2 is only a finite-schema axiomatizable theory.

[15] For simplicity, we concentrate on kinematics of relativity, but by the same methods one can extend the investigations to, say, mechanics. A motivation for sticking with kinematics is that by using *only* kinematics we can prove things which are usually proved by using notions like e.g. mass.

[16] We use the expression "*observer*" in the sense of the physics book d'Inverno [75, pp.17,21]. So, in our sense, an observer "*coordinatizes*" the set of events and as we will later (in §6) call it, an observer coordinatizes what will be called there "space-time". Other books (e.g. Hraskó [139, p.32], Landau-Lifsic [158], Misner-Thorne-Wheeler [196, p.327]) use the expression "*reference frame*" for what we call observer. Still other books use a more abstract notion of observer such that for them

(special bodies). Given an observer m, time instance t and location s, observer m may "observe" a certain body b as being present at $\langle t, s\rangle$ while m may observe other bodies b_1 as not being present at $\langle t, s\rangle$. This simply means that from the point of view of m, b is present at location s at time t. We treat this concept of observing as primitive and denote it as $b \in w_m(t, s)$. That is, $w_m(t, s)$ is defined to be the set of bodies present at $\langle t, s\rangle$ from the point of view of m. We should emphasize that this kind of observing has (almost) nothing to do with the intuitive notion of observing in the form of, say, seeing optically.

- What are time instances t and locations s?

- Our first answer is that they are "labels" used by observers. But sooner or later we will have to be more specific. So let us see what t is.

We agree that, for an observer m, time instances are "quantities" like 100, 500, 1/2. To be faithful to the spirit of the axiomatic method, we *do not decide* what quantities are, we only postulate that they satisfy some simple axioms which in themselves are intuitively convincing. Namely, we assume that quantities form an ordered field $\mathfrak{F} = \langle F, +, \cdot, \leq\rangle$, that is, $\mathfrak{F}$ satisfies the usual axioms of ordered fields (to be recalled in section 2.1 from, e.g., [59]). The time scale of observer m is simply $\mathfrak{F}$ itself, the neutral element 0 of $\mathfrak{F}$ means "*now*", $t > 0$ represents "*future*" and $t < 0$ represents "*past*". For simplicity, we agree that locations s are represented by triplets of quantities $s = \langle s_1, s_2, s_3\rangle \in {}^3F$.

So far, we agreed on representing locations by triplets of quantities, or by triplets of "coordinates" from the field $\mathfrak{F}$. It is pairs $p = \langle t, s\rangle$ of time instances and locations for which we say that a body b occurs there (at $\langle t, s\rangle$) for observer m. We call such pairs points of our <u>*coordinate-system*</u> ${}^4\mathfrak{F}$, which we also denote by ${}^4F$. Therefore points of our coordinate-system are of the form $p = \langle p_0, p_1, p_2, p_3\rangle \in {}^4F$. We call p_0 the <u>*time coordinate*</u> and $\langle p_1, p_2, p_3\rangle$ the <u>*space coordinates*</u> of p.

Although our coordinate-system is four-dimensional, many of the ideas (and proofs) can be illustrated already in two or three dimensions. We will try to keep our presentation as simple as possible. Therefore we will sometimes *pretend* that our coordinate-system is 2-dimensional *but* we will go up to 3 or 4 dimensions as soon as the higher dimensional case would behave differently.

As we said, to each point $p \in {}^4F$ of our coordinate-system, an observer m

"reference frame" = "observer + coordinatization" becomes the case. For us, this is *only* a matter of choosing words, no issue of ideology is involved; and since we had to make a choice, we decided to follow d'Inverno's terminology where "observer" is basically the same as "reference frame". In passing we note that it is our impression that Einstein used the word "observer" in the same sense as d'Inverno does and we do, cf. [80, §9]. Cf. also Taylor-Wheeler [256, §I.4 (the definition of observers)].

associates a set $w_m(p)$ of bodies which, for m, are present at point p. Therefore, to each observer m, we associate a so called world-view function $w_m : {}^4F \longrightarrow \mathcal{P}(B)$ mapping our coordinate-system ${}^4F$ into the powerset $\mathcal{P}(B)$ of the set B of bodies. We call the elements of $\mathcal{P}(B)$ "*events*". Matolcsi [190] calls them occurrences. For us an event is nothing but information telling us which bodies are present and which are absent.[17] (This is why [190] calls them *occurrences*.) Therefore we can identify an event by a subset of B.

On terminology: Sometimes we might write sloppily *space-time* for our coordinate-system ${}^4\mathfrak{F}$. However we need to reserve the expression "space-time" for a similar but slightly different structure. Namely, in a later part of this work we will use the word space-time for a structure whose *elements* are the *events* (roughly, the universe of this structure is $\mathcal{P}(B)$) and whose *structure* will be induced by that of ${}^4\mathfrak{F}$ via the world-view functions $w_m : {}^4F \longrightarrow \mathcal{P}(B)$ belonging to the observers. Cf. the geometry chapter §6. In the *simplest* cases of special relativity, space-time will be *isomorphic* with our coordinate-system ${}^4\mathfrak{F}$. However, in order to be prepared for generalizations coming in the more advanced chapters of the present work, we need to treat space-time[18] as a structure strictly *different* from ${}^4\mathfrak{F}$.

Some connections with the literature. To our knowledge, the first attempt at a deductive treatment of relativity is due to Reichenbach [223], but we mention also Robb [225] which is earlier but which seems to be an "only the heart" approach. Although no explicit logical framework is present in [223], that work can be considered a second-order logic approach analogous with Hilbert's second-order logic discussion of Euclidean geometry in [133].[19] The requirement of using basic, observation-related terms as primitives is made explicit by Reichenbach in his general philosophy of natural sciences.

The project of a strictly deductive presentation of relativity theory can be compared to the similar development in Euclidean geometry. As well known, the first

[17] Misner & Thorne & Wheeler [196, p.6] (cf. Figure 1.2 therein) uses basically the same notion of an event as we do. They also give there detailed intuitive motivation for this definition of an event. For completeness, we note the following. In §6.9 ("On what we learned (so far) about choosing our first order language for relativity") of the present work we will arrive at a more abstract, more sophisticated notion of an "event" cf. item (402) on p.1210 and the explanation following it. The intuition behind that notion, however, is basically the same as the present one.

[18] Space-time will be a structure $\langle Mn, \ldots \rangle$ with $Mn \subseteq \mathcal{P}(B)$ the set of "observable" events. (In this connection, we note that e.g. Friedman [90, p.32, lines 4–5] defines space-time as "the set of … all actual and possible events".)

[19] Some definitions and axioms suggest in Reichenbach's work the impossibility of a first-order translation. Reichenbach did not aim at a first-order logic formalization.

comprehensive and rigorous treatment of geometry is due to Hilbert [133], which is taken as a second-order logic formalization.[20] Emphasizing the benefits of a first-order approach, Tarski formalized Euclidean geometry in first-order logic. Cf. e.g. [251], [237].

The first logic-oriented results related to relativity are due to Robb [225], who aimed at deriving the geometrical structure induced (in some sense) by the binary relation *being after* over events (in the sense indicated above). Despite the apparent similarity of Minkowskian geometry to Euclidean geometry, however, the absence of a comprehensive axiomatization allowing foundational and metamathematical discussions of the former is pointed out by Suppes [243], who proposes the idea of a first-order formalization of Minkowskian geometry. (He might also be interpreted as proposing a broader project of a first-order axiomatization of special relativity. The identification of special relativity with its theoretical core, Minkowskian geometry, is common in the literature. As we have already mentioned, we consider this identification as unfortunate.) Such a treatment of Minkowskian geometry was provided in turn by Goldblatt [108]. From the point of view of special relativity as a comprehensive physical theory, Goldblatt's study can be regarded as an "only the heart" approach. We mention also Schutz [236], whose axiomatization is of second-order, but is distinguished by the discussion of the independence of its axioms; Ax [35], who aims at deriving Minkowskian geometry from observational primitives similar to those in Reichenbach's approach; and Mundy [199], who presents a systematically simplified second-order axiomatization related to Robb's treatment. We should also mention Montague [198, §11] which represents a *model theoretical* (hence also logical) approach to physical theories of motion (Montague was a student of Tarski and became famous for successfully applying the methodology of model theory outside of pure mathematics.) The present list of references to related work is far from being complete. Further references can be found in the bibliographies of the works we quoted.

The question naturally arises: What is new in the present work (relative to the above references)?

A short answer is that we continue where our precursors stopped. More concretely, most of what we outlined in items (II)-(VI), (IX)-(X) on pp. 6-12 way above seem to be new (or almost new). To be more precise, the idea of starting theory building from the observational side (of the observational/theoretical distinction), sketched in item (IX), appears already in Reichenbach's work (but is not implemented there in first-order logic). The idea of restricting our tools strictly and consistently to (many-sorted) first-order logic is carried through in Goldblatt [108],

[20]For brevity we will sometimes write 'first-order approach' for 'first-order logic approach', and similarly for 'second-order' approach.

but he does not seem to go beyond the "only the heart" approach (cf. item (IV) in §1.1).

There seems to be a point where most of the above quoted authors seem to stop. This is, more or less, the following.[21] Roughly speaking, they write up axiom systems, then prove that the axiom systems have certain desirable properties.[22] But sooner or later they seem to stop. With some exaggeration one might say that in the present work the real fun begins after we have written up some suitable axiom systems and after we have proved that these have the desirable properties.

In connection with the above we would like to point out the following. If we want to do the logical analysis of a theory (which is not yet in logical form), say of special relativity, then the first step is to build an axiom system in the language of the logic we have chosen, which will be our "logicized" version of the theory in question. Then we prove that this "logicized" theory is indeed about the subject matter we wanted to analyze (and not about something else). Let us call this Step 2. However, it is only *after* Step 2 that we can really start applying the methods of mathematical logic to analyze the so obtained logic-based theory of whatever we wanted to study, e.g. of special relativity. In passing we note that during this analysis, among other things, we will probably experiment with changing the axioms, so e.g. we end up with having several concurrent logic-based versions of special relativity.[23] Cf. again

[21]We quoted so many works that it is hard to make categorical statements about them. Therefore what we write here is intended to be a "general impression" only, allowing exceptions etc. and is not a careful critical study of the literature.

[22]E.g. if the author's aim was to axiomatize Minkowskian geometry, then he proves, say, that every model of the axiom system is representable by a Minkowskian geometry over some real closed field, etc.

[23]An example for what we are saying is Tarski's and his followers' (Tarski et al.'s for short) first-order logic-based approach to geometry. They too begin with writing up axiom systems for geometry and proving so-called representation theorems (which prove that the axiom systems describe those mathematical structures which the authors wanted to study). This is what we called Step 2 above. Indeed, it is only *after* this Step 2 (and on the basis of Step 2) that the main bulk (the main results etc.) of the theory developed by Tarski et al. unfolds (or in other words, is developed). Further, this mathematical logic-based theory of geometry (initiated by Tarski et al.) is not finished or "closed down" even today; it is still under development; it still provides new and new insight into the original subject matter (and into related subjects).

Another example is provided by Tarski's theory of cylindric algebras. Tarski wrote up the axioms of this theory long time ago, and then he proved a representation theorem, saying that locally finite cylindric algebras are exactly those structures which he originally wanted to axiomatize, cf. [129, Part I]. This part of the theory could be written up and fully proved in not more than 50 pages. However, the main bulk of the theory of cylindric algebras came into existence *after* these Step 2-type results were obtained, and already in 1985 they filled two volumes, which together take up almost 1000 pages (cf. [129, Parts I, II]). Ever since then new and new results are added to the theory of cylindric algebras leading to deeper and deeper understanding of the subject matter for items (I)-(X) in §1.1.

On some of our aims and an outline of this work:

As already indicated, all our axioms will be formulas of first-order logic. We do not want to make our axioms generate a complete theory.[24] Our purpose is the opposite: we want to make our axioms as weak (and intuitively acceptable and convincing) as possible while still strong enough for proving interesting theorems of relativity theory.[25]

When introducing a new axiom, say **Ax**, we will investigate why **Ax** is plausible, why we (or the student) should believe in **Ax**, why we need it, and what would happen if we omitted it. This way we will obtain a relatively small set, called **Basax** (for basic axioms) of convincing (almost self-evident) axioms. **Basax** will be our *first* "possible" axiom system. Later, as a result of studying **Basax**, we will introduce and study a *hierarchy* of axiom systems (or of possible special relativity theories) in which hierarchy **Basax** will be neither the strongest nor the weakest theory. As we already said, we will investigate how many different complete theories $Th \supseteq$ **Basax** exist, which are possible consistent extensions of **Basax**. We will also attempt a structural description of the essentially different kinds of models of **Basax**.

In Chapter 2, we introduce and discuss **Basax**. We also study it there to some extent, e.g. we prove that **Basax** is consistent, that in dimension two it permits faster than light (FTL) observers, which in turn lead to time-travel like phenomena, and that the latter do *not* lead to logical paradoxes i.e. "**Basax** + there are FTL observers" is consistent (in dimension 2). In this chapter we also prove from **Basax** what we call the "paradigmatic effects" of relativity: moving clocks slow down, moving clocks get out of synchronism, and that moving meter-rods shrink. In section 2.8 we experiment with adding a strong symmetry axiom, **Ax(symm)**, to **Basax**. This

which Tarski initiated the study. (Very roughly, this subject matter can be summarized as the development and understanding of the theories of quantifier logics in a structuralist (or algebraic) perspective.)

[24] A theory T is called *complete* iff for every sentence φ in the language of T, exactly one of φ and $(\neg\varphi)$ follows from T.

[25] The situation is somewhat analogous with the difference between classical number theory studying the standard model $\underline{Z} = \langle Z, +, \cdot, -, 0, 1\rangle$ consisting of the set Z of integers, contrasted with, say, a part of abstract algebra, e.g. ring theory (or the theory of fields) where we study a broad class K of all rings of which $\underline{Z}$ is only a very special element. Sometime when we prove theorems about K, we say (or feel) that we understand more (or better) why that theorem is true for $\underline{Z}$. In this analogy, classical, standard special relativity is analogous with the complete theory of $\underline{Z}$ while the version we are describing here is analogous with the algebraic theory of K. (We note, however, that this analogy is imperfect, as often happens with analogies.)

symmetry axiom can be considered as an "instance" of Einstein's (special) principle of relativity. We will find that adding one, very natural and transparent axiom to **Basax** yields a theory which completely reproduces the usual, standard version of special relativity (i.e. the one based on Minkowskian geometry).[26] Actually this extra axiom will be **Ax(symm)**.

In Chapter 3, we continue studying **Basax**. Further, we introduce *refined versions* of **Basax** (e.g. **Newbasax**) and also *weaker* versions (e.g. **Bax**) of **Basax**. These weaker versions are used for the purposes listed in §1.1 e.g. for finding answers to the "why"-type questions. E.g. we will find that a surprisingly small fragment of **Basax** is sufficient for proving a theorem to the effect that there are no FTL observers in higher than 2 dimensions. We also analyse the *reason* for this no FTL theorem.[27] In the same chapter we study and classify the models of **Basax**. In §§ 3.8, 3.9 we return to studying the possible symmetry axioms (connected with Einstein's special principle of relativity) which can be added to **Basax** (and to its variants like **Bax**). This will lead us to a complete and decidable extension "**BaCo+Ax(rc)**" of **Basax** which describes the standard "textbook-version" of special relativity. In passing, we note that a difference between **Basax** and **Newbasax** is that in **Basax**-models the coordinate system ${}^4\mathfrak{F}$ is isomorphic with "space-time" while in **Newbasax** this is not necessarily the case (similarly to the situation in general relativity).

In Chapter 4, we continue refining our "generic" theories $\mathbf{Basax} + \mathbf{Ax(symm)}$ and **Basax**. We elaborate a *hierarchy* of theories ranging from the very weak $\mathbf{Bax}^{--}$ (in which practically no relativistic effect is provable) through stronger theories like e.g. "**Flexible-Basax**" to $\mathbf{Basax} + \mathbf{Ax(symm)}$ in which stages more and more of the relativistic effects will be provable. Cf. e.g. Figure 223 on p.653. Special sections are devoted to the Reichenbach-Grünbaum version of relativity (saying that we can measure only the two-way speed of light) and its connections with the Einsteinian version. The emphasis in Chapter 4 is on studying a hierarchy of theories (its internal dynamics etc.) as opposed to studying a single theory. For motivation for doing this we refer to §1.1, p.40, §3.4.2 and Figures 180 and 223 on pp.552, 653. In addition we note that §4 illustrates the "modularity" principle or "lego" principle of the axiomatic approach. By the latter we mean that §4 is intended to provide a dynamic picture of the possible (special) relativity theories, i.e. it concentrates on

[26] The connections between standard Minkowskian geometry and our more flexible (or more general) versions of relativity will be discussed in greater detail in Chapter 6, but cf. also §3.9 ("Symmetry axioms").

[27] We also investigate how one could avoid the no FTL result e.g. by changing or refining the axioms or changing the definition of an observer etc. These are done e.g. in §3.4.2.

"theory change"[28] via taking theories apart and putting them together differently and investigating the consequences (as opposed to studying a single theory in a "static" way).

In Chapter 5, we continue the process of weakening our speed of light axioms so much that eventually they completely disappear. The so obtained photon-free relativity theory **Relnoph** remains strong enough to prove most of the paradigmatic effects of usual relativity. We will see that **Relnoph** is a photon-free theory (or axiom system) which can be considered as an adequete axiomatization of special relativity (in the just indicated sense of proving most of the interesting theorems). Connections with ideas from non-standard analysis is also discussed (cf. also §4.1, Figure 133 on p.450).

In Chapter 6, we "discover" that there is an "observer independent" geometry sitting inside each model $\mathfrak{M}$ of, say, **Bax**. If $\mathfrak{M}$ is a model of the complete extension of **Basax** mentioned above (cf. §3.9), then this geometry agrees with the standard Minkowskian geometry. Further, we elaborate a so-called duality theory acting between the "world of certain kinds of geometries" on the one side, and the world of our observational-oriented models $\mathfrak{M}$ on the other side.

In Chapter 7 we investigate decidability questions (of our special relativity theories) and connections with Gödel's two incompleteness theorems. E.g. we find that for a natural extension, call it $\mathbf{Basax}^+$, of **Basax**, the conclusions of both Gödels's incompleteness theorems hold, e.g. the consistency $\mathbf{Con(Basax^+)}$ of $\mathbf{Basax}^+$ can be formalized in the language of $\mathbf{Basax}^+$ and is independent[29] of $\mathbf{Basax}^+$.

In Chapter 8 we study accelerated observers. To this end, we change (enrich) our first-order language for relativity and we refine and enrich our axiom system **Newbasax**.

How to read this work (interdependence of major parts): We tried to keep Chapters 1,2 easily readable and of an introductory character. (E.g. we tried to postpone more technical definitions or axioms to later parts.) The later chapters, i.e. Chapters 3,4 etc., were designed to be readable *independently* of each other, but they presuppose Chapter 2. So, after having read Chapter 2, the reader may continue with e.g. Chapter 4 or whatever chapter he may prefer. We tried to encourage independent reading of later chapters (after having read Chapter 2, of course) by including

[28] in the sense of the school of theory dynamics, theory change represented by e.g. [268], [95], [97], [124].

[29] I.e. is neither provable nor disprovable from $\mathbf{Basax}^+$.

an *Index of symbols and defined terms* and a *List of axioms and axiom systems*, cf. p.1253.

We plan to make this work accessible in the form of Parts I, II, III etc. Part I consists of Chapters 1,2 together with the Index, List of axioms and References. Except for Part I, each Part consists of one chapter. Further exception is the last Part which consists of appendices, various lists of definitions, index, references, etc. As we said, only Part I is a prerequisite for the others.

Acknowledgements: Chapters 3, 5, 6 of this work were written by J. Madarász. Chapters 1, 2 were written by Andréka, Madarász and Németi with contributions from Sain; Chapter 4 was written by Madarász with advisory help from Andréka and Németi. In these chapters (1, 2, 4) the roles of Andréka, Németi and Sain were of an advisory character: Chapters 1–6 will be used as material for Madarász's dissertation. The mathematical results in Chapters 3-6 were obtained by Madarász (and were checked for correctness by Andréka and Németi). Attila Andai read the material with the special purpose of comparing it with the generally accepted ideas and norms of contemporary physics. The contributions of Gábor Sági are in Chapter 8, while those of Ildikó Sain involve Chapter 1, Sections 2.1–2.4 and Chapter 8. Csaba Tőke's contributions involve Sections 3.9 and 3.10, and also he read the material with a similar eye as Attila Andai, i.e. with the eye of the physicist. Thanks go to Márta Fehér for checking this material from the point of view of philosophy of science. Andréka, Madarász and Németi are grateful to Attila Andai, Gábor Sági, Ildikó Sain and Csaba Tőke for comments, suggestions etc. concerning the whole material. They are also grateful to László Pólos and Sándor Vályi for various kinds of very useful help. Thanks also go to the participants and organizers of the 1998 Spring course at University of Amsterdam.

2 Special Relativity

Contents of Chapter 2

2.1 Frame language of relativity theory; world-view function

Some set theoretical notation and convention:

ω denotes the set of all natural numbers $\{0, 1, \ldots, n, \ldots\}$. We use von Neumann's concept of natural numbers, that is,

$0 \stackrel{\text{def}}{=} \emptyset$ ($\emptyset$ denotes the empty set) and

$n + 1 \stackrel{\text{def}}{=} n \cup \{n\} = \{0, \ldots, n\}$ for every $n \in \omega$. Therefore, in this spirit we will often write

$i \in n$ for $i < n$, where $i, n \in \omega$.

R denotes the set of all real numbers, and

$\mathfrak{R} = \langle \mathrm{R}, +, \cdot, \leq \rangle$ denotes the ordered field of real numbers (where $+, \cdot, \leq$ are the usual ones).

Z denotes the set of all integers.

For any set H, $\mathcal{P}(H)$ denotes the powerset of H, that is,

$\mathcal{P}(H) = \{ X : X \subseteq H \}$.

If R is a binary relation, i.e. set of (ordered) pairs, then $Dom(R)$ and $Rng(R)$ denote its domain and range, respectively. That is:

$Dom(R) \stackrel{\text{def}}{=} \{ a : \exists b \, \langle a, b \rangle \in R \}$ and

$Rng(R) \stackrel{\text{def}}{=} \{ b : \exists a \, \langle a, b \rangle \in R \}$.

A function is a binary relation f with the property that for each $x \in Dom(f)$ there is only one y such that $\langle x, y \rangle \in f$. As usual, $f(x)$ denotes this unique y.

$f : A \longrightarrow B$ or $A \xrightarrow{f} B$ denote that f is a function, $Dom(f) = A$ and $Rng(f) \subseteq B$.

For an arbitrary set H and $n \in \omega$, we often identify the set

${}^nH \stackrel{\text{def}}{=} \{ f : (f : n \longrightarrow H) \}$ with the Cartesian power

$\underbrace{H \times \ldots \times H}_{n\text{-times}} \stackrel{\text{def}}{=} \{ \langle h_0, \ldots, h_{n-1} \rangle : (\forall i < n) h_i \in H \}$. Thus, in particular,

${}^2H = H \times H$.

If R and S are binary relations, then their composition $R \circ S$ is defined as

$R \circ S \stackrel{\text{def}}{=} \{ \langle a, b \rangle : (\exists c)[\langle a, c \rangle \in R \ \wedge \ \langle c, b \rangle \in S] \}$.

Therefore, in particular if f and g are functions with $Rng(f) \subseteq Dom(g)$ then we write their *composition* the following way[30]:

$(f \circ g)(x) \stackrel{\text{def}}{=} g\Big(f(x)\Big)$ for every $x \in Dom(f)$.

[30]This is usually used in the reverse order in the literature.

For a binary relation R and a set X, the R-image $R[X]$ of X is defined as

$R[X] \stackrel{\text{def}}{=} \{ b \,:\, (\exists a \in X)\, \langle a, b\rangle \in R \}$.

Therefore in particular for a function f,

$f[X] = \{ f(x) \,:\, x \in Dom(f) \cap X \}$.

For a binary relation R, its inverse is

$R^{-1} \stackrel{\text{def}}{=} \{ \langle b, a\rangle \,:\, \langle a, b\rangle \in R \}$.

$\mathrm{Id}_A \stackrel{\text{def}}{=} \{ \langle x, x\rangle \,:\, x \in A \}$ is the *identity function* on A, for any set A. When A is understood from the context we will write Id in place of Id_A.

The following is a notation for defining functions. Let $expr(x)$ be an expression involving x, and let D be a set. Then

$\langle\, expr(x) \,:\, x \in D \,\rangle \stackrel{\text{def}}{=} \{ \langle x, expr(x)\rangle \,:\, x \in D \}$.

$f \restriction C \stackrel{\text{def}}{=} \{ \langle x, y\rangle \in f \,:\, x \in C \}$ is the restriction of the function f to the set C, for any function f and set C.

$\lhd$

Before giving the definition of our frame-language, we recall from [59] some of the standard notation and terminology used in (many-sorted) first-order logic.

By a *first-order language* we understand a language[31] of first-order logic. Similarly for *3-sorted first-order language* or *many-sorted first-order language*. We will often use the word *vocabulary* instead of a first-order language (to avoid ambiguity arising from the fact that "language" could also refer to the set of all formulas of some theory). A vocabulary is a collection of sort-symbols, relation-symbols and function-symbols.

In the present work we will use *many-sorted first-order logic*. We hope that the reader having some familiarity with one-sorted first-order logic will find the transition from one-sorted to many-sorted easy to make. Indeed, throughout the literature it is emphasized that many-sorted (first-order) logic is only a convenient "notational dialect" of one-sorted first-order logic and that anyone familiar with the one-sorted version will easily understand the many-sorted version without studying it separately.

By *many-sorted logic* we understand the many-sorted version of first-order logic. I.e. for brevity, we will omit the adjective "first-order" (so in this work many-sorted automatically implies first-order). Many-sorted logic is so close to one-sorted first-order logic, that most logic books study and discuss the one-sorted case first and

[31] Let us recall from the literature of logic that a *language of (many-sorted) first-order logic* or a "*vocabulary*" or a "*similarity type*" are different names for the same thing. The details can be found in any logic book e.g. in Monk [197, p.14] or Enderton [82]. Cf. §6.3 for more information on this.

then they formulate the generalization to the many-sorted case as an *exercise* left to the reader. Of course, for this exercise they explain how many-sorted logic can be reduced to the one-sorted case. (The fact that here we allow finitely many sorts only makes this reduction easier and "stronger", cf. footnote 37 on p.29.)

For an introduction to many-sorted logic and for its reduction to one-sorted first-order logic we refer to almost any logic book, e.g. to Enderton [82, §4.3, pp.277-281 (but the whole of §4 will be useful later)] or Manzano [185] or Monk [197]. We note that the whole book Meinke-Tucker [193] is devoted to many-sorted logic and its connections with higher-order logic. For completeness, we note that further useful information on this subject is available in the book Barwise & Feferman [43] on pp.25-27, pp.33-34, and item 7.1.2 (p.68).[32] We would like to reassure the reader that for understanding the present work, looking into [43] is not a prerequisite. (At a second reading of the present work, the just quoted parts of [43] might improve appreciation of certain "fine details".) Looking into [43] might also help seeing the connections of our approach with second-order logic. At this point we would like to emphasize that throughout the present work we are staying <u>*strictly within first-order*</u> logic.[33]

Let Fm and M denote, respectively, the set of all formulas and the class of all models of an arbitrary first-order language.

Then $\models (\subseteq \mathsf{M} \times Fm)$ denotes the *validity relation* of this language. We extend $\models$ to $\mathcal{P}(\mathsf{M}) \times \mathcal{P}(Fm)$ the usual way: Let $\mathsf{K} \subseteq \mathsf{M}$ and $\Sigma \subseteq Fm$. Then

$$\mathsf{K} \models \Sigma \ \text{ iff } \ (\forall \mathfrak{M} \in \mathsf{K})(\forall \varphi \in \Sigma)\mathfrak{M} \models \varphi\,.$$

We will write $\mathsf{K} \models \varphi$ in place of $\mathsf{K} \models \{\varphi\}$ and $\mathfrak{M} \models \Sigma$ when $\mathsf{K} = \{\mathfrak{M}\}$.

$$\mathsf{Th}(\mathsf{K}) \stackrel{\text{def}}{=} \{\, \varphi \in Fm \ : \ \mathsf{K} \models \varphi \,\}$$

is the *theory* of K, and

$$\mathsf{Mod}(\Sigma) \stackrel{\text{def}}{=} \{\, \mathfrak{M} \in \mathsf{M} \ : \ \mathfrak{M} \models \Sigma \,\}$$

is the *class of all models* of Σ. Let $\varphi \in Fm$. Then we say that φ is a *semantical consequence* of Σ, in symbols $\Sigma \models \varphi$, iff $\mathsf{Mod}(\Sigma) \models \varphi$.

$$\mathsf{Th}(\mathfrak{M}) \stackrel{\text{def}}{=} \mathsf{Th}(\{\mathfrak{M}\})$$

[32] For that book page numbers are important because it has no index and is 893 pages long.

[33] If at some point the reader would have the impression that this is not the case, then there must be a misunderstanding at that point which needs to be clarified. (This is so because the issue of our staying within first-order logic [or in one of its equivalent forms] is an important one from our methodological point of view. For reasons see the appendix.) Cf. Enderton [82, §4] or Manzano [185] for reducing higher-order logic to *first-order* many-sorted one.

is the (first-order) theory of the model $\mathfrak{M}$.

We will start our formal exposition of relativity theory with fixing a 3-sorted first-order language. We will call this language the *frame-language of relativity theory*.[34] We will use this language for formulating our first axiom systems for special relativity (this way producing our first formalized versions of the theory).[35]

In this chapter we introduce a relatively rich language because we want to use this language throughout the present work. At the beginning, and especially throughout Part I, we could have used a much simpler language, e.g. the one introduced in [16]. More specifically, in part I (and in chapters 1-5) we will not really need G, E , *Ib* introduced in Def.2.1.1 below.

Definition 2.1.1 (frame-language of relativity theory)
Let B, Q and G denote three sorts called *bodies*, *quantities* and *lines* or *geometry*, respectively. Let a natural number $n > 1$ be fixed.[36] Intuitively, n will be the dimension of our "space-time".

We are defining a first-order language with sorts[37] B, Q, G by first defining its models, as follows. $\mathfrak{M}$ is a *model (of dimension n) of this language* iff

$$\mathfrak{M} = \langle B^{\mathfrak{M}}, F^{\mathfrak{M}}, G^{\mathfrak{M}};\ Obs^{\mathfrak{M}}, Ph^{\mathfrak{M}}, Ib^{\mathfrak{M}}, +^{\mathfrak{M}}, \cdot^{\mathfrak{M}}, \leq^{\mathfrak{M}}, \mathrm{E}^{\mathfrak{M}}, W^{\mathfrak{M}}\rangle\,,$$

also denoted as

$$\mathfrak{M} = \langle B, F, G;\ Obs, Ph, Ib, +, \cdot, \leq, \mathrm{E}, W\rangle$$

for brevity[38], where:

- B is a nonempty set, it is $\mathfrak{M}$'s universe of sort B. B is called the set of *bodies* (of $\mathfrak{M}$).

[34] Later we will expand our frame-language with e.g. a kind of *pseudo-metric* $d : {}^nF \times {}^nF \longrightarrow F$, also called *distance*, see §8.1. Our choice of language will be re-considered in §6.9 ("On what we learned (so far) about choosing our first order language for relativity").

[35] Because of the purposes explained on p.21 ("On ... aims ... of this work") in later chapters we develop several axiom systems.

[36] We will be interested only in the case $n \in \{2,3,4\}$, but we give definitions and lemmas for arbitrary n if this does not cost any extra effort.

[37] Many-sorted logic is known to be reducible to one-sorted logic the following way (cf. Monk [197], Enderton [82]): One uses the union $\mathrm{B} \cup \mathrm{Q} \cup \mathrm{G}$ of the universes of the sorts of the many-sorted model as the universe of our new one-sorted model and one calls B, Q, G unary predicates.

[38] As is usual in logic, B, F, G, Obs etc. are *symbols* (sort symbols and relation symbols) of the language of $\mathfrak{M}$ while $B^{\mathfrak{M}}, \ldots, Obs^{\mathfrak{M}}$ etc. are *objects denoted* by these symbols according to the model $\mathfrak{M}$. *If* and where there is no danger of confusion, we will identify the symbols with the objects they denote (hence we write B for $B^{\mathfrak{M}}$ etc).

- F is $\mathfrak{M}$'s universe of sort Q. Intuitively, F serves both to be our "*time scale*" and "*space scale*". Relations
$+, \cdot, \leq$ are of sort Q, hence $\langle F, +, \cdot, \leq \rangle$ forms a structure. We will assume that $\mathfrak{F} := \langle F, +, \cdot, \leq \rangle$ is a linearly ordered field.[39] That is, the following set $\mathbf{Ax}_{\mathrm{OF}}$ of axioms is satisfied by $\langle F, +, \cdot, \leq \rangle$.

  $\mathbf{F} := \langle F, +, \cdot \rangle$ is a field[40]
  $\langle F, \leq \rangle$ is a linear order, and for every $a, c \in F$,
  $a \leq c \;\Rightarrow\; (\forall d \in F)(a + d \leq c + d)$ and
  $(a \leq c \text{ and } d > 0) \;\Rightarrow\; (d \cdot a \leq d \cdot c)$ hold.

  0 and 1 denote the usual zero and unit elements of the field. Further, for every $a \in F$, $|a|$ denotes the absolute value of a, that is,
  $|a| \stackrel{\text{def}}{=} \max\{a, -a\}$ (where "$-$" is the usual group theoretic inverse operation determined by $+$).

  We will denote the ordered field $\langle F, +, \cdot, \leq \rangle$ by $\mathfrak{F}$ and its field reduct $\langle F, +, \cdot \rangle$ by $\mathbf{F}$. Often we write $\mathfrak{F}^{\mathfrak{M}}$ for $\mathfrak{F}$ ($\mathbf{F}^{\mathfrak{M}}$ for $\mathbf{F}$) when we want to indicate explicitly that we look at $\mathfrak{F}$ ($\mathbf{F}$) as the "quantity part" of $\mathfrak{M}$. $\mathfrak{F}^{\mathfrak{M}}$ is called the *ordered field reduct* of $\mathfrak{M}$, following the standard notation and terminology of many-sorted model theory. We note that every linearly ordered field is infinite.

- G is a nonempty set, it is $\mathfrak{M}$'s universe of sort G. G is called (the set of) *lines* (or geometry, but geometry will be used in Chapter 6 in a slightly different and more comprehensive sense).[41] Intuitively, lines represent motion (in the form of "life-lines") of inertial bodies.

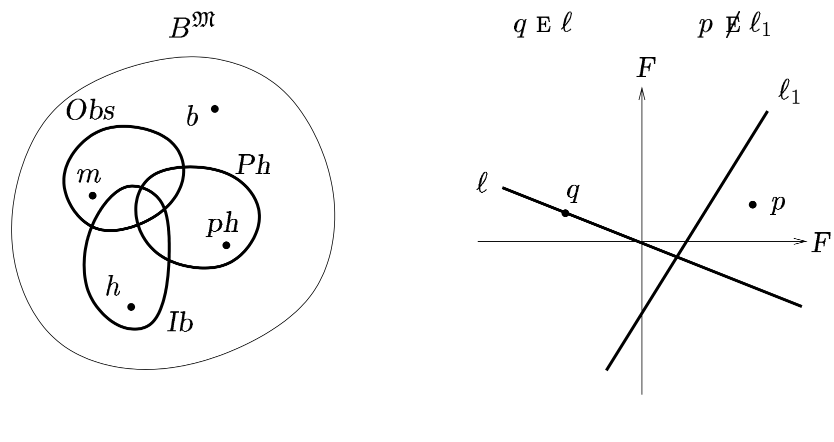
- $Obs, Ph, Ib \subseteq B$ are unary relations (of sort B). Their names are: set of

[39]This is why the universe of sort Q of $\mathfrak{M}$ is denoted by $F^{\mathfrak{M}}$ instead of $Q^{\mathfrak{M}}$. Occasionally we may refer to sort Q as sort F or as the field-sort of $\mathfrak{M}$. (Since in standard mathematical practice Q often denotes the field of rationals, there is a potential danger for ambiguity here for which we apologize to the reader. Anyway, we will *not* use Q to denote the rationals.)

[40]For completeness, we recall here the definition of a field. $\langle F, +, \cdot \rangle$ is called a *field* iff
$\langle F, + \rangle$ is a commutative group, we let 0 denote its neutral element;
$\langle F \setminus \{0\}, \cdot \rangle$ is a commutative group, we let 1 denote its neutral element;
$\cdot$ distributes over $+$, that is, $a \cdot (c + d) = a \cdot c + a \cdot d$ holds for every $a, c, d \in F$.
Sometimes we think of a field as a structure $\mathbf{F} = \langle F, +, \cdot, -, 0, 1 \rangle$, we hope this will cause no confusion. (We omitted 0, 1 and "$-$" from the original definition because they are first-order definable from $+$ and "$\cdot$". One thing that can be slightly influenced by this omission is the set of homomorphisms between two fields.)

[41]So the acronym G refers to geometry, but to avoid misunderstandings in Chapter 6, we pronounce it simply as "lines".

observers, set of *photons*, and set of *inertial bodies* (or *lonely bodies*), respectively. See the left-hand side of Figure 1.



Bodies of $\mathfrak{M}$ Quantities and lines of $\mathfrak{M}$

Figure 1: Bodies, quantities and lines of a model $\mathfrak{M}$.

- $\mathrm{E} \subseteq {}^nF \times G$ is an $(n+1)$-ary relation of sort $\langle \mathrm{Q}, \dots, \mathrm{Q}, \mathrm{G} \rangle$. Intuitively, for $p = \langle p_0, \dots, p_{n-1} \rangle \in {}^nF$ and $\ell \in G$, $\mathrm{E}\,(p_0, \dots, p_{n-1}, \ell)$ expresses that the point $p \in {}^nF$ is on the line ℓ. If p and ℓ are as above, we abbreviate $\mathrm{E}\,(p_0, \dots, p_{n-1}, \ell)$ by $p \,\mathrm{E}\, \ell$. We postulate axiom $\mathbf{Ax}_{\mathrm{G}}$ below, called the axiom of *extensionality* of lines.

  $$\mathbf{Ax}_{\mathrm{G}} \quad (\forall \ell_1, \ell_2 \in G)\Big((\forall p \in {}^nF)(p \,\mathrm{E}\, \ell_1 \;\Leftrightarrow\; p \,\mathrm{E}\, \ell_2) \;\Rightarrow\; \ell_1 = \ell_2 \Big).$$

  Here we note that the axiom of extensionality allows us to identify $\ell \in \mathrm{G}$ with a subset of ${}^nF$. (See the right-hand side of Figure 1.) Indeed, *we will identify* ℓ with the set $\{\, p \in {}^nF \,:\, p \,\mathrm{E}\, \ell \,\}$ (which is sometimes called the extension of ℓ). By this identification we may assume that $G \subseteq \mathcal{P}({}^nF)$ and E is the real "element-of" relation, $\in$. We will do this from now on, cf. Convention 2.1.2 (p.35).[42]

[42] This is a standard technique for handling higher-order objects of a logic.

- Let $p = \langle p_0, \dots, p_{n-1}\rangle \in {}^nF$. Then, p_0 is called the *time component* of p, while $\langle p_1, \dots, p_{n-1}\rangle$ is the *space component* of p.[43] Often we write p_t, p_x, p_y, p_z for p_0, p_1, p_2, p_3 respectively.

  ${}^nF$ is called the *coordinate-system* of $\mathfrak{M}$. We refer to p as a *point* or a (space-time) *location*. $\langle p_1, \dots, p_{n-1}\rangle$ is a *(space) location*. We will use the word location ambiguously.

- $W \subseteq B \times {}^nF \times B$, that is, W is an $n+2$-ary relation of sort $\langle \mathrm{B}, \underbrace{\mathrm{Q}, \dots, \mathrm{Q}}_{n\text{-times}}, \mathrm{B}\rangle$.

  W is called the *world-view relation* (of $\mathfrak{M}$). The most important part of our model is this relation. Intuitively, for n=4, $W(m,t,x,y,z,b)$ means that the observer m "observes" or "sees"[44] the body b at time t at (space) location $\langle x, y, z\rangle$. From the $(n+2)$-ary relation W and arbitrary observer $m \in Obs$ we define the *world-view function* $w_m : {}^nF \longrightarrow \mathcal{P}(B)$ as follows:

  $$w_m(p) \stackrel{\text{def}}{=} \{\, b \in B \;:\; W(m,p,b)\,\} \qquad \text{for every } p \in {}^nF\,,$$

  see Figure 2.

  For $p \in {}^nF$, we call the set $w_m(p)$ of bodies the *event* "happening" at location p as seen by m.[45] Intuitively, w_m defines the "subjective reality" of m. That is, w_m tells us how observer m "arranges" the events (elements of $\mathcal{P}(B)$) in the coordinate-system ${}^nF$; in other words, w_m tells us how m *"coordinatizes"* the set of events $\mathcal{P}(B)$. See Figure 3.

---

[43] It is important to emphasize here that ${}^nF$ is only the coordinate system of $\mathfrak{M}$ as opposed to being say "space-time" itself of $\mathfrak{M}$. Space-time will not be one of our primitive (i.e. basic) concepts, instead, it will be a derived "theoretical" concept and it will appear e.g. in §6. Cf. the observational/theoretical duality in §1.1, on p.11.

[44] We want to emphasize that here "observing" or "seeing" has nothing to do with the intuitive notion of observing in the form of measurement, or with the everyday notion of seeing via photons. In the present text, "observer" and "observing" are technical expressions which we use for historical reasons. Our "observing" is really a kind of coordinatizing, i.e. when we say that observer m observes event e at coordinates t, x, y, z, we mean only to say that m associates coordinates t, x, y, z to event e. (As opposed to "real observing", this is a very abstract act only.) By the word "observer" we mean what is sometimes called *frame of reference* or "system of reference" (or coordinate-system), cf. Remark 2.2.5 (p.54).

[45] Two or more bodies occupying the same space at the same time might contradict the physical intuition. However, presently we abstract away from the sizes of the bodies and therefore we permit two or more bodies to be at the same place at the same time. We also note the following. The reader may ask "why is an event a set of bodies". Motivation for this definition of an event can be found e.g. in Misner-Thorne-Wheeler [196, p.6], and Friedman [90, p.31] starting with line 9 therein.

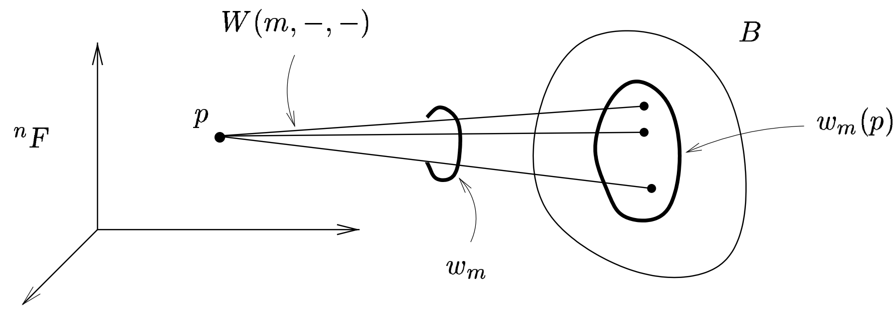


Figure 2: The world-view function w_m.

In the literature sometimes ^nF is called space-time, and sometimes the set of events $\mathcal{P}(B)$ is called space-time. The reason for calling $\mathcal{P}(B)$ space-time is that ^nF is only a coordinate-system (consisting of labels), using which observers coordinatize the set of events $\mathcal{P}(B)$.[46] On the long run it will be more fruitful to use the word space-time for the thing which is being coordinatized, that is for $\mathcal{P}(B)$.[47] We will see more reasons for calling the set of events space-time in the geometry chapter §6, pp.770–1169. The sets $B^{\mathfrak{M}}, F^{\mathfrak{M}}, G^{\mathfrak{M}}$ are also called the *universes* of $\mathfrak{M}$ (of sorts B, Q, G respectively).

Summing up: The similarity type of our first-order language consists of

– the sort symbols B, Q, G;

– the unary relation symbols Obs, Ph, Ib (most often, their interpretations in models are denoted by Obs, Ph, Ib as well);

– the symbols $+, \cdot, \leq$ of the ordered field $\mathfrak{F}$ (the neutral elements 0 and 1 of $+$ and $\cdot$, respectively, and "$-$" will also be treated as basic symbols);

– the (n+1)-ary relation symbol E, which we will systematically replace by the set theoretic "$\in$" (cf. Convention 2.1.2);

– the (n+2)-ary relation symbol W. Further:

The reduct $\langle B, Obs, Ph, Ib\rangle$ of $\mathfrak{M}$ is purely of sort B (body);

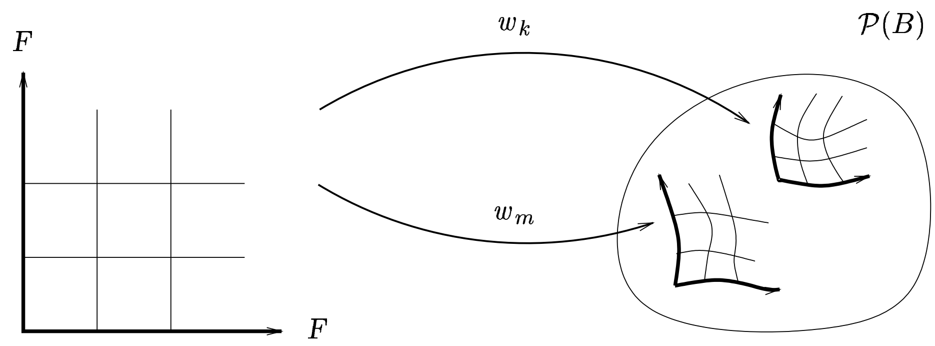
[46] A location $p \in {}^nF$ functions only as an "address" or "label" used by an observer m in labeling those events which exist for m.

[47] To help the reader's intuition we note that the world-view function w_m connects up the coordinate-system ^nF with the set of events $\mathcal{P}(B)$. Therefore if for someone it were easier to imagine ^nF as space-time then he/she can use the world-view function w_m to translate his/her intuition for viewing set of events as space-time.

$\mathfrak{F} = \langle F, +, \cdot, \leq \rangle = \langle \mathbf{F}, \leq \rangle$ is purely of sort Q (quantities);

G is the universe of sort G (lines), and there are no relation or function symbols which would be purely of sort G.

E (which we will replace by $\in$) acts between sorts Q and G, while W involves B and Q.



coordinate-system (^nF)

world-view functions ($w_m \ : \ m \in Obs$)

set of events or space-time (subsets of B)

The heart of our model is W, which is represented by functions $w_m : {}^nF \to \mathcal{P}(B)$ for each $m \in Obs$.

Figure 3: This is the way one should visualize a model $\mathfrak{M}$.

Variables ranging over the universes B, F, G of $\mathfrak{M}$ are most often chosen as follows. For arbitrary $i \in \omega$,

$b, b_i, h, h_i, k, k_i, m, m_i, ph, ph_i \in B$;
$a, a_i, c, c_i, d, d_i, t, t_i, x, x_i, y, y_i, z, z_i, \varepsilon, \lambda, \eta \in F$;
$\ell, \ell_i \in G$.

Let us recall that
$\mathbf{Ax}_{\mathrm{OF}} \cup \{\mathbf{Ax}_{\mathrm{G}}\} =$
$\{$the axioms postulating that $\mathfrak{F}$ is a linearly ordered field, axiom of extensionality$\}$. Now the <u>*frame-language*</u> of relativity theory of dimension n is defined to be the 3-sorted first-order language built up from the above symbols the usual way. A model

$\mathfrak{M} = \langle B, F, G; Obs, Ph, Ib, +, \cdot, \leq, \mathrm{E}, W\rangle$ is called a *frame model* (of relativity theory, of dimension n) iff

$$\mathfrak{M} \models \mathbf{Ax}_{\mathrm{OF}} \cup \{\mathbf{Ax}_{\mathrm{G}}\} \cup \{W(m,p,b) \to Obs(m)\}\,.$$ [48]

We denote the class of all frame models by FM. We call $\mathbf{Ax}_{\mathrm{OF}} \cup \{\mathbf{Ax}_{\mathrm{G}}\} \cup \{W(m,p,b) \to Obs(m)\}$ the *frame theory of special relativity theory* (or *frame theory* for short). By $\models^{\mathrm{OFG}}$ we denote semantical consequence within our present frame theory $\mathbf{Ax}_{\mathrm{OF}} \cup \{\mathbf{Ax}_{\mathrm{G}}\} \cup \{W(m,p,b) \to Obs(m)\}$. That is, for two sets Σ and Γ of formulas in our frame language,

$$\Sigma \models^{\mathrm{OFG}} \Gamma \iff (\forall \mathfrak{M} \in \mathsf{FM})(\mathfrak{M} \models \Sigma \ \Rightarrow\ \mathfrak{M} \models \Gamma)\,.$$

Also we define

$$\mathsf{Mod}_{\mathrm{OFG}}(\Sigma) \stackrel{\mathrm{def}}{=} \mathsf{FM} \cap \mathsf{Mod}(\Sigma)\,.$$

For brevity, throughout this work, we will write $\mathsf{Mod}(\Sigma)$ for $\mathsf{Mod}_{\mathrm{OFG}}(\Sigma)$. We hope that this causes no confusion, since we never want to talk about models (of type of our frame language) in which $\mathbf{Ax}_{\mathrm{OF}}$, $\mathbf{Ax}_{\mathrm{G}}$, or $(W(m,p,b) \to Obs(m))$ would fail.

Similarly, throughout we denote $\models^{\mathrm{OFG}}$ simply by $\models$, and we will never use $\models$ in its original purely logical sense *in the context* of our frame language (to avoid misunderstanding). Of course, when talking about structures or formulas of a different similarity type like e.g. $\mathfrak{F}$, then we use "$\models$" in its usual logical sense.

END OF DEF.2.1.1 (FRAME LANGUAGE).

$\lhd$

CONVENTION 2.1.2 As we indicated on p.31, below the definition of $\mathbf{Ax}_{\mathrm{G}}$, we will identify our E with the set theoretic membership relation "$\in$". As it was indicated there, this causes no loss of generality because every frame model $\mathfrak{M}$ is isomorphic to a frame model $\mathfrak{N}$ such that $\mathrm{E}^{\mathfrak{N}}$ coincides with the set theoretic "$\in$". Therefore throughout the rest of this work a frame model is of the form

$$\mathfrak{M} = \langle B, F, G;\ Obs, Ph, Ib, +, \cdot, \leq, \in, W\rangle\,.$$

[48] We use the standard convention from logic that an axiom $\varphi(x)$ automatically means its universal closure $\forall x\, \varphi(x)$. Throughout we write p for $p_0, \ldots, p_{n-1}$, hence $W(m,p,b)$ abbreviates $W(m, p_0 \ldots, p_{n-1}, b)$.

Throughout, we use the semicolon ";" to separate the sorts of a model from its relations and functions, as in the above equality. Often, we will use the more concise notation

$$\mathfrak{M} = \langle (B;\ Obs, Ph, Ib), \mathfrak{F}, G;\ \in, W \rangle\ .$$

If we want to indicate that a universe (or sort) like B or a relation like W comes from a particular model $\mathfrak{M}$, then we use the superscript $B^{\mathfrak{M}}, W^{\mathfrak{M}}$ respectively. This is why on p.29 we wrote $\mathfrak{M} = \langle B^{\mathfrak{M}}, \ldots, Obs^{\mathfrak{M}}, \ldots, W^{\mathfrak{M}} \rangle$. However, if $\mathfrak{M}$ is understood from context, then we will usually omit the superscript. All this ($B^{\mathfrak{M}}$ etc.) is standard notation from model theory and universal algebra, cf. e.g. Hodges [136], Monk [197], Grätzer [112], McKenzie&McNulty&Taylor [192], [129], Barwise-Feferman [43, p.27]. (As an exception, Chang-Keisler [59] uses lower indices like $B_{\mathfrak{M}}$ instead of $B^{\mathfrak{M}}$. But the general style and notational philosophy remains the same in [59], too, as adopted here.) $\triangleleft$

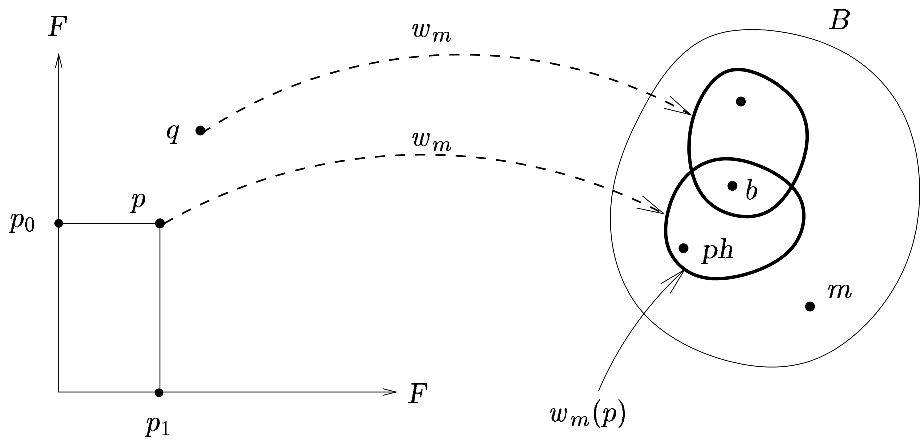
As we said, intuitively, n is the dimension of our space-time. If $n = 2$, then we have one time-dimension, and one space-dimension, i.e. space is one-dimensional. If $n = 3$, then space is two-dimensional, and $n = 4$ represents our usual 4-dimensional space-time, i.e. space is three-dimensional. In the case of $n = 2$ it is rather easy to illustrate things, so we will often use $n = 2$ in our drawings. When $n = 3$, one still can illustrate ideas by drawings quite well. Many ideas can be better seen in the case $n = 2$ and work completely analogously for arbitrary n. Some statements, however, are true for $n = 2$ and not true for $n = 3, 4$. In these cases we will emphasize that $n = 2, n = 3$, or $n = 4$. (Sometimes [but not frequently], the cases of 3 and 4 behave differently. In such cases, of course, one emphasizes this difference. But most of the time, for understanding the key ideas, we will concentrate on the case of $n = 3$.)[49] There is another reason why it may be useful to allow the dimension of space-time to vary. Later we may devise models in which not all observers coordinatize events with the same dimensional coordinate-system ${}^nF$. E.g. we could allow that most of the observers coordinatize events in 4-dimension, while some special (e.g. faster-than-light) observers coordinatize events with 2-dimensional coordinate-system only.

Figures 1 - 4 illustrate the structure of an arbitrary model $\mathfrak{M}$ (of dimension 2) in the sense of Definition 2.1.1. Consider the coordinate-system in the right hand side of Figure 1 (or in the left-hand sides of Figures 2, 3, 4). Intuitively, the first (vertical) axis is the <u>*time scale*</u> while the second (horizontal) axis represents <u>*space*</u>. The straight lines ℓ and ℓ_1 represent "lines" in Figure 1. The world-view relation W, which is the heart of our model, is illustrated in Figures 2–4. W is represented

[49] Sometimes it is worth contemplating why the proofs are different for different dimensions.

by the system of world-view functions $\langle\, w_m : m \in Obs \,\rangle$, cf. Figure 3. In Figure 4, $w_m(p) = \{b, ph\}$ means that m "sees" at time p_0 at location p_1 two bodies: b and ph. I.e, $W(m,p,b)$, $W(m,p,ph)$ are true, while e.g. $W(m,p,m)$ is not true.

For the time being we do not have a structure on the set $\mathcal{P}(B)$ of events, which we also call space-time. In the geometry chapter §6 we will put some structure on our space-time, too. Sometimes, the structure in Figure 3 is mathematically modeled by a so called manifold.[50]



The world-view relation W and world-view functions w_m.
m "sees" at time p_0 at location p_1 two bodies: b and ph.

Figure 4: Second drawing of the world-view function w_m.

We will use the following notation. For $Obs(b), Ph(b), Ib(b)$ we often write $b \in Obs$, $b \in Ph$, $b \in Ib$, respectively. Moreover, we will reserve the variables m, m_i, k, k_i to denote observers; we reserve ph, ph_i for photons; finally we use the symbols p, q, r, s to denote elements of ${}^nF$. Thus we have[51]

$m, m_i, k, k_i \in Obs;$
$ph, ph_i \in Ph;$
$p, q, r, s \in {}^nF.$

[50]The manifold structure is not particularly relevant at the present point, but it will be relevant in later developments.

[51]Sometime we will deviate from this convention though, for lack of enough letters. E.g. sometimes we will use m, k to denote natural numbers also.

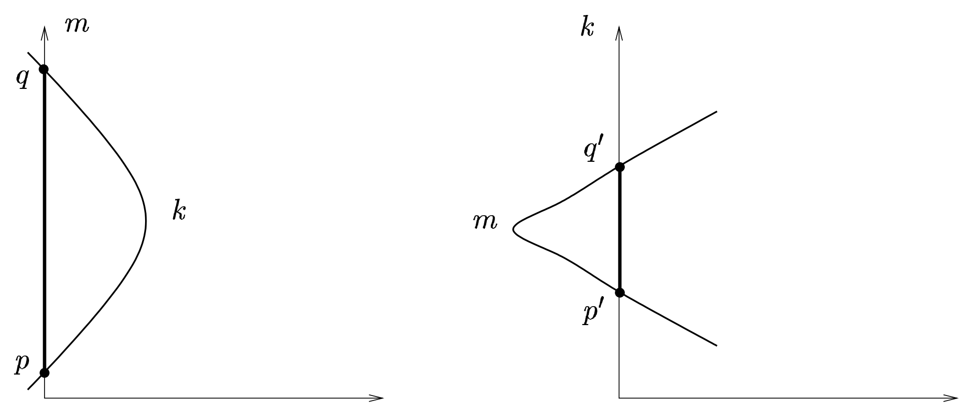
Using the terminology of vector spaces, elements of ${}^nF$ will often be referred to as *vectors*. As we mentioned, we use the convention from logic that

$\varphi(m)$ when used as an axiom, means $(\forall m \in Obs)\varphi(m)$.

(This is based on our convention above that m ranges over elements of Obs, and not B.)

Consider a frame model $\mathfrak{M}$ and its ordered field reduct $\mathfrak{F}$. We will sometimes impose the condition on our $\mathfrak{M}$ that $\mathfrak{F} = \mathfrak{R}$, the ordered field of real numbers.

We close this section with giving a possible formulation of the so called "*twin paradox*", as an example for a formula in our frame language.[52] Intuitively, the twin paradox says that if one of two twin brothers leaves the other (accelerating) and returns to him later, then the brother who stayed behind will be *older* at the time of their reunion. That is, more time has passed for the "non-moving" brother than for the traveling one.



m: "non-moving" (inertial) brother $\qquad$ k: traveling (accelerated) brother

Figure 5: The "twin paradox".

[52]We use natural abbreviations here, as well as later. E.g. we write "$m, k \in w_m(p) \cap w_m(q)$" in place of the longer "$W(m,p,m) \ \wedge \ W(m,q,m) \ \wedge \ W(m,p,k) \ \wedge \ W(m,q,k)$".

(TwP) $(\forall m \in Obs \cap Ib)(\forall k \in Obs \setminus Ib)(\forall p, q, p', q' \in {}^4F)$
$\Big(m, k \in w_m(p) \cap w_m(q) \ \wedge \ w_m(p) = w_k(p') \ \wedge \ w_m(q) = w_k(q')\Big) \ \Rightarrow$
$|p_t - q_t| > |p'_t - q'_t|$. [53]

See Figure 5. It is not a coincidence that on Figure 5, the life-line of m as seen by k is more "exotic" than that of k as seen by m. (The acceleration of m is sometimes negative and sometimes positive.) We will discuss the reason for this in §2.8, p.145 and in §8.

[53] The notation p_t, q_t was introduced on p.31.

2.2 Basic axioms Basax

Our next task is to postulate axioms in our frame language, expressing parts of our intuition about physical reality. Our first set of axioms to be proposed shortly will be called **Basax**, and it serves as one possible starting point for axiomatizing *special* relativity theory.

Before presenting **Basax**, we would like to say a few words about its place in the hierarchy of axiom systems which will be studied in the present work.

In section 8.1 we will introduce a further (actually a more "advanced") set **Acc** of axioms, in which we will allow accelerated observers, and accordingly, in **Acc** we will modify some of the postulates of **Basax** (e.g. we will modify item 7 below).[54] In section 3.3 we will define variants of the axioms of **Basax**, and variants of **Basax** itself (e.g. **Newbasax**).[55] These new versions will be more "balanced" in a sense, and will make it easier to move toward having accelerated observers, i.e. toward **Acc**. (On the other hand, our first choice, **Basax** has the advantage that its axioms are easy to formulate and understand, so it might be considered as a good starting point.) In later parts we will introduce stronger as well as weaker (than **Basax**) axiom systems. As we indicated in §1.1 (Broad Introduction), a plurality of competing axiom systems (or relativity theories) is an essential feature of logical analysis of a theory like relativity. Accordingly, in §3.4.2 and in §4 we introduce several axiom systems for the purposes indicated in §1.1. One of these purposes is conceptual analysis (started e.g. in Friedman [90] and Rindler [224]) which asks which axiom of relativity is responsible for what conclusion of the theory. Another purpose of this plurality is to study such variants of relativity as e.g. the Reichenbach-Grünbaum version and to compare them with the standard version. Also, we want to "fine-tune" our axiom systems in various regards. A further, but not negligible purpose in studying weaker axiom systems is to prove stronger theorems. For more on the motivation for having a plurality of axiom systems we refer the reader to §§1.1, 3.4.2, 4. See also Figures 180 and 223 on pp.552, 653. Finally we note, that besides

[54]**Acc** (and its theory) can be considered as a *first step* in the direction of *experimenting* with the idea of treating general relativity (in **Acc** we will have gravity, event horizons etc) in the framework of first-order logic in a spirit analogous with that of the present work.

[55]It belongs to the spirit of the axiomatic method that we start out with a simple set of axioms (like **Basax**), investigate its properties, prove some theorems from it, and then we use our so obtained experience for modifying this axiom system. After that, we restart the "cycle", i.e. we start investigating the new axiom system etc.

weakening (and/or modyfying) **Basax** we will also study the possibility of making it <u>*stronger*</u> by adding a few new, natural axioms, cf. e.g. §§ 2.8, 3.8, 3.9. In §3.8, we will also study an extension "**BaCo**+**Ax**($\sqrt{\ }$)" of **Basax**, which completely describes the standard, Minkowskian models of special relativity.

Before presenting **Basax**, we emphasize that it is only our first and simplest variant of an axiom system for special relativity. Later, we will also have: (i) axiom systems in which accelerated observers are permitted (i.e. informal postulate 2 below will be withdrawn), (ii) systems in which for different observers different events may exist (i.e. postulate 7 below will be withdrawn), (ii) systems in which the speed of light will be not the same for all osbservers, (iv) the Reichenbachian version of relativity where there is even less restriction on the speed of light, (v) systems in which the coordinate-system of an orbserver may be not the whole of ${}^4F$ but only a subset of ${}^4F$, etc.

Informally, about a model $\mathfrak{M} = \langle (B;\ Obs, Ph, Ib), \mathfrak{F}, G;\ \in, W \rangle$, **Basax** will postulate the following.

> $\mathfrak{F}$ is a linearly ordered field; we can thus define straight lines of the usual, Cartesian geometry over $\mathfrak{F}$, i.e. of ${}^nF$ (which, intuitively, are "*life-lines*" or "*traces*" of the motions of inertial bodies), and we can define angles of straight lines (which represent "*speeds*" of inertial bodies). In this sense of the word we will postulate the following:
>
> 1. G is the set of straight lines of the Cartesian geometry over $\mathfrak{F}$.
> 2. Observers and photons are inertial bodies.
> 3. The "trace" of an inertial body h as seen[56] by any observer m is in G.
> 4. Any observer m sees itself as being at rest in the origin.
> 5. Any observer sees some observer on each "slow" line.[57]
> 6. Each line which could be the life-line of a photon (according to item 8 below) is indeed the life-line of a photon.
> 7. Any two observers see the same events.
> 8. All observers see all photons moving with the same speed.

[56]Below, and later on, we will use the word "see" as a kind of intuitive (or "animated") way of referring to the act of <u>*observing*</u> via the world-view function, as we already indicated this (cf. footnote 44 on p.32).

[57]A line is called slow if its "speed" (i.e. angle with the time axis) is smaller than that of a photon.

In items 5 and 6 above by existence we mean only *potential* existence. I.e. when we say that on each slow line there exists an observer, what we mean is that potentially there can exist an observer, but in reality all these potential observers and photons need not be really there. The same applies to the existence of "potential" photons in item 6.[58]

For the formal definition of **Basax**, we need some preparation. We start with recalling some basic notions of linear algebra e.g. from Halmos [122] or Kostrikin-Manin [155] or Hausner [125] or [228] (or any other textbook on linear algebra).

If $p \in {}^nF$ for some set F and $n \in \omega$ then, for any $i < n$, p_i denotes the i-th component (projection) of p. Thus $p = \langle p_0, \ldots, p_i, \ldots, p_{n-1}\rangle = \langle p_i\rangle_{i<n}$.

Recall from any textbook on vector spaces (e.g. [122]) that, to any field $\mathbf{F} = \langle F, +, \cdot\rangle$ and natural number $n \in \omega$, an n-dimensional *vector space* ${}^n\mathbf{F}$ can be associated the following natural way. Defining $+^V : {}^nF \times {}^nF \longrightarrow {}^nF$ by

$$(\forall p, q \in {}^nF)\ p +^V q \stackrel{\text{def}}{=} \langle p_i + q_i\rangle_{i<n}\,,$$

$\langle {}^nF, +^V\rangle$ turns out to be a commutative group with neutral element

$$\bar{0} \stackrel{\text{def}}{=} \langle 0\rangle_{i<n}$$

and inverse $-^V p = \langle -p_i\rangle_{i<n}$ for any $p \in {}^nF$. With defining "multiplication by scalars" $\cdot^V : F \times {}^nF \longrightarrow {}^nF$ by

$$a \cdot^V p \stackrel{\text{def}}{=} \langle a \cdot p_i\rangle_{i<n} \qquad \text{for each } a \in F \text{ and } p \in {}^nF\,,$$

$\langle {}^nF, +^V\rangle$ becomes a vector space over the field $\mathbf{F}$. We denote this vector space by ${}^n\mathbf{F}$. We note that any n–dimensional vector space over $\mathbf{F}$ is isomorphic to ${}^n\mathbf{F}$ (see e.g. Halmos [122]). In universal algebra, there are two ways for making the notion of a vector space like ${}^n\mathbf{F}$ precise. These are the "one-sorted" and the "two-sorted" versions, defined below. The one-sorted version is defined as follows:

$${}^n\mathbf{F}_1 \stackrel{\text{def}}{=} \langle {}^nF, +^V, -^V, \bar{0}, f_a\rangle_{a\in F}$$

with f_a unary and $f_a(p) \stackrel{\text{def}}{=} a \cdot^V p$ for $p \in {}^nF$ and $a \in F$. Cf. Burris-Sankappanavar [54]. The two-sorted version is the structure

$${}^n\mathbf{F}_2 \stackrel{\text{def}}{=} \langle \mathbf{F}, {}^nF; +^V, -^V, \bar{0}, \cdot^V\rangle\,,$$

[58]We will return later to clarifying the issue of these potentially existing entities (observers, photons) which exist only potentially but need *not* exist *actually*. This can be made precise e.g. by using first-order modal logic as a framework as will be discussed soon.

where the operations $+^V$, $-^V$, $\bar{0}$ are defined on sort ${}^nF$ while $\cdot^V$ is of mixed sort, i.e. $\cdot^V : F \times {}^nF \longrightarrow {}^nF$. Throughout this work, ${}^n\mathbf{F}$ denotes either ${}^n\mathbf{F}_1$ or ${}^n\mathbf{F}_2$ depending on context. Occasionally we will explicitly indicate which one is meant. So whenever ${}^n\mathbf{F}$ shows up, it denotes the n-dimensional vector space over $\mathbf{F}$ without specifying whether we mean the one-sorted or the two-sorted version (the reader is asked to use the context if he wants to decide this).

We note that the notation ${}^n\mathbf{F}$ is slightly ambiguous (from a different point of view too) because ${}^n\mathbf{F}$ can denote the vector space over the field $\mathbf{F}$ but also (by the standard notation of universal algebra) it can denote the n'th direct (or Cartesian) power of the algebraic structure $\mathbf{F}$. This direct power happens to be a ring. Therefore we might talk about the vector space ${}^n\mathbf{F}$ or the ring ${}^n\mathbf{F}$ (they are not the same because they have different operations). If we do not indicate which one is meant then, by *default*, we mean the vector space. I.e. if the symbol ${}^n\mathbf{F}$ appears in the text (without an indication of whether we mean a vector space or a ring) then it *denotes a vector space*. A completely analogous convention applies to $\mathfrak{F}$ in place of $\mathbf{F}$.

As usual, we will often write $p -^V q$ in place of $p +^V (-^V q)$ for simplicity. Further, we will often omit the index V from $\cdot^V$, $+^V$ and $-^V$, and hope that context will always save us from misunderstandings.

CONVENTION 2.2.1

(i) Throughout, $\mathfrak{F}$ $(= \langle \mathbf{F}, \leq \rangle)$ denotes an arbitrary linearly ordered field. However, this is a context sensitive convention in the following sense: If there is a frame model $\mathfrak{M}$ around, then automatically $\mathfrak{F}$ denotes the ordered field reduct of $\mathfrak{M}$. A similar convention applies to the field $\mathbf{F}$, its universe F, coordinate system ${}^nF$, and vector space ${}^n\mathbf{F}$, e.g. if there is an $\mathfrak{F}$ around then automatically F denotes its universe etc. In the other direction if we talk about, say, F then implicitly we assume that there is an $\mathfrak{F}$ in the background etc.

(ii) As we already said in Def.2.1.1, when we work in ${}^n\mathbf{F}$ $(2 \leq n \leq 4)$, to match the physical intuition, we call the 0-th coordinate p_0 of a point $p = \langle p_0, \ldots, p_{n-1} \rangle \in {}^nF$ the *time coordinate* or *time component* of p. Accordingly, when drawing coordinate systems, we call the 0-th axis of it the *time axis* or *$\bar{t}$-axis*. The rest of the coordinates are the *space coordinates* or *space components*. We denote the first four coordinate axes as follows:

$$\bar{t} \stackrel{\text{def}}{=} F \times {}^{n-1}\{0\} \; (= F \times \{0\} \times \ldots \times \{0\}),$$
$$\bar{x} \stackrel{\text{def}}{=} \{0\} \times F \times {}^{n-2}\{0\},$$

$$\bar{y} \stackrel{\text{def}}{=} \{0\} \times \{0\} \times F \times {}^{n-3}\{0\}\,, \text{ and}$$
$$\bar{z} \stackrel{\text{def}}{=} \{0\} \times \{0\} \times \{0\} \times F \times {}^{n-4}\{0\}\,.$$

In general $\bar{x}_i$ denotes the i'th coordinate axis, that is
$$\bar{x}_i \stackrel{\text{def}}{=} {}^{i}\{0\} \times F \times {}^{n-i-1}\{0\}\,.$$

Also, we put

$$p_t \stackrel{\text{def}}{=} p_0,$$
$$p_x \stackrel{\text{def}}{=} p_1,$$
$$p_y \stackrel{\text{def}}{=} p_2,$$
$$p_z \stackrel{\text{def}}{=} p_3,$$

for each $p \in {}^nF$.

(iii) Throughout this work, the dimension n ($\in \omega$) of our space-time is a parameter of almost all of our concepts. Therefore a possibility for a rigorous presentation would be to indicate n in the name of each concept we introduce, e.g., by putting something like "(n)" after it. But then the text would become too complicated. Therefore we chose omitting the "(n)"-s except when this would lead to misunderstanding or when we want to emphasize the presence of n.

But sometimes we will define or state things for one particular n only (e.g., for just $n = 2$). In these cases we will indicate this fact by putting the particular number, in parenthesis, after the name of the concept involved. For example, we will formulate an axiom **Ax1**, where n will be a parameter of **Ax1**. Then the instance of **Ax1** for the case $n = 2$ will be denoted by **Ax1**(2).

Throughout this work $n > 1$. Therefore, we will not mention this explicitly.

We will treat some other parameters likewise. E.g., we will sometimes state things for a collection of models from FM such that all $\mathfrak{M} \in \mathsf{FM}$ share the same ordered field $\mathfrak{F}$ as their "quantity part". Then we will denote this collection by $\mathsf{FM}(\mathfrak{F})$.

In cases when we will need more than one parameter we will list them in parentheses, separated by commas. For example,

$$\mathsf{FM}(3, \mathfrak{R}) = \left\{\, \mathfrak{M} \in \mathsf{FM}(3) \;:\; \mathfrak{F}^{\mathfrak{M}} = \mathfrak{R} \,\right\}\,.$$

That is, $\mathfrak{M} \in \mathsf{FM}(3, \mathfrak{R})$ iff $\mathfrak{M}$ is of dimension 3 and the quantity part of $\mathfrak{M}$ is the ordered field $\mathfrak{R}$ of real numbers.

$\lhd$

Besides our frame language introduced in section 2.1, we will also use the language of the vector space ${}^n\mathbf{F}_2$ (as an extension of our frame language) for expressing ideas concisely. (E.g., for $r, s \in {}^nF$ we may mention the vectors $r + s$ or $3 \cdot r$.) We are allowed to do this since the ${}^n\mathbf{F}_2$ formulas are translatable to our frame language. As a first example for this and for the other natural abbreviations we will use, we introduce our first axiom **Ax1** both as a formula in a concise style translatable to our frame language[59] and, equivalently, as a (longer) formula written purely in the frame language.

The set of straight lines of ${}^n\mathbf{F}$ in the usual Euclidean sense is denoted by $\mathsf{Eucl} := \mathsf{Eucl}(n, \mathfrak{F}) := \mathsf{Eucl}(n, \mathbf{F})$, that is,

$$\ell \in \mathsf{Eucl}(n, \mathbf{F}) \stackrel{\text{def}}{\Longleftrightarrow} (\exists r, s \in {}^nF)\Big(s \neq \bar{0} \ \wedge \ \ell = \{\, r + a \cdot s \ : \ a \in F \,\}\Big)\,.$$ [60]

Ax1 in a concise language:
$G = \mathsf{Eucl}(n, \mathbf{F})$.

Ax1 in the frame language of relativity theory:

Ax1' $(\forall r_0, \ldots, r_{n-1}, s_0, \ldots, s_{n-1} \in F)$
$\Big(\{s_0, \ldots, s_{n-1}\} \neq \{0\} \ \Rightarrow \ (\exists \ell \in G)(\forall p_0, \ldots, p_{n-1} \in F)\,(\ \mathrm{E}\ (p_0, \ldots, p_{n-1}, \ell) \Leftrightarrow$
$(\exists a \in F) \ \bigwedge_{i<n} p_i = r_i + a \cdot s_i)\Big)$
and
$(\forall \ell \in G)(\exists r_0, \ldots, r_{n-1}, s_0, \ldots, s_{n-1} \in F)$
$\Big(\{s_0, \ldots, s_{n-1}\} \neq \{0\} \ \wedge \ (\forall p_0, \ldots, p_{n-1} \in F)(\ \mathrm{E}\ \ (p_0, \ldots, p_{n-1}, \ell) \Leftrightarrow$
$(\exists a \in F) \ \bigwedge_{i<n} p_i = r_i + a \cdot s_i)\Big)$.

Here we emphasize that **Ax1** is designed to serve the purposes of special relativity only. In later parts when dealing with more general theories of relativity, **Ax1** will be changed.

[59] I.e. in **Ax1** we use convenient abbreviations reducible to our frame language.

[60] Note that after this definition the formula $\ell \in \mathsf{Eucl}(n, \mathbf{F})$ counts as a formula of our frame language. Namely it abbreviates the following formula of our frame language:
$\ell \in G$ and $(\exists r, s \in {}^nF)\Big(s \neq \bar{0} \ \wedge \ (\forall p \in {}^nF)[p \in \ell \ \Leftrightarrow \ (\exists a \in F)\, p = r + a \cdot s]\Big)$.

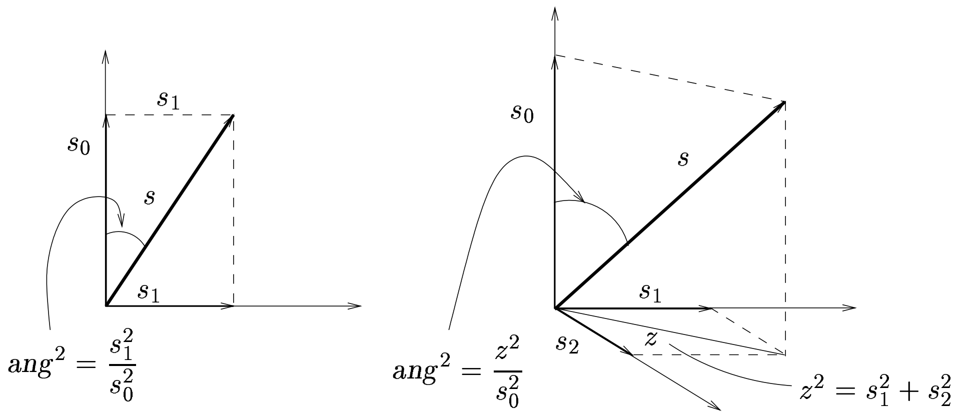


Figure 6: Angle of a line.

If $\ell \in \mathsf{Eucl}(n, \mathbf{F})$, then we can consider the angle between ℓ and the time axis. By $ang^2(\ell)$ we denote the square of the tangent of the angle between ℓ and the time axis.[61] Thus, for $\ell = \{\, r + a \cdot s \;:\; a \in F \,\} \in \mathsf{Eucl}(n, \mathbf{F})$,

$$\begin{array}{lcll} ang^2(\ell) & \stackrel{\text{def}}{=} & \dfrac{s_1^2 + s_2^2 + \ldots + s_{n-1}^2}{s_0^2} & \text{if } s_0 \neq 0, \text{ and} \\ ang^2(\ell) & \stackrel{\text{def}}{=} & \infty & \text{if } s_0 = 0. \end{array}$$

(It will cause no problem that infinity ∞ is not an element of F.) Thus $0 \leq ang^2(\ell) \leq \infty$. $ang^2(\ell) = 0$ means that ℓ is vertical, $ang^2(\ell) = 1$ intuitively means that the angle between ℓ and the time axis is 45^o, and $ang^2(\ell) = \infty$ means that ℓ is horizontal. The definition of $ang^2(\ell)$ is illustrated in Figure 6.

Definition 2.2.2 (life-line (or trace), speed)
Let $\mathfrak{M}$ be a frame model as in Definition 2.1.1. Let $m \in Obs$ and $b \in B$ be arbitrary but fixed. Recall from Definition 2.1.1 that the world-view function $w_m : {}^nF \longrightarrow \mathcal{P}(B)$ of m was defined as follows:

$$w_m(p) = \{\, b \in B \;:\; \langle m, p, b \rangle \in W \,\} \qquad \text{for every } p \in {}^nF\,.$$

[61]We consider the square of the tangent (instead of the tangent itself) of this angle because, in general, we do not assume that square-roots exist in $\mathbf{F}$.

(i) By the *life-line (*or *trace) of* b *as seen by* m (or life-line (or trace) of b by the world-view of m) we mean the set

$$tr_m(b) \stackrel{\text{def}}{=} \{\, p \in {}^nF \; : \; b \in w_m(p) \,\} \;\; = \{\, p \in {}^nF \; : \; W(m,p,b) \,\}.$$

(ii) If $tr_m(b) \in \mathsf{Eucl}(n, \mathbf{F})$, then by the *speed of* b *as seen by* m we mean

$v_m(b) \stackrel{\text{def}}{=} ang^2(tr_m(b)),$

cf. Figure 7.

The formula $v_m(b) = a$ will abbreviate that

$tr_m(b) \in \mathsf{Eucl}(n, \mathbf{F}) \qquad \text{and} \qquad ang^2(tr_m(b)) = a.$

$\lhd$

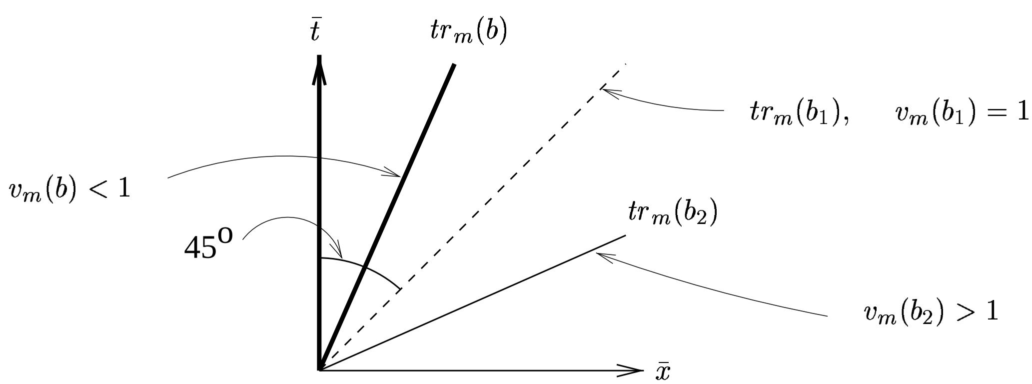


Figure 7: Traces and speeds.

In Figure 7, the line $tr_m(b)$ illustrates the life-line of a body b (in case $n = 2$). The acronym "*tr*" stands for "*trace*". If $tr_m(b) = \{\, \langle t, x, y, z \rangle \; : \; t \in F \,\}$, then m always, at each time instance $t \in F$, sees b at location $\langle x, y, z \rangle$, i.e. m sees the body b at rest at location $\langle x, y, z \rangle$. Thus, $tr_m(b)$ is a vertical line (a line parallel with the time axis), i.e. $v_m(b) = 0$, means that "b is at rest, as seen by m". Similarly, the bigger $v_m(b)$ is, the more "speed" b is moving with, as seen by m, cf. Figure 7.

As we said $v_m(b)$ is called the *speed* of b as seen by m. To be more precise it is the square of the usual speed (since we used ang^2 instead of ang). The reason for using the square of quantities (in place of the original quantities) is that we do not want to assume that square-roots exist in $\mathbf{F}$. So, speed is a scalar (i.e. element of F). As opposed to speed, the *velocity* $\vec{v}_m(b)$ of b as seen by m is an $(n-1)$-vector, i.e. $\vec{v}_m(b) \in {}^{n-1}F$, defined as follows (cf. Figures 8, 9). Let $\mathfrak{M}$ be a frame model. Let $m \in Obs$ and $b \in B$ such that $tr_m(b) = \ell = \{\, r + a \cdot s \ : \ a \in F \,\} \in \mathsf{Eucl}$, for some r and $s \neq \bar{0}$. Assume that $s_0 > 0$. Then

$$\vec{v}_m(b) \stackrel{\text{def}}{=} \vec{v}_m(\ell) \stackrel{\text{def}}{=} \langle s_1/s_0, \dots, s_{n-1}/s_0 \rangle \ (= \langle s_1, \dots, s_{n-1} \rangle / s_0).$$

If $s_0 = 0$ then

$$\vec{v}_m(b) \stackrel{\text{def}}{=} \{a \cdot s : a \in F\}.$$ [62]

If $s_0 = 0$, then $\bar{v}(m)(b)$ is infinite (i.e. $\bar{v}(m)(b) = \infty$), therefore we cannot represent $\vec{v}_m(b)$ as a finite vector. Therefore, the information content of $\vec{v}_m(b) = \ell$ (where ℓ is in the space part of ${}^nF$) remains that b is moving in direction ℓ with infinite speed both "forward" and backward". We note that the speed $v_m(b)$ is the (square of) distance covered by b in unit time; while the velocity $\vec{v}_m(b)$ is the vector representing the change of location which happened in unit time, see Figure 8, assuming $v_m(b) \neq \infty$. For more on the distinction between speed and velocity cf. e.g. Gardner [98, p.7].

We are ready to postulate axioms **Ax2**–**Ax6**.

Ax2 $Obs \cup Ph \subseteq Ib$.

That is, observers are inertial bodies; and so are photons.

Ax3 $(\forall h \in Ib)(\forall m \in Obs)\ \Big(tr_m(h) \in G\Big)$.

That is, the life-line of any inertial body h as seen by any observer m must be a "line".

Ax4 $(\forall m \in Obs)\ \Big(tr_m(m) = \bar{t} \quad (= F \times {}^{n-1}\{0\})\Big)$.

Ax4 states that the life-line $tr_m(m)$ of an observer as seen by itself is the 0-th axis (the time axis). Thus **Ax4** says that each observer sees itself to be a body at rest (not moving) at (space) location $\langle 0, \dots, 0 \rangle$. In particular, $v_m(m) = 0$. This is one of the basic axioms of relativity theory. This was a "relativistic axiom" already before Einstein.[63] It expresses that each observer

[62] We note that the $s_0 = 0$ case of this definition in *not* important (i.e. we could have said that $\vec{v}_m(b)$ is undefined if $s_0 = 0$).

[63] Sometimes a slightly stronger form of this is referred to as Galileo's relativity principle. Galileo's relativity principle says a bit more than just **Ax4**. Cf. e.g. Geroch [101], pp.32-39, in particular §3 entitled "The Galilean View".

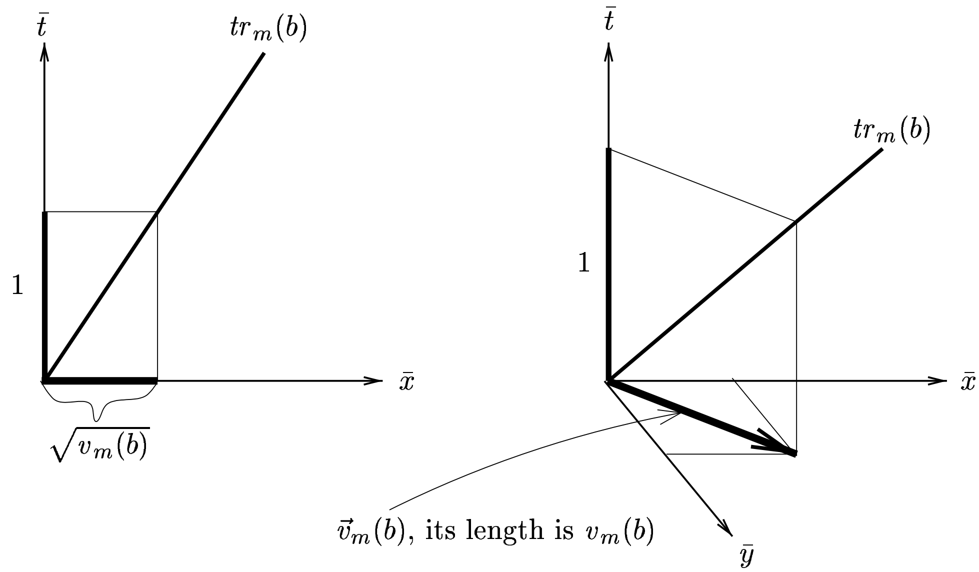


Figure 8: Speed and velocity.

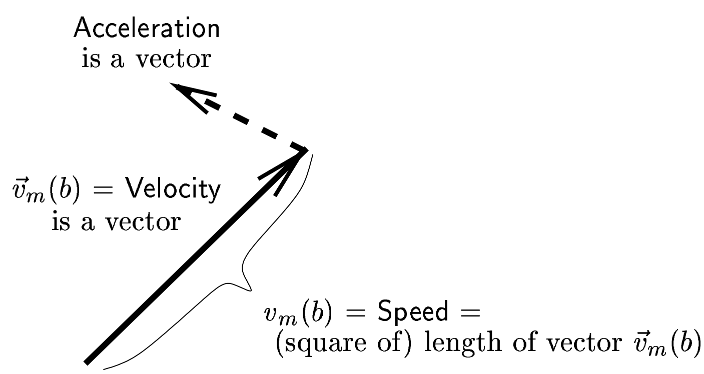


Figure 9: Velocity, speed, and acceleration represented purely in space (the time dimension is suppressed). (Of this figure, acceleration will be relevant only in §8.)

can "think" that he is at rest and all other bodies are moving. The first step toward general relativity theory will be that we will extend **Ax4** to accelerated observers, too[64]: then even accelerated observers can "think" that they are at rest (and then, in a poetic language, gravity will come into the picture to explain certain strange behavior of other bodies).[65] On an intuitive level, the principle on which special relativity is based is quoted as "all inertial observers [or reference frames] are equivalent" (at this point our future **AxE** will play a role, too); and the principle of general relativity is quoted as "all observers [including the accelerated ones] are equivalent". (Here "equivalent" means only that each of these observers may imagine that he is not moving and it is the rest of the universe which moves, accelerates etc.)

Ax5 $(\forall m \in Obs)(\forall \ell \in G)\ \Big(ang^2(\ell) < 1 \ \Rightarrow\ (\exists k \in Obs)\, \ell = tr_m(k)$ and

$$ang^2(\ell) = 1 \ \Rightarrow\ (\exists ph \in Ph)\, \ell = tr_m(ph)\Big).$$

Ax5 makes sense only in the presence of **Ax1** (because $ang^2(\ell)$ is not defined otherwise). Then it states that we have the tools for (performing) thought-experiments: on any appropriate straight line *we can assume* there is an observer; and the same for photons.[66] Later we will weaken the first part of **Ax5** to say that there is a positive c such that in every direction for every positive $\lambda < c$ there is an observer going in that direction and with speed λ. (I.e., $(\forall m)(\exists c > 0)(\forall \ell)[ang^2(\ell) < c \Rightarrow (\exists k \in Obs)\ell = tr_m(k)]$. Cf. **Ax(5nop)**$^{-+}$ in §5, pp.761, 763.) This weaker form of the axiom is sufficient for many purposes.

Ax6 $(\forall m, k \in Obs)\ \Big(Rng(w_m) = Rng(w_k) \Big).$

Ax6 states that all observers see the same set of events. I.e. whenever an observer m sees a set E of bodies at some time point t and space location s, any other observer k must see the same set E of bodies at *some* time point t' and space location s'. In still other words, the same events "exist" or "are

[64] According to e.g. Friedman [90], p.5, general relativity begins with the study of accelerated observers (or accelerated reference frames) (at least when they are treated "equivalently" with inertial reference frames). In this sense, our chapter 8.1 deals with the (first steps of the) generalization of our (logic based) method from special relativity to general relativity.

[65] Cf. e.g. p.147 and Figure 48 in the discussion of the twin paradox in §2.8.

[66] In a later version of the present work we will see a (first-order) modal logic refinement (or variant) of our axioms (and formalism) in which **Ax5** sounds "less radical" (that is, sounds more convincing intuitively). The modal version of **Ax5** avoids making space-time "overcrowded" with observers and photons.

available" for all observers. **Ax6** is quite strong. In particular, it will not be true in our theory of accelerated observers (or in general relativity).[67] Later we will weaken **Ax6** to $\mathbf{Ax6_{00}}$ such that the new version will be true for our accelerated observers, too. The new version $\mathbf{Ax6_{00}}$ will say that if m sees an event E on the trace of the observer k, then k itself sees this event E.

Our last axiom in the present section is the most distinguished one in relativity theory:[68]

AxE $(\forall m \in Obs)(\forall ph \in Ph)\ v_m(ph) = 1.$

AxE ("*Einstein's axiom*") states that the speed of a photon ph, as seen by any observer m, is always 1. In **Basax**, we choose the "speed of light" to be 1. This is a rather ad-hoc decision, the important part of **AxE** is that all observers see all photons as having the same speed. Later, e.g. in §4, we will weaken **AxE** in several ways.[69] We will see that *already* most of these weak forms of **AxE** will be enough for proving the majority of the important consequences of **Basax**. In particular, we will see that the weaker postulates saying that in each direction there is a photon going forwards and that "photons do not race with one another like bullets do" in place of **AxE** are already sufficient (together with the other axioms, of course) to prove most of the interesting theorems of special relativity theory.

Definition 2.2.3 (Basax) We define

$$\mathbf{Basax} \stackrel{\text{def}}{=} \{\, \mathbf{Ax1}, \mathbf{Ax2}, \mathbf{Ax3}, \mathbf{Ax4}, \mathbf{Ax5}, \mathbf{Ax6}, \mathbf{AxE} \,\},$$

where the axioms **Ax1**–**Ax6**, **AxE** were defined above.

[67] One reason for this is that if observer k accelerates (in m's world) so fast that its clock will never reach 12 o'clock as seen by m, then the "event" seen by k at 12 o'clock (or after 12) will not be "seen" by m.

[68] One could refer to e.g. the Michelson-Morley experiment for motivation, but instead of doing that, we refer to the introduction of Friedman [90].

[69] One of these says that each observer m sees all photons with the same speed, another one is the Reichenbach-Grünbaum version of **AxE** etc. Cf. §4. Moreover, following an idea of Gyula Dávid [71], in §5 we will see a variant of **Basax** which (proves most of usual relativity and) does not need **AxE** at all.

Here is a summary of the axioms in **Basax**:

Ax1 $G = \mathsf{Eucl}(n, \mathbf{F})$.

Ax2 $Obs \cup Ph \subseteq Ib$.

Ax3 $(\forall h \in Ib)(\forall m \in Obs)\ \Big(tr_m(h) \in G\Big)$.

Ax4 $(\forall m \in Obs)\ \Big(tr_m(m) = \bar{t}\Big)$.

Ax5 $(\forall m \in Obs)(\forall \ell \in G)\ \Big(ang^2(\ell) < 1\ \Rightarrow\ (\exists k \in Obs)\, \ell = tr_m(k)$ and

$$ang^2(\ell) = 1\ \Rightarrow\ (\exists ph \in Ph)\, \ell = tr_m(ph)\Big).$$

Ax6 $(\forall k, m \in Obs)\ \Big(Rng(w_m) = Rng(w_k)\Big)$.

AxE $(\forall m \in Obs)(\forall ph \in Ph)\ v_m(ph) = 1.$ $\lhd$

It follows from **Ax2**,**Ax3** that the trace of any observer is a line. Since we will often use this conclusion, we are giving it a name:

(geod) $(\forall m, k \in Obs)\ tr_m(k) \in G$.

Statement **(geod)** together with **Ax1** imply that $tr_m(k)$ is a *Euclidean* straight line. Later, **Ax1** will be generalized so that G will be a more general geometry-like structure, e.g. G might consist of the geodesics of some structure. Beginning with §8 where we will have accelerated observers too, **(geod)** will be restricted to inertial observers.

The world-view function w_m can be recovered from the family of traces of all bodies (from $\langle tr_m(b)\ :\ b \in B\rangle$), and the world-view-relation W can be recovered from all the world-view functions (from $\langle w_m : m \in Obs\rangle$). Thus we can "represent" the function w_m by the <u>world-view</u> of m, which is just the indexed family $\langle tr_m(b) : b \in B\rangle$, and which, in turn, we represent by drawing the traces of bodies that we are interested in. See Figure 10.

Assuming **Basax**, we can (and will often) draw the world-view of an observer m as shown in Figure 11. In this figure and in similar pictures, most often we simply write the name of a body h instead of writing out the long expression $tr_m(h)$, when indicating the life-line of h (as seen by m).

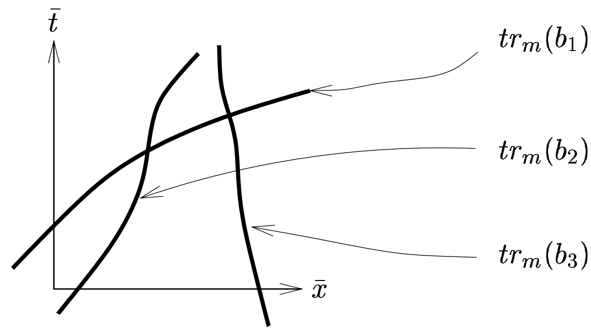


Figure 10: World-view of m.

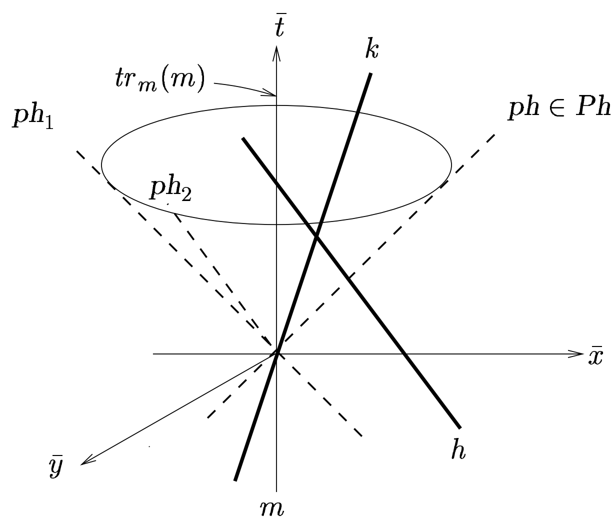


Figure 11: The world-view of an observer m in a model of **Basax**.

We will sometimes use the following.

FACT 2.2.4

(i) *Assume* **Basax**. *Let* $h \in Ib$ *be an* <u>inertial body</u> *with* $v_m(h) \neq \infty$. *Then* $tr_m(h) : F \longrightarrow {}^{n-1}F$ *is a function everywhere defined on* F, *where we think of* F *as the time axis* $\bar{t}$ *and of* ${}^{n-1}F$ *as "space".*[70]

(ii) *Statement (i) above remains true if we replace* **Basax** *with* $\{\mathbf{Ax1}, \mathbf{Ax3}\}$.

We omit the **proof**. ∎

By the *space part* S of ${}^nF$ we understand the subspace $S \stackrel{\text{def}}{=} \{ \langle 0, q_1, \ldots, q_{n-1} \rangle \,:\, q \in {}^nF \} (= \{q \in {}^nF \,:\, q_0 = 0\})$. Throughout chapter 2 we will *identify* S with ${}^{n-1}F$ to simplify notation. In later chapters we do *not* identify S with ${}^{n-1}F$.

Remark 2.2.5 (Terminology: Observers, reference frames, "slim observers", "fat observers")
We call the (sometimes partial[71]) function $w_m : {}^nF \longrightarrow \mathcal{P}(B)$ the *world-view* function of observer m.

(i) Some authors call w_m the *reference frame* of observer m, cf. e.g. d'Inverno [75]. We could have used that word instead of world-view function, it is only a historical accident that we chose the other name.[72]

(ii) Some authors eliminate "observers" and talk *only* about reference frames (i.e. world-view functions) w's (with $w : {}^nF \longrightarrow \mathcal{P}(B)$), instead. This is absolutely justified, because given a world-view function $w : {}^nF \longrightarrow \mathcal{P}(B)$ we can *recover* an observer, call it m, from w such that, after some modifications, basically w will be the world-view function of m. In more detail: We "create" a *new body* m by *postulating* that the set of events in which m is present should be $w[\bar{t}]$. Next, we *expand* all the world-view functions of our model with this new m. With this all properties of m *as a body* are defined. Now, we *raise* m to the *rank* of being an

[70] More precisely, we can regard the relation $tr_m(h) \subseteq {}^nF$ as a function $tr_m(h) : F \longrightarrow {}^{n-1}F$ by identifying $F \times {}^{n-1}F$ with ${}^nF$.

[71] w_m will become a partial function in much later chapters e.g. in the chapter on accelerated observers, §8. Also, when preparing the present framework for generalization to general relativity we will make w_m partial.

[72] We think, that in mathematics the choice of words is not so terribly important; the important thing is the definitions we supply for them, and the way we use the so defined terms in the subsequent discussions.

observer by postulating that the world-view function w_m of m is *defined* to be w. This construction shows that a reference frame w completely determines an observer m such that m's world-view function is w. The above illustrates that *if we wish* we could forget the observers m and talk about reference frames w instead. Then instead of a set Obs, another set Rfm of reference frames would be given as one of our primitives. (We could let $Rfm := \{ w_m : m \in Obs \}$.) The above train of thought shows that our approach and the "only reference frames" approach are equivalent (inter-definable) and it is not important whether we start out with observers (Obs) or reference frames (Rfm) in our basic vocabulary.

(iii) Our observers are "slim" in the respect that their life-lines (or traces) are thin curves in ${}^nF$. This again is not important, it is again *only* a *choice of words*: Namely, we could identify observer m with its world-view function w_m, and then it would cease to be "slim" in the above sense. In passing, we also note that instead of a single body m, we could have used as an observer *m together with a set K* of bodies [slim observers] such that $(\forall k \in K)$ ($tr_m(k)$ would be parallel with $\bar{t}$).[73] But since the final mathematical effects would remain[74] more or less the same[75] (via interdefinability), we decided to stick with an observer being a single body $m \in B$ and whenever we would need a "fat observer" like K above, we will simply recover it from the reference frame (i.e. world-view) w_m of m.

$\triangleleft$

Remark 2.2.6 Throughout, we will use the standard practice from logic of introducing new relation and function symbols by defining them, and then treating them as if they were symbols of our original language. E.g. we defined the function w_m and then we used it in our axioms (as if it was part of our language). We believe that translating the so enriched language back to the original first-order language is straightforward. For such translating algorithms see e.g. Monk [197, pp. 206–210] or Bell-Machover [44, p.97].

$\triangleleft$

[73] In this case we could think of an element k of K as a "partner" of m representing a time-like coordinate-line for m. Then $m \in K$ could be called the "central partner" in $\langle m, K \rangle$. Such an observer $\langle m, K \rangle$ could be visualized as a cloud of "particles" floating in space and each particle having a clock. Etc.

[74] at least from the point of view of questions investigated in this work

[75] As we said, on the long run we allow $w_m : {}^nF \longrightarrow \mathcal{P}(B)$ to be a partial function, i.e. $Dom(w_m) \subsetneq {}^nF$ is allowed.

2.3 Some properties of Basax, world-view transformation

In this section we introduce the notion of world-view transformations. We discuss some simple consequences of our basic axioms – to get a feel for them – and then we investigate those functions that occur as world-view transformations in models of **Basax**. We close this section with listing some basic properties of **Basax** as a logical theory (like consistency, independence, categoricity).

Definition 2.3.1 (world-view transformation)
Given $m, k \in Obs$, we define the *world-view transformation* f_{mk} as follows:

$$\mathsf{f}_{mk} \stackrel{\text{def}}{=} w_m \circ w_k^{-1} .$$

$\triangleleft$

We note that ${w_k}^{-1}$ is a relation, hence the composition $w_m \circ {w_k}^{-1}$ is again a relation, cf. the definition of composition on p.26. Thus $\mathsf{f}_{mk} \subseteq {}^nF \times {}^nF$ and

$$\mathsf{f}_{mk} = \{ \langle p, q \rangle \in {}^nF \times {}^nF \ : \ w_m(p) = w_k(q) \} ,$$

see Figure 12. Thus f_{mk} is a binary relation on the coordinate system ${}^nF$; two points are f_{mk}-related when m and k see the same "events" at those points. See also Figure 13.

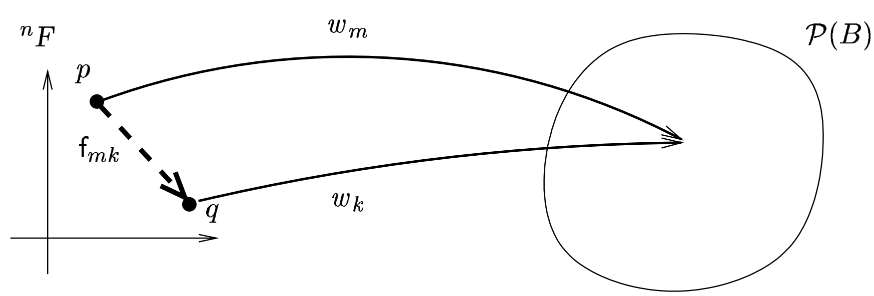


Figure 12: The world-view transformation.

The name "world-view *transformation*" suggests that f_{mk} is a function, i.e. to any $p \in {}^nF$ there is at most one q such that p is f_{mk}-related to q. This indeed will be the case in models of e.g. **Basax**, see Prop.2.3.3(v).[76] In arbitrary frame models, f_{mk} can be an arbitrary binary relation.[77] As we said, in models of **Basax**, f_{mk} cannot be an arbitrary binary relation, e.g. it has to be a function. Towards the end of this section we *characterize* those functions that occur in models of **Basax**(2) as world-view transformations cf. Thm.2.3.12, and also there we give some hints for the $n > 2$ case.

Figure 13 illustrates the world-view transformation f_{mk} for the 2-dimensional case. We drew the picture under the assumption that $\mathsf{f}_{mk} : {}^2F \longrightarrow {}^2F$, and we indicated two copies of ${}^2F$, the usual coordinate system way. The world-view of m

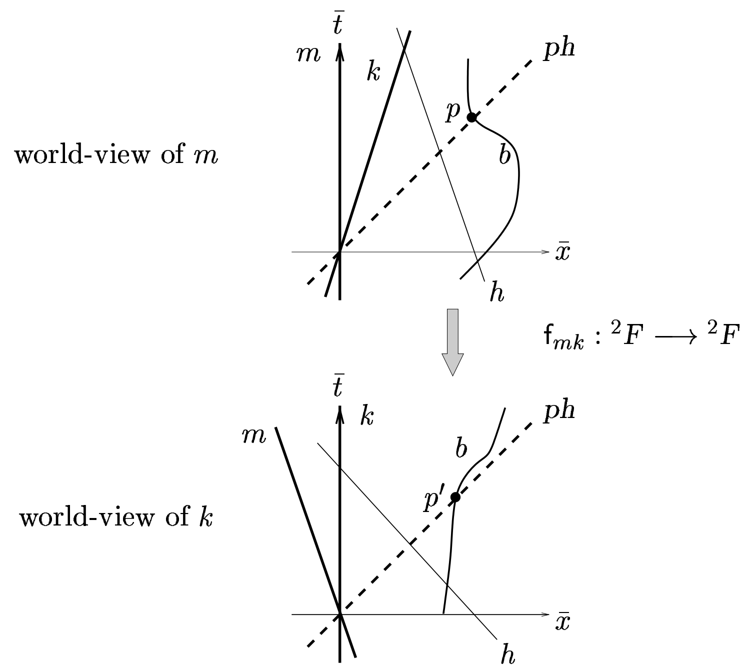


Figure 13: World-view transformation. The event when "m, k and ph are together" happens at $\bar{0}$ both for m and for k, hence $\mathsf{f}_{mk}(\bar{0}) = \bar{0}$. The event when ph and b are together happens at p for m and at p' for k; thus $\mathsf{f}_{mk}(p) = p'$.

[76] f_{mk} will be a *partial function* in all of the axiom systems, besides **Basax**, studied in the present work.

[77] By this we mean that for any ordered field $\mathfrak{F}$ and a binary relation $R \subseteq {}^nF \times {}^nF$, there are a frame model $\mathfrak{M}$ and two observers m, k in $\mathfrak{M}$ such that $R = \mathsf{f}_{mk}$.

is illustrated in the top coordinate system, and the world-view of k is in the bottom coordinate system (we did not represent in the picture all traces and all points of the world-views).

As a warm-up we begin with simple statements about our axiom system **Basax**. Let us recall that $\mathsf{Eucl} = \mathsf{Eucl}(n, \mathfrak{F})$ is the set of straight lines defined on p.45.

Notation 2.3.2 We define the sets of n-dimensional *slow-lines* SlowEucl and *photon-lines* PhtEucl over an ordered field $\mathfrak{F}$ as follows.

$$\begin{aligned} \mathsf{SlowEucl} &\stackrel{\text{def}}{=} \mathsf{SlowEucl}(n,\mathfrak{F}) \stackrel{\text{def}}{=} \{\, \ell \in \mathsf{Eucl}(n,\mathfrak{F}) \ :\ ang^2(\ell) < 1 \,\}\ , \\ \mathsf{PhtEucl} &\stackrel{\text{def}}{=} \mathsf{PhtEucl}(n,\mathfrak{F}) \stackrel{\text{def}}{=} \{\, \ell \in \mathsf{Eucl}(n,\mathfrak{F}) \ :\ ang^2(\ell) = 1 \,\}\,. \quad \lhd \end{aligned}$$

In connection with Prop.2.3.3(x) below, let us recall from p.27, that Id is the identity function on ${}^nF$.

PROPOSITION 2.3.3 *Let $\mathfrak{M}$ be a frame model of* **Basax**. *Then the following are true for all $m, k, h \in Obs$, $ph \in Ph$ and $b \in B$.*

(i) $Obs \cap Ph = \emptyset$, *i.e. no photon can be an observer.*

(ii) $tr_m(k) \neq tr_m(ph)$, *i.e. no observer can travel together with a photon.*

(iii) $v_m(k) \neq 1$, *i.e. the speed of an observer is never* 1.

(iv) *The world-view function w_m is an injection (i.e. one-one). That is, no observer "sees" the same event at two different space-time locations.*

(v) *The world-view transformation f_{mk} is a bijection (i.e. one-one, defined on ${}^nF$ and onto ${}^nF$).*

(vi) $w_m = \mathsf{f}_{mk} \circ w_k$. *I.e. we get the world-view of m from that of k by "applying f_{mk}" to it; f_{mk} is the "conversion" between m's and k's world-views.*

(vii) f_{mk} *takes the trace of a body as seen by m to the trace of the body as seen by k, i.e.* $\mathsf{f}_{mk}[tr_m(b)] = tr_k(b)$.

(viii) f_{mk} *takes slow-lines to straight lines, i.e. if $\ell \in \mathsf{SlowEucl}$, then $\mathsf{f}_{mk}[\ell] \in \mathsf{Eucl}$.*

(ix) f_{mk} *takes photon-lines to photon-lines, i.e. if $\ell \in \mathsf{PhtEucl}$, then* $\mathsf{f}_{mk}[\ell] \in \mathsf{PhtEucl}$.

(x) $\mathsf{f}_{mm} = \mathrm{Id}$, $\mathsf{f}_{mk} = \mathsf{f}_{km}^{-1}$, *and* $\mathsf{f}_{mk} = \mathsf{f}_{mh} \circ \mathsf{f}_{hk}$.

All the statements in Proposition 2.3.3 can be expressed with (first-order) formulas in our frame-language. We note that none of (i)-(ix) in Prop.2.3.3 is true without assuming (at least part of) **Basax**. We invite the reader to construct frame models in which these statements fail. We will prove the items in Prop.2.3.3 one-by-one, so that we can single out the axioms we need for proving them.

Claim 2.3.4 $\{\mathbf{Ax4}, \mathbf{AxE}\} \models Obs \cap Ph = \emptyset$.

Proof: Assume that $m \in Obs \cap Ph$. Look at $v_m(m)$. By **Ax4** we have that $v_m(m) = 0$, and by **AxE** and $m \in Ph$ we have that $v_m(m) = 1$. Since in all fields 0 and 1 are different elements, we reached a contradiction. ∎

Claim 2.3.5 $\{\mathbf{Ax4}, \mathbf{Ax6}, \mathbf{AxE}\} \models tr_m(k) \neq tr_m(ph)$.

Proof: Assume that $tr_m(k) = tr_m(ph)$. Then $tr_k(k) = \bar{t}$ and $v_k(ph) = 1$ by **Ax4** and **AxE**; in this connection note that $v_k(ph) = 1$ implies that $tr_k(ph) \in \mathsf{Eucl}$ by the convention on p.47. Thus $tr_k(k) \neq tr_k(ph)$. Then k sees an event in which k is present but ph is not present (namely such is $w_k(p)$ for any $p \in tr_k(k) \setminus tr_k(ph)$). However, m does not see such an event by $tr_m(k) = tr_m(ph)$. This contradicts **Ax6**, proving the proposition. See Figure 14. ∎

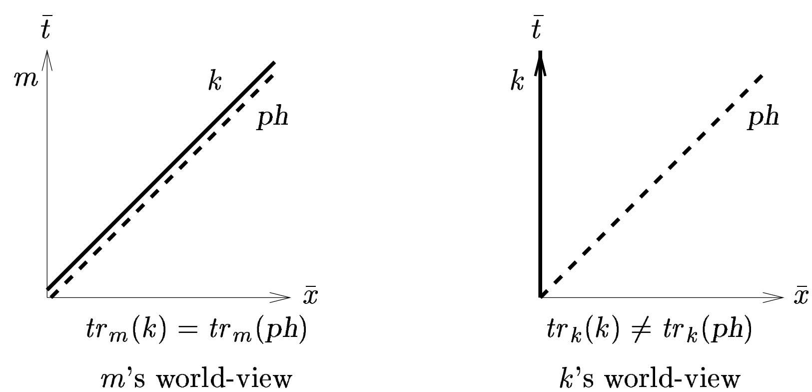


Figure 14: An observer cannot travel together with a photon.

Claim 2.3.6 $\{\mathbf{Ax1}, \mathbf{Ax4}, \mathbf{Ax5}, \mathbf{Ax6}, \mathbf{AxE}\} \models v_m(k) \neq 1$.

Proof: Assume that $v_m(k) = 1$ for some $m, k \in Obs$. Then $ang^2(tr_m(k)) = 1$, thus by **Ax5**, $tr_m(k) = tr_m(ph)$ for some $ph \in Ph$. This contradicts Claim 2.3.5. ∎

Claim 2.3.7 $\{\mathbf{Ax1}, \mathbf{Ax5}\} \models (\forall m \in Obs)(w_m$ is an injection).

Proof: Let $m \in Obs$ and assume that $p, q \in {}^nF$, $p \neq q$. Then, by **Ax1** and by the properties of $\mathsf{Eucl}(n, \mathbf{F})$, $(\exists \ell \in G)(p \in \ell \ \wedge \ q \notin \ell \ \wedge \ ang^2(\ell) < 1)$. By **Ax5**, $(\exists k \in Obs)\ell = tr_m(k)$. For such a k, $k \in w_m(p)$ but $k \notin w_m(q)$. ∎

Claim 2.3.8

(i) $\{\mathbf{Ax1}, \mathbf{Ax5}, \mathbf{Ax6}\} \models (\mathsf{f}_{mk}$ is a bijection $\mathsf{f}_{mk} : {}^nF \longrightarrow {}^nF)$.

(ii) $\{\mathbf{Ax1}, \mathbf{Ax5}\} \models (\mathsf{f}_{mk}$ is a (possibly) partial one-to-one function).

(iii) $\{\mathbf{Ax1}, \mathbf{Ax5}, \mathbf{Ax6}\} \models (\mathsf{f}_{mm} = \mathrm{Id}, \mathsf{f}_{mk} = \mathsf{f}_{km}^{-1}, \mathsf{f}_{mk} = \mathsf{f}_{mh} \circ \mathsf{f}_{hk})$.

Proof: That f_{mk} is one-to-one follows from Claim 2.3.7. That f_{mk} is defined everywhere and is onto ${}^nF$ follows from **Ax6**. $\mathsf{f}_{mm} = \mathrm{Id}$, $\mathsf{f}_{mk} = \mathsf{f}_{km}^{-1}$ and $\mathsf{f}_{mk} \supseteq \mathsf{f}_{mh} \circ \mathsf{f}_{hk}$ follow from the definition of the world-view transformation relations. Assume $\mathbf{Ax1}, \mathbf{Ax5}, \mathbf{Ax6}$, let $p \in {}^nF$, and $\mathsf{f}_{mk}(p) = q$, i.e. $w_m(p) = w_k(q)$. By **Ax6** there is $p' \in {}^nF$ such that $w_m(p) = w_h(p')$. Now $\mathsf{f}_{mh}(p) = p'$ by $w_m(p) = w_h(p')$ and $\mathsf{f}_{hk}(p') = q$ by $w_h(p') = w_m(p) = w_k(q)$. Thus $\mathsf{f}_{mk}(p) = \mathsf{f}_{hk}(\mathsf{f}_{mh}(p))$. ∎

Remark 2.3.9 By Claim 2.3.8 we have that if the set $Wtm \overset{\text{def}}{=} Wtm^{\mathfrak{M}} \overset{\text{def}}{=} \{\mathsf{f}_{mk} : m, k \in Obs^{\mathfrak{M}}\}$ of the world-view transformations is closed under composition $\circ$, then $\langle Wtm, \circ, {}^{-1}, \mathrm{Id}\rangle$ forms a group (under assuming **Ax1**,**Ax5**,**Ax6**). In Def.3.6.11 (p.269) we will define a class GM of models of **Basax**, such that for some $\mathfrak{M} \in \mathsf{GM}$ we have that Wtm is not closed under composition.[78] However, in section 3.9 we will introduce a "symmetry axiom" **Ax□1** and we will see that if $\mathfrak{M} \models \mathbf{Basax} \cup \{\mathbf{Ax\Box 1}\}$, then $\langle Wtm^{\mathfrak{M}}, \circ, {}^{-1}, \mathrm{Id}\rangle$ is a group.

◁

The proof of Prop.2.3.3(viii) consists of noting that every slow-line is the trace of some observer k_1 as seen by m, and that $tr_k(k_1)$ is a straight line. Similarly, the proof of Prop.2.3.3(ix) consists of noting that every photon-line is a trace of some photon ph_1 as seen by m (by **Ax5**), and that $tr_k(ph_1)$ is a photon-line again (by **AxE**). The proofs of Proposition 2.3.3 (vi), (vii) are similar to those of Proposition 2.3.3 (i)–(v), (x). We leave them to the reader.

[78] For more on models $\mathfrak{M}$ of **Basax** in which Wtm is not a group cf. section 3.10. Cf. also [259].

By Claim 2.3.8(ii), in most of the situations we will investigate, f_{mk} will be a *function*. This will remain so, even when we will study refinements of our axiom system **Basax**, or even when we will omit some or most of our axioms, f_{mk} will be at least a *partial function* ${}^nF \supseteq Dom(\mathsf{f}_{mk}) \xrightarrow{\mathsf{f}_{mk}} {}^nF$. Therefore, we would like to use the standard notation $\mathsf{f}_{mk}(p)$ when $p \in {}^nF$ as if f_{mk} were a (partial) *function* symbol. But then (since in our original frame-language f_{mk} is only a relation symbol) we have to define a translation mechanism ensuring that the formulas involving notation like $\mathsf{f}_{mk}(p)$ remain formulas of our frame language. To ensure this we make the following convention.

CONVENTION 2.3.10 We introduced f_{mk} as a binary relation symbol (in the extended version of our frame-language). Since in models of **Basax** it is a function (cf. Prop.2.3.3(v)), we will also use f_{mk} as if it were a unary function symbol. There is a well known practice of doing this; a precise translation algorithm can be found e.g. in Monk [197, pp. 206–210] or Bell-Machover [44, p.97] ("Elimination of function symbols"). However, *later* we want to treat theories where f_{mk}'s will be *only partial* functions. Therefore, instead of the algorithms for translating total functions given e.g. in Monk [197], we want to use a slightly more general translation algorithm suitable for handling partial functions as well, see e.g. Andréka-Németi [27]. This translation is quite intuitive: whenever we write "$\mathsf{f}_{mk}(p)$" we mean "f_{mk} is defined on p, i.e. there is a unique q such that $\langle p, q\rangle \in \mathsf{f}_{mk}$, and $\mathsf{f}_{mk}(p)$ denotes this unique q".

In more detail: Let τ, $\sigma(p)$ be terms and R be a relation symbol like "$=$" or "$\leq$" in our frame language (expanded, for convenience, with the language of the vector space[79] ${}^n\mathbf{F}_2$). Let us recall that p, q are variable symbols ranging over ${}^nF$. Then an atomic formula of the "shape" $\mathsf{f}_{mk}(p) = \tau$ means

$$\exists! q \ (\langle p, q\rangle \in \mathsf{f}_{mk}) \quad \wedge \quad \exists q \ (\langle p, q\rangle \in \mathsf{f}_{mk} \ \wedge \ q = \tau),$$

where q is a new variable and "$\exists!$" means "there is a unique". That is, the new formula says, f_{mk} is defined on the argument p and is a function on $\{p\}$ and $\mathsf{f}_{mk}(p) = \tau$.[80]

Similar convention applies to more general atomic formulas like $R(\mathsf{f}_{mk}(p), \tau)$ or $\sigma\Big(\mathsf{f}_{mk}(p)\Big) = \tau$. In both cases the new formula begins with $\exists! q(\langle p, q\rangle \in \mathsf{f}_{mk})$. E.g. the translated version of the second formula is

$$\exists! q \ (\langle p, q\rangle \in \mathsf{f}_{mk}) \quad \wedge \quad \exists q \ (\langle p, q\rangle \in \mathsf{f}_{mk} \ \wedge \ \sigma(q) = \tau),$$

[79] As we already said, ${}^n\mathbf{F}_2$ formulas are translatable to our frame language.

[80] The first subformula $\exists! q \ \langle p, q\rangle \in \mathsf{f}_{mk}$ means, simply, that $\mathsf{f}_{mk}(p)$ is uniquely defined.

where q does not occur in τ or $\sigma(p)$.

Let Tr denote the "translation function" which we are in the process of defining, which is defined on formulas, and which eliminates *function-symbol* style occurrences of the f_{mk}'s. So far we described how to translate atomic formulas, call them φ_i, possibly containing f_{mk}'s as function symbols to new formulas $Tr(\varphi_i)$ in which f_{mk}'s do not occur as function symbols (and hence $Tr(\varphi_i)$ is truly in our frame language). Now, if we want to translate a *complex formula*, call it ψ, the same way (i.e. eliminate using f_{mk}'s as functions), then first we translate all the atomic formulas φ_i occurring in ψ, and then we put together the translations exactly as ψ was put together. E.g. $Tr(\varphi \wedge \psi) = Tr(\varphi) \wedge Tr(\psi)$, $Tr(\neg\varphi) = \neg Tr(\varphi)$, $Tr(\exists x\varphi) = \exists x(Tr(\varphi))$.

$\lhd$

Now, we turn to characterizing the world-view transformations in models of **Basax**(2). Figures 15 and 16 illustrate these transformations, and give perhaps a hint for why we will call such transformations later "rhombus transformations". Their relationship with the literature (Lorentz transformations, Poincaré transformations) is discussed in §2.9.[81] In Figure 16 the world-view transformation f_{mk} is illustrated in such a way that the world-views of both k and m are drawn in the same copy of ${}^2F$. I.e. k's coordinate system is drawn into m's world-view, cf. also Figure 15.

Before giving the characterization (of the world-view transformations), we cite a theorem from the next chapter.

THEOREM 2.3.11 *Assume* **Basax**. *Let $m, k \in Obs$. Then f_{mk} takes straight lines to straight lines, that is,* $(\forall \ell \in \mathsf{Eucl})\ \mathsf{f}_{mk}[\ell] \in \mathsf{Eucl}$.

We will **prove** the above theorem as Thm.3.1.1 in §3.1. ∎

Throughout, by a *transformation* f (of ${}^n\mathbf{F}$) we mean a function $\mathsf{f} : {}^nF \longrightarrow {}^nF$.[82] By a *photon-preserving* transformation f (of ${}^n\mathbf{F}$) we mean a bijective transformation such that both f and f^{-1} take photon-lines to photon-lines. Further, by a *collineation* f (of ${}^n\mathbf{F}$) we mean a transformation (of ${}^n\mathbf{F}$) which takes

[81] Using that terminology, world-view transformations in models of **Basax** are exactly the Poincaré transformations composed with expansions and with functions induced by field-automorphisms. Cf. Theorem 2.9.4.

[82] Often we write mapping or map instead of transformation like e.g. photon-preserving mapping or linear mapping.

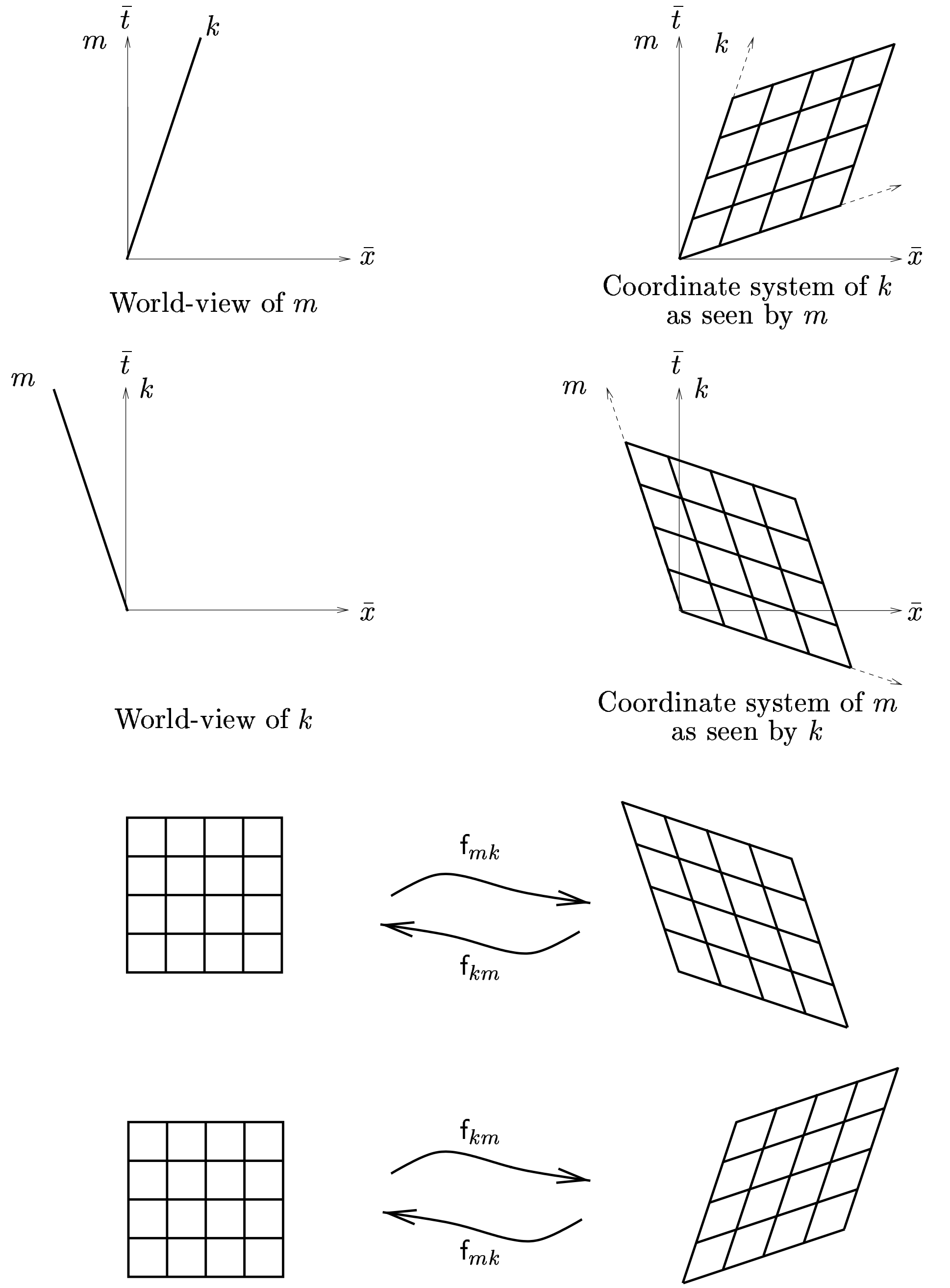


Figure 15: World-view transformation in two space-time dimensions assuming **Basax**.

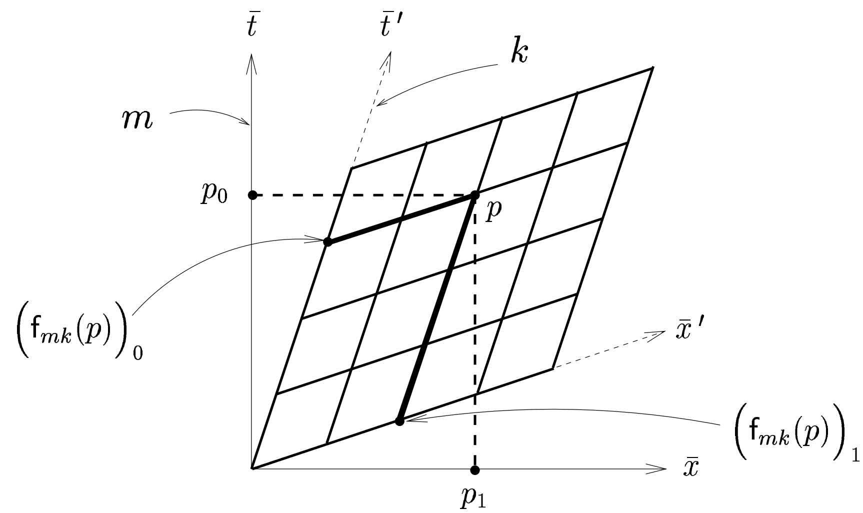


Figure 16: Two-dimensional word-view transformation in **Basax**(2).

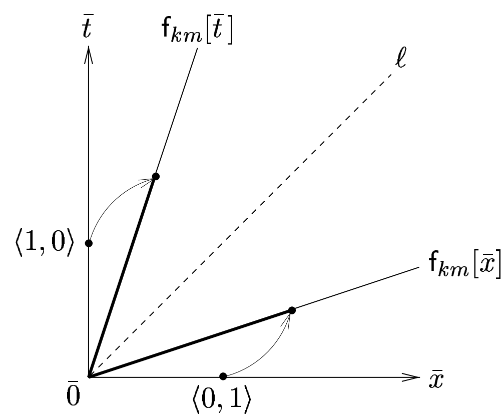


Figure 17: Two-dimensional world-view transformation in **Basax**(2).

straight lines to straight lines, i.e. which preserves Eucl. We recall from the standard literature of algebra that by a *linear transformation* of a vector space ${}^n\mathbf{F}$ we understand a *homomorphism* of the one-sorted vector space ${}^n\mathbf{F}_1$ into itself, cf. e.g. Halmos [122]. (The homomorphisms of the two-sorted vector space ${}^n\mathbf{F}_2$ into itself are something else, cf. Remark 2.3.13.)

THEOREM 2.3.12 (Characterization of world-view transformations in Basax(2).) *Let $\mathfrak{F} = \langle \mathbf{F}, \leq \rangle$ be any ordered field, and $\mathsf{f} : {}^2F \longrightarrow {}^2F$.*

1. *Assume first that $\mathbf{F}$ has no (nontrivial) automorphisms*[83] *and $\mathsf{f}(\bar{0}) = \bar{0}$. Then (i)–(iii) below are equivalent.*

   (i) f *is a world-view transformation in some model of* $\mathbf{Basax}(2)$ *whose ordered field reduct is $\mathfrak{F}$.*[84]

   (ii) f *is like on Figures 15 and 16, i.e. f is a bijective linear transformation of the vector-space ${}^2\mathbf{F}$ such that $\mathsf{f}[\bar{t}\,]$ and $\mathsf{f}[\bar{x}]$ are mirror images of each other w.r.t. a photon-line passing through $\bar{0}$. Moreover, the vectors $\mathsf{f}(\langle 1, 0\rangle)$ and $\mathsf{f}(\langle 0, 1\rangle)$ are of the same length.*[85] *Cf. Figure 17.*

   (iii) f *is a photon-preserving bijective collineation (i.e. f is bijective, takes straight lines and photon-lines to straight lines and photon-lines respectively).*

2. *In the more general case when $\mathbf{F}$ is permitted to have (nontrivial) automorphisms, we still have that (i), (iii) above are equivalent (with each other and) with both (ii)' and (ii)* below:*

   (ii)' $\mathsf{f} = \gamma \circ \mathsf{g}$ *where g is like f was in (ii) above and there is an automorphism $\varphi : \mathbf{F} \longrightarrow \mathbf{F}$ of $\mathbf{F}$ such that $\gamma(p) = \langle \varphi(p_0), \varphi(p_1)\rangle$ for all $p \in F^2$.* [86]

   (ii)* f *is a bijective collineation such that $\mathsf{f}(\langle 1, 0\rangle)$ and $\mathsf{f}(\langle 0, 1\rangle)$ are mirror images of each other w.r.t. a photon-line passing through $\bar{0}$. I.e. f is like on Figures 15–17.*

---

[83] An *automorphism* of a structure is an injective and surjective homomorphism of that structure into itself (cf. p.160 for more detail). Let us note that the property of $\mathfrak{M}$ that "$\mathbf{F}^{\mathfrak{M}}$ has no (nontrivial) automorphisms" cannot be expressed by a set of (first-order) formulas in our frame-language, since this property is not preserved under taking ultrapowers. We also note that the field of reals (real numbers) and the field of rational numbers enjoy this property.

[84] I.e. $(\exists \mathfrak{M} \in \mathsf{Mod}(\mathbf{Basax}(2)))\,[(\exists m, k \in Obs)\ \mathsf{f} = \mathsf{f}_{km}$ and $\mathfrak{F}^{\mathfrak{M}} = \mathfrak{F}]$.

[85] We use $p_0^2 + p_1^2$ for the length of $p \in {}^2F$. (We do not take square roots because no axiom ensures their existence *yet*.)

[86] To help the reader's intuition we note that $\gamma \circ \mathsf{g}$ on the points with rational coordinates, like e.g. $p = \langle 1, 1\rangle$, is the same as g. (Let us recall that, for any $\mathfrak{F}$ the rational numbers can be considered as elements of $\mathfrak{F}$.)

3. *If in 2 above we drop the assumption* $\mathsf{f}(\bar{0}) = \bar{0}$*, then (ii)' and (ii)** *have to be changed to (ii)" and (ii)****, respectively, below.*

(ii)" f *is a composition of a function* f' *which is like in (ii)' and a translation, i.e.* $\mathsf{f} = \mathsf{f}' \circ \tau$ *where* f' *is exactly like* f *was in (ii)' and* $\tau : {}^2F \longrightarrow {}^2F$ *is a translation*[87]*.*

(ii)** f *is a bijective collineation such that* $\mathsf{f}(\langle 1,0\rangle)$ *and* $\mathsf{f}(\langle 0,1\rangle)$ *are mirror images of each other w.r.t. a photon-line passing through* $\mathsf{f}(\bar{0})$*.*

Before proving Thm.2.3.12, we include the following two remarks.

Remark 2.3.13 The bijective collineations of ${}^n\mathbf{F}$ came up in the above theorem (and they will keep on coming up later, too). Therefore, we note that the $\bar{0}$-preserving *bijective collineations* are exactly the *automorphisms* of the two-sorted version ${}^n\mathbf{F}_2$ of the vector space ${}^n\mathbf{F}$.

Another characterization (of the bijective collineations preserving $\bar{0}$) is that they are exactly the maps obtainable as a composition of a bijective linear transformation (i.e. an automorphism of the one-sorted version ${}^n\mathbf{F}_1$ of the vector-space) and a map induced by an automorphism of the field $\mathbf{F}$. Cf. Lemma 3.1.6 on p.163.

$\triangleleft$

Remark 2.3.14 The above theorem (characterizing the f_{mk}'s) involves *field automorphisms*. Intuitive (as well as mathematical) discussion of field automorphisms with examples, pictures, and their roles in **Basax** models, in collineations and in the world-view transformations (the f_{mk}'s) will be discussed in a separate item in §3 in a later version of the present work. We note that a partial version of the just promised dicussion (of field automorphisms etc.) can be found in the 1997 October 27 version of the present work, pp. 25–26. The just promised discussion will include e.g. the following: (i) In any **Basax** model $\mathfrak{M}$, if $\mathfrak{F}^{\mathfrak{M}}$ is Archimedean[88] and Euclidean then the f_{mk}'s are affine transformations[89].[90] (ii) There are **Basax**(2)

[87] A translation is a map of the form $\langle p+q : p \in {}^nF\rangle$, where $q \in {}^nF$ is fixed.

[88] $\mathfrak{F}$ is *Archimedean* iff to each positive $x \in F$ there is a natural number $\varrho \in \omega$ which is larger than x, i.e. $\varrho > x$. (We note that for every ordered field the set ω of the natural numbers can be considered as a subset of the ordered field, or in more careful wording ω is embeddable into the ordered field in a natural way.) For brevity, by *"Archimedean field"* we mean "Archimedean ordered field". We further note that $\mathfrak{F}$ is *Archimedean* iff it is embeddable into (i.e. isomorphic to a subfield of) $\mathfrak{R}$.

[89] Affine transformations are linear transformations composed with translations, as we will discuss this in §2.9.

[90] For undefined terminology the reader is referred to the Index.

models with Archimedean ordered field reducts containing non-betweenness preserving hence non-continuous and not affine world-view transformations.[91] (iii) If, to **Basax**(2), we add the axiom that the f_{mk}'s are betweenness preserving then we will obtain a strictly stronger and natural version (of **Basax**(2)). (For $n > 2$, **Basax** implies that the f_{mk}'s are betweeness preserving, cf. Prop.6.6.5 on p.1028). (iv) We guess that in **Basax** models the assumption that the f_{mk}'s are betweenness preserving implies that they are continuous, but we did not have time to check this. (v) There are **Basax** models with Euclidean ordered field reducts in which some of the f_{mk}'s are not affine, for every $n \geq 2$. (vi) There are **Basax** models with Euclidean ordered field reducts where some of the f_{mk}'s are continuous collineations which are still not affine transformations. This means that if we add to **Basax** continuity of the f_{mk}'s as an extra axiom we still cannot force all the f_{mk}'s to be affine. (vii) If $n > 2$ and $\mathfrak{F}$ is a reduct of a **Basax** model then all the automorphisms of $\mathbf{F}$ are order preserving, i.e. using a standard notation of universal algebra $Aut(\mathbf{F}) = Aut(\mathfrak{F})$, cf. Corollary 6.7.12 on p.1142.

$\lhd$

Proof of Thm.2.3.12: The main idea of the proof is illustrated in Figure 18. We note that a more carefully polished proof will be included at a later stage of development.

Assume first that $\mathbf{F}$ has no (nontrivial) automorphisms and $\mathsf{f}(\bar{0}) = \bar{0}$.

(i) $\Rightarrow$ (iii): f is a bijection and photon-preserving by Prop.2.3.3(v),(ix); and f is a collineation by Thm.2.3.11.

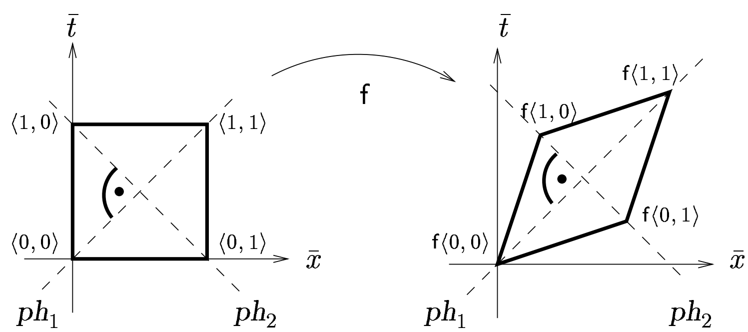


Figure 18: This is the main idea of the proof of Thm.2.3.12.

[91] We note that "affine $\Rightarrow$ continuous $\Rightarrow$ betweenness preserving" (for f_{mk}'s of **Basax** models if $\mathfrak{F} = \mathfrak{R}$).

(iii) $\Rightarrow$ (ii): Since **F** has no (nontrivial) automorphisms and $\mathsf{f}(\bar{0}) = \bar{0}$, a bijective collineation is a linear transformation, cf. Remark 2.3.13. If we do not assume that **F** has no (nontrivial) automorphisms – but we still assume $\mathsf{f}(\bar{0}) = \bar{0}$ – then f is like in (ii)', i.e. f is a composition of a linear transformation with a map coming from a field automorphism, cf. Remark 2.3.13. If we do not assume $\mathsf{f}(\bar{0}) = \bar{0}$ either, then f is like in (ii)", i.e. we have to compose with a translation also. The main idea of the rest of the proof is illustrated in Figure 18.

For any two distinct points $p, q \in {}^2F$, $\overline{pq}$ denotes the Euclidean line containing both p and q.

Consider the two photon-lines (in Figure 18) illustrated on the left-hand copy of ${}^2F$, they are $\overline{\langle 0,0\rangle\langle 1,1\rangle}$ and $\overline{\langle 1,0\rangle\langle 0,1\rangle}$. These two photon-lines are taken to $\overline{\mathsf{f}\langle 0,0\rangle\mathsf{f}\langle 1,1\rangle}$ and $\overline{\mathsf{f}\langle 1,0\rangle\mathsf{f}\langle 0,1\rangle}$.[92]

These last two are photon-lines because f is photon-line preserving. They cannot be parallel, because the original two photon-lines are not parallel. Thus they have to be orthogonal (in the usual Euclidean sense) to each other because we are in dimension 2. The two pairs of lines $\overline{\mathsf{f}\langle 0,0\rangle\mathsf{f}\langle 1,0\rangle}$, $\overline{\mathsf{f}\langle 0,1\rangle\mathsf{f}\langle 1,1\rangle}$ and $\overline{\mathsf{f}\langle 0,0\rangle\mathsf{f}\langle 0,1\rangle}$, $\overline{\mathsf{f}\langle 1,0\rangle\mathsf{f}\langle 1,1\rangle}$ are parallel because the original lines are so. Thus the square with vertices $\langle 0,0\rangle$, $\langle 1,0\rangle$, $\langle 0,1\rangle$, $\langle 1,1\rangle$ is taken to the parallelogram with vertices $\mathsf{f}\langle 0,0\rangle$, $\mathsf{f}\langle 1,0\rangle$, $\mathsf{f}\langle 0,1\rangle$, $\mathsf{f}\langle 1,1\rangle$. The latter parallelogram is indeed a rhombus, because its diagonals are orthogonal. This implies (ii).

(ii) $\Rightarrow$ (i): We will prove this in the next section, as Thm.2.4.2. ∎

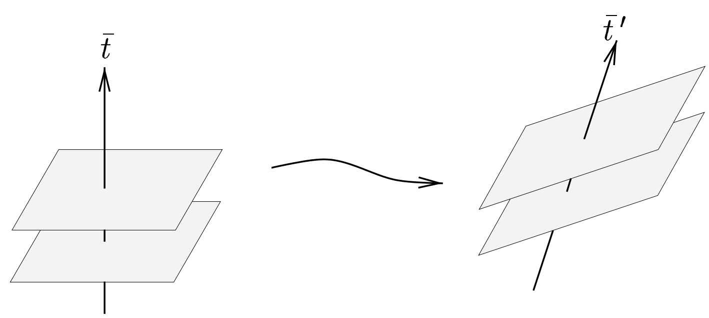


Figure 19: World-view transformation in three space-time dimensions, cf. Figures 15 and 16.

[92]Sometimes we write $\mathsf{f}p$ for $\mathsf{f}(p)$ like $\mathsf{f}\langle 0,0\rangle$ for $\mathsf{f}(\langle 0,0\rangle)$.

A large part of Thm.2.3.12 remains true in higher dimensions (i.e. for $\mathbf{Basax}(n)$ in place of $\mathbf{Basax}(2)$), e.g., under a mild extra assumption[93] on $\mathfrak{F}$, (i) and (iii) remain equivalent, cf. Thm.3.6.16 on p.273. We now generalize the kind of transformations described in (ii) (of Thm.2.3.12) to arbitrary dimensions $n \geq 2$; we will call such transformations "rhombus transformations". Cf. Figures 19, 20.

In dimension 2, the trace of an observer, as everything else, is in the plane of the time-axis and the $\bar{x}$-axis. In higher dimensions this is not so. Below we will single out a special case in higher dimensions that resembles the 2-dimensional case, and we will call it "standard configuration".

Notation 2.3.15

(i) For every $i \in n$, $1_i \in {}^nF$ denotes the <u>*unit vector*</u> pointing in direction of the i'th coordinate axis $\bar{x}_i$, that is,

$$1_i \stackrel{\text{def}}{=} \langle \underbrace{0,\ldots,0}_{i}, 1, \underbrace{0,\ldots,0}_{(n-i-1)} \rangle .$$

Usually, we will write

$$1_t, 1_x, 1_y, 1_z \quad \text{for} \quad 1_0, 1_1, 1_2, 1_3, \quad \text{respectively.}$$

(ii) $\mathsf{Plane}(\bar{t}, \bar{x})$ denotes the plane determined by lines $\bar{t}$ and $\bar{x}$, that is,

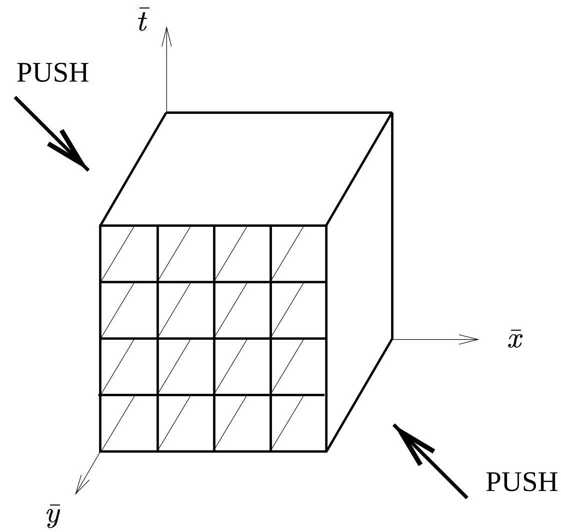
$$\mathsf{Plane}(\bar{t}, \bar{x}) \stackrel{\text{def}}{=} \{\, p \in {}^nF \;:\; (\forall 1 < i \in n)\; p_i = 0 \,\} .$$

I.e. $\mathsf{Plane}(\bar{t}, \bar{x}) = F \times F \times {}^{n-2}\{0\}$.

$\lhd$

We are ready to define standard configuration. We will write about the intuitive meaning of standard configuration after the definition.

[93] This extra assumption is that the square roots of positive elements exist in $\mathfrak{F}$ (i.e. that $\mathfrak{F}$ is Euclidean).



Preparation for drawing 3-dimensional f_{mk}.

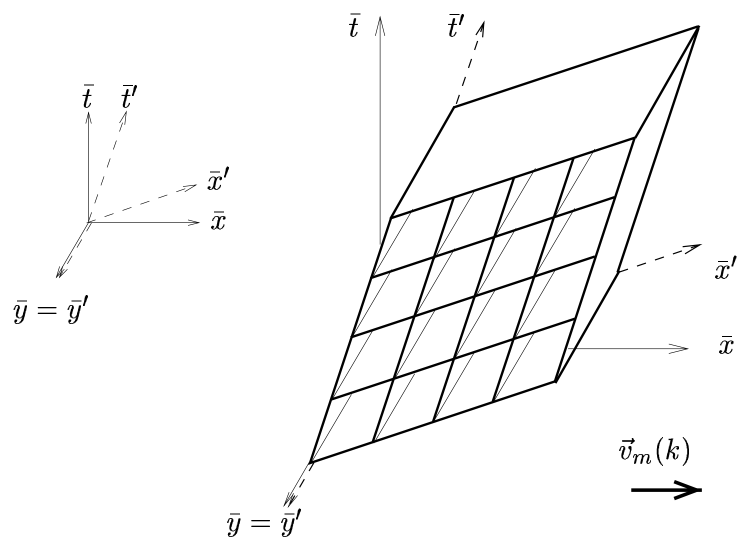


Figure 20: 3-dimensional world-view transformation f_{mk} in "standard" configuration, cf. Figures 14, 16. For the notion of standard configuration cf. Def.2.3.16 and Figures 21, 22.

Definition 2.3.16 (Standard configuration)

(i) Let $\mathfrak{M}$ be a frame model. Let $m, k \in Obs$. We say that m and k are in <u>*standard*</u> configuration if

$$\mathsf{f}_{mk}[\mathsf{Plane}(\bar{t},\bar{x})] = \mathsf{Plane}(\bar{t},\bar{x}) \quad \text{and} \quad (\forall 1 < i \in n)(\exists 0 < \lambda \in F)\mathsf{f}_{mk}(1_i) = \lambda \cdot 1_i.$$

(ii) We say that m and k are in <u>*strict standard*</u> configuration if in addition to the above we have $\mathsf{f}_{mk}(1_x)_x > 0$.

See Figures 21, 22. Cf. also Figure 20.

$\lhd$

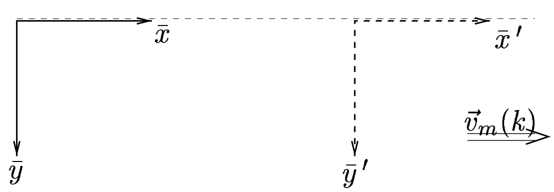


Figure 21: Standard configuration. Here $\bar{x}$ and $\bar{y}$ are space axes of m while $\bar{x}\,'$ and $\bar{y}\,'$ are space axes of k. The spatial coordinate system $\{\bar{x}\,', \bar{t}\,'\}$ of k is moving relative to that of m.

The next proposition says that m and k are in standard configuration iff they meet at $\bar{0}$, they see each other moving in direction 1_x (forwards or backwards), and they see each other's unit-vectors other than $\bar{t}, \bar{x}$ as perhaps shrinking or growing but pointing in the same direction.

PROPOSITION 2.3.17 *Assume* **Ax1** – **Ax5**. *Then m and k are in standard configuration iff (i)-(iv) below hold.*

(i) $\mathsf{f}_{mk}(\bar{0}) = \bar{0}$

(ii) $tr_m(k), tr_k(m) \subseteq \mathsf{Plane}(\bar{t},\bar{x})$

(iii) *If* $v_m(k) = 0$ *then* $\mathsf{f}_{mk}[\bar{x}] \subseteq \mathsf{Plane}(\bar{t},\bar{x})$

(iv) *Let* $1 < i \in n$. *Then* $\mathsf{f}_{mk}(\underbrace{0,\dots,0}_{i}, 1, 0, \dots, 0) = \mathsf{f}_{mk}(\underbrace{0,\dots,0}_{i}, \lambda, 0, \dots, 0)$ *for some* $0 < \lambda \in F$.

$\triangleleft$

We note that **Ax3** and **Ax5** in the above Proposition 2.3.17 can be replaced with their much weaker forms $\mathbf{Ax3_0}$ and $\mathbf{Ax(5nop)^{-+}}$ i.e. with $(\forall k \in Obs)[tr_m(k) = \emptyset$ or $tr_m(k) \in G]$ and with $(\exists c > 0)(\forall \ell)[ang^2(\ell) < c \Rightarrow (\exists k \in Obs)\ell = tr_m(k)]$ respectively. Thus in later parts when we deal with weaker axiom systems, (i)-(iv) in Proposition 2.3.17 will still give an equivalent definition of standard configurations (because the weaker axioms that we mentioned will be included in all our weak axiom systems).

We note that being in standard configuration is a symmetric relation, i.e. if m and k are in standard configuration, then k and m are also in standard configuration. Very often it simplifies the discussion if we assume that m and k are in standard configuration. (Sometimes, in intuitive discussions we may assume that m and k are in standard configuration without explicitly mentioning this.)

The reader is invited to contemplate Figures 15–20. They all represent cases of a natural kind of transformations $\mathsf{f} : {}^nF \longrightarrow {}^nF$ which we will call rhombus transformations, their set will be denoted by *Rhomb*, cf. Def.2.3.18 below. They are generalizations of the kind of functions occurring in Thm.2.3.12(ii); they will be strongly related to what we will call Lorentz transformations in standard configuration, cf. Thm.2.9.6 on p.156.

Now we turn to a common generalization of the transformations illustrated in Figures 15–20.

Definition 2.3.18 (Rhombus transformation, *Rhomb*)
Assume $\mathfrak{F}$ is an ordered field and $n \geq 2$.
By a *rhombus transformation* (of ${}^n\mathfrak{F}$)[94] we understand a bijective linear transformation $\mathsf{f} : {}^nF \longrightarrow {}^nF$ of the vector space ${}^n\mathbf{F}$ satisfying (i)-(iii) below.

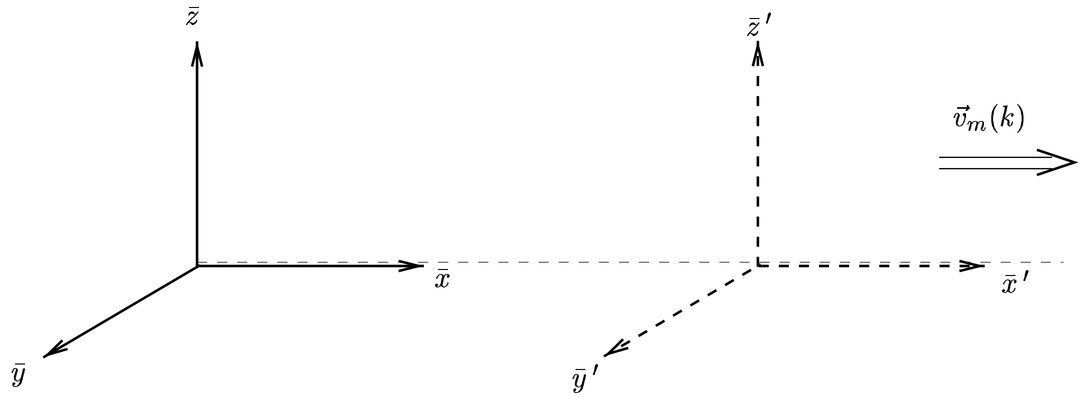
(i) $\mathsf{f}(1_t)$ and $\mathsf{f}(1_x)$ are both in $\mathsf{Plane}(\bar{t}, \bar{x})$ and are mirror images of each other w.r.t. a photon-line ℓ with $\bar{0} \in \ell \subseteq \mathsf{Plane}(\bar{t}, \bar{x})$.[95]

(ii) $(\forall 1 < i \in n)$ $(\mathsf{f}(1_i) = \lambda \cdot 1_i$, for some $0 < \lambda \in F)$.[96]

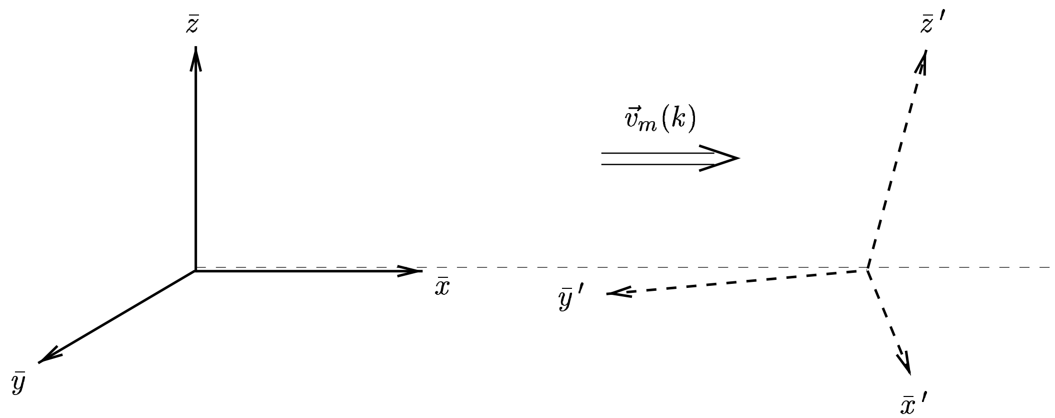
[94] Occasionally we mention this symbol ${}^n\mathfrak{F}$. Since $\mathfrak{F}$ is an algebraic structure, so is its Cartesian power ${}^n\mathfrak{F}$ (which happens to be a partially ordered ring). However, in this work, we think of ${}^n\mathfrak{F}$ as a *partially ordered* vector space $\langle {}^n\mathbf{F}, \leq \rangle$ where the partial ordering $\leq$ of ${}^n\mathfrak{F}$ is induced by $\leq^{\mathfrak{F}}$ of $\mathfrak{F}$ in the usual, "Cartesian power" style. (In particular, the coordinate axes like $\bar{t}$ are linearly ordered by this partial order $\leq$ of ${}^n\mathfrak{F}$.)

[95] This mirror image part means that if $\mathsf{f}(1_t) = \langle p_0, p_1, 0, \ldots, 0 \rangle$ then either $\mathsf{f}(1_x) = \langle p_1, p_0, 0, \ldots, 0 \rangle$ or $\mathsf{f}(1_x) = \langle -p_1, -p_0, 0, \ldots, 0 \rangle$.

[96] For completeness we note that more on the choice of λ can be found in §§ 3.2, 3.5. However, we emphasize that the above definition makes sense (i.e. is complete) without any further discussion of the choice of λ.



Standard configuration



A nonstandard configuration, which in "animated" form is drawn below.
The picture shows a spaceship flying in the indicated "nonstandard" direction.

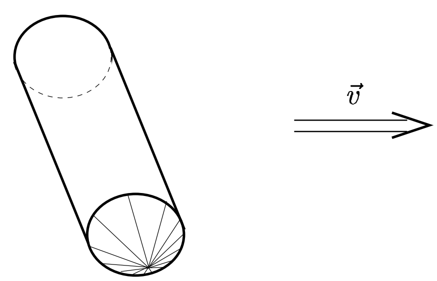


Figure 22: A standard, and a nonstandard configuration.

(iii) f preserves the set of photon-lines, i.e. $(\forall \ell \in \mathsf{PhtEucl})\ \mathsf{f}[\ell] \in \mathsf{PhtEucl}$.

Condition (iii) is needed only if $n > 2$. The role of (iii) is to regulate the choice of λ in (ii).

$Rhomb = Rhomb(n, \mathfrak{F})$ denotes the set of rhombus transformations of ${}^n\mathfrak{F}$.

$\lhd$

We note that rhombus transformations will play a central role in proving that **Basax**(n) is consistent, cf. §2.4 and §3.5.

Remark 2.3.19 Assume that square roots of positive elements of $\mathfrak{F}$ exist, that is $(\forall 0 < x \in F)(\exists y \in F)\ x = y^2$. Assume $n > 2$. In chapter 3 we will see that[97] for any slow-line ℓ with $\bar{0} \in \ell \subseteq \mathsf{Plane}(\bar{t}, \bar{x})$ there is a rhombus transformation taking $\bar{t}$ to ℓ. The idea will be that first we choose $\mathsf{f}(1_t)$ and $\mathsf{f}(1_x)$ so that they are mirror-images of each other like in (i) of Def.2.3.18, and $\mathsf{f}(1_t)$ is on ℓ. $\mathsf{Plane}(\bar{t}, \bar{x}_i)$ is the plane determined by $\bar{t}$ and $\bar{x}_i$ in an analogous way as $\mathsf{Plane}(\bar{t}, \bar{x})$ was defined. Then for every $i \in n$, $i > 1$ there is a unique λ making (ii) (of Def.2.3.18) true so that photon-lines in $\mathsf{Plane}(\bar{t}, \bar{x}_i)$ are mapped to photon-lines. These now fix our linear transformation f. Finally, we have to check that (iii) of Def.2.3.18 is satisfied, i.e. that *every* photon-line is mapped (by f) to a photon-line, and not only those in $\mathsf{Plane}(\bar{t}, \bar{x}_i)$.

$\lhd$

We note that if for observers m and k we have $\mathsf{f}_{km} \in Rhomb$ then m and k are in standard configuration.

In §2.9 we will recall from the literature the so called Lorentz transformations. A special case of the latter will be called Lorentz transformations in standard configuration. The elements of the above introduced *Rhomb* will turn out to be generalizations of Lorentz transformations in standard configuration, cf. Thm.2.9.6 on p.156. At this point we would like to suggest that the reader go through Figures 15–20 and compare them with the definition of *Rhomb*.

Connections between the world-view transformations f_{mk} and Lorentz transformations will be discussed in §2.9. It will turn out that for establishing these connections it is enough to assume **Basax**. Roughly speaking, these connections will say that every f_{mk} is a composition of a *Lorentz transformation*, an *"expansion"*, and a map induced by a *field automorphism*.

[97] Cf. §3.2, cf. also Lemma 3.8.46.

***

Now we turn to listing some (logical) properties of **Basax** as a first-order theory.

According to our Convention 2.2.1(ii), **Basax**(2) denotes **Basax** in the 2-dimensional case. Next, in section 2.4, we will see that **Basax**(2) is *consistent*, that is, there exist frame models satisfying **Basax**(2). In sections 3.2, 3.5 we will see that **Basax**(3) is also consistent, and that generally, **Basax**(n) is consistent for all $n \geq 3$ (cf. Definition 3.5.5, Thm.3.5.6).

The next two properties "count as logical" in the sense that the above property (consistency) concerns the existence of models while the next two properties concern existence of special kinds of models (namely models with faster than light observers, and models with special ordered field reducts).

We will see, in section 2.4, that there are models of **Basax**(2) in which there *are* observers moving *faster than light*, while if $n > 2$ then there are no such models of **Basax**(n) (i.e. for $n > 2$, $\mathbf{Basax}(n) \models (\forall m, k \in Obs) v_m(k) < 1$, see Thm.3.4.1, while $\mathbf{Basax}(2) \not\models (\forall m, k \in Obs) v_m(k) < 1$).

We will see that *every* linearly ordered field is the ordered field reduct of some model of **Basax**(2), while the ordered field reducts of **Basax**(3) are exactly the Euclidean ordered fields (i.e. those in which square roots of positive elements exist). For $n > 3$, we do not know exactly which ordered fields occur as ordered field reducts of **Basax**(n) models, but we know that all Euclidean ordered fields do occur.

An axiom system **Th** is called *independent* if no axiom of **Th** follows from the rest of **Th**, i.e. if $\mathbf{Th} \setminus \{\mathbf{Ax}\} \not\models \mathbf{Ax}$ for all $\mathbf{Ax} \in \mathbf{Th}$. **Basax**($n$) is independent for every $n > 1$.[98] We omit the proof of this, but cf. [16]. To make this independence statement about **Basax** precise we have to make the formulation of **Ax5** a little-bit more careful. Namely we have to replace the subformula $ang^2(\ell) < 1$ with the formula $(\ell \in \mathsf{Eucl} \;\wedge\; ang^2(\ell) < 1)$; similarly for the subformula $ang^2(\ell) = 1$.

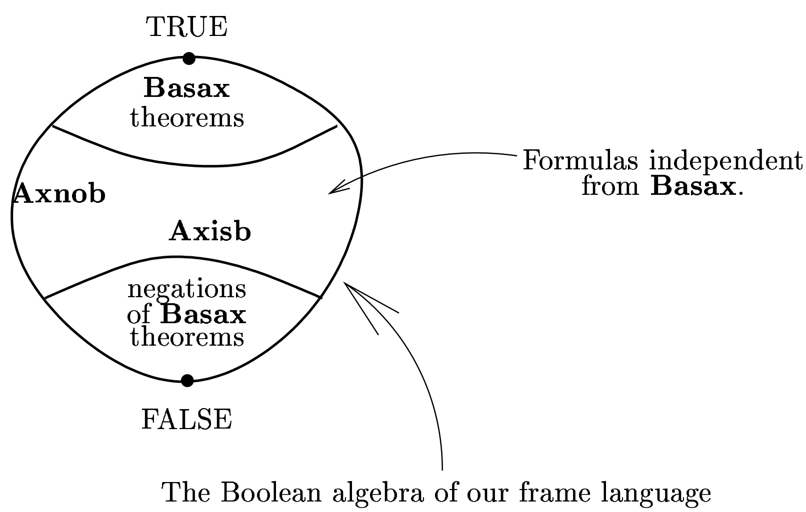
We now list some further *logical* properties of **Basax**.[99] We already stated that **Basax** is consistent and independent. We will classify the models of **Basax**, and we will see that there are continuum many non-elementarily equivalent models of **Basax** (such that they have the same ordered field reduct $\mathfrak{F}$), cf. Thm.3.8.18. Hence, **Basax** is not complete (even if we add the theory $\mathsf{Th}(\mathfrak{F})$, for any choice of $\mathfrak{F}$); **Basax** is non-categorical in any cardinality even if we fix the reduct $\mathfrak{F}$ (it has non-isomorphic

[98] Let **Basax**$'$ be the axiom system obtained from **Basax** by replacing **Ax2** and **Ax3** with a single axiom $(\forall h \in Obs \cup Ph)\ tr_m(h) \in G$. Then **Basax**$'$(2) is independent, **Basax**$'$(3) is not independent, and we do not know whether **Basax**$'$(n) is independent for $n > 3$. These properties of **Basax**$'$ are proved in [16], taken together with Thm.3.6.17.

[99] For the notions from logic used below (like categorical theory, complete theory, theory generated by a set of axioms etc.) we refer the reader to §3.8 and to §7.

models [with a common ordered field reduct] of the same cardinality, for each infinite cardinality), cf. Thm.3.8.18. We will prove that the theory generated by **Basax**, i.e. the set of first-order consequences of **Basax**, is undecidable cf. §7. This will also prove that **Basax** is not complete (hence not categorical in any cardinality, since its models are infinite), because **Basax** is finite. We will define some natural axioms, call them **Axnob** [100] and **Axisb** and we will show that **Basax** ∪ {**Axnob**} is complete[101] (cf. §3.8), while **Basax** ∪ {**Axisb**} is hereditarily undecidable, thus no finite extension of it can be complete, cf. §7.[102] Moreover, the conclusion of Gödel's second incompleteness theorem also applies to **Basax** ∪ {**Axisb**}.

Definability issues related to **Basax** and its variants will be discussed in §6.7. In more detail, in §6.6 we will see that **Basax** admits a nice "duality theory" acting between models of **Basax** and certain geometries.[103] This duality theory involves, among others, "representation theorems" (in the Tarskian sense[104]). So in a sense **Basax** admits a kind of "geometrization"[105]. Studying this duality theory will lead us (in §6.7) to *definability* properties of **Basax** (and its geometric counterpart) in the sense of the chapter of model theory called *definability theory*.



[100] To be precise, we note that **Axnob** is only a schema of axioms (as opposed to being a single axiom).

[101] and also *categorical* in *some* natural sense to be made precise in §3.8

[102] The name **Axnob** refers to the fact that this axiom says, among others, that there are no accelerated bodies. On the other hand **Axisb** refers to the fact that the key part of this axiom says that there do exist accelerated bodies.

[103] For this we first add a few natural axioms to **Basax**, and then we find that this duality theory works already for "fragments" and variants of **Basax**.

[104] bringing together Tarski's approach to geometry and his approach to algebraic logic.

[105] This is analogous with "algebraization" of logics in Tarskian algebraic logic.

Summing up:

- **Basax** is *independent*.
- **Basax** is *consistent* (cf. §§ 2.4, 3.2, 3.5).
- **Basax** has *many non-elementarily equivalent* models (even if we add to it $Th(\mathfrak{R})$), cf. Thm.3.8.18.
- We will give a classification of the models of **Basax** (cf. §3.6).
- The first-order theory T(**Basax**) generated by **Basax** is *undecidable* hence not complete, cf. §7.
- Adding an extra axiom-schema can make T(**Basax**) *complete* hence *decidable*, since **Basax** is finite, cf. §3.8.
- Adding a different extra axiom can make T(**Basax**) *hereditarily undecidable* hence hereditarily not complete. The conclusions of *Gödel's incompleteness theorems* apply to the so extended version of **Basax**. Cf. §7.
- Adding an extra axiom-schema makes **Basax** equivalent with the standard, "textbook version" of "Einsteinian" special relativity, cf. §§ 2.8, 2.9, 3.8, 6.2, 6.6.
- Other distinguished versions like the Reichenbach-Grünbaum version of relativity can (and will) be formalized in first-order logic (and compared with the Einsteinian version) by appropriately modifying **Basax** (cf. §§ 3.4.2, 4.4, 6.6 and the section [on "Reichenbachian relativity"] of §4 (§4.5)).

2.4 Models for Basax in dimension 2

In this section we show that **Basax**(2) is consistent, via defining a frame model $\mathfrak{M}$ and showing that $\mathfrak{M} \models$ **Basax**(2). We will also give a model of **Basax**(2), in which there are faster than light observers.

First, let us have some *intuitive* considerations on why **Basax**(2) is consistent. (Later we will give a formal proof.) The main reason why **Basax**(2) is consistent is the following:

($\star$) for each slow-line ℓ there is a photon-preserving bijective collineation taking $\bar{t}$ to ℓ.

The reader is invited to study Figures 15–20 (pp. 63–70) to convince himself that ($\star$) is true, and then use ($\star$) the following way to show that **Basax**(2) is consistent.

(I) Assume, we are given a "partial model"

$$\mathfrak{M} = \langle (B;\ \{m_0\}, Ph, Ib), \mathfrak{F}, G;\ \in, W \rangle\,,$$

which satisfies all the axioms in **Basax** *except* for the observer-part of **Ax5**. Let us use the notation **Ax5** = **Ax5(Obs)** $\wedge$ **Ax5(Ph)**. Then

$$\mathfrak{M} \models (\mathbf{Ax1}\text{–}\mathbf{Ax4}, \mathbf{Ax5(Ph)}, \mathbf{Ax6}, \mathbf{AxE})\,.$$

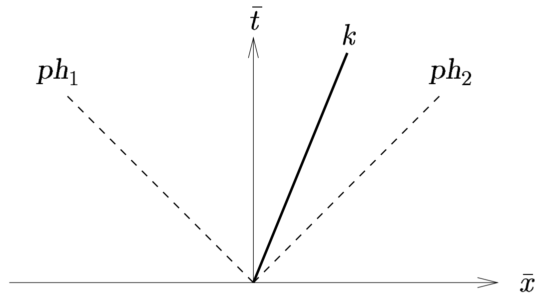
Assume further $\mathfrak{F} = \mathfrak{R}$, and that

$$(\forall \ell \in G)(\exists b \in Ib)\ \ell = tr_{m_0}(b)\,.$$

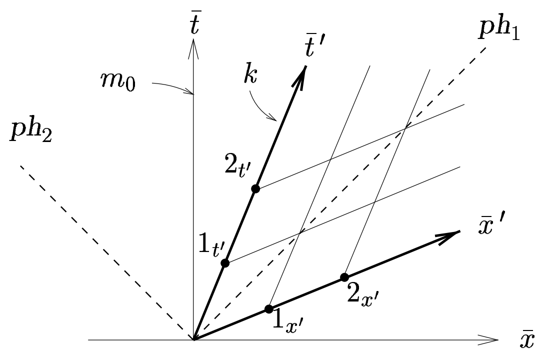
Constructing such a partial model is easy, and is left to the reader.

(II) Next, we would like to add new observers to $\mathfrak{M}$ so that eventually **Ax5(Obs)** would become true without destroying validity of the other axioms (hence **Basax** would become true).

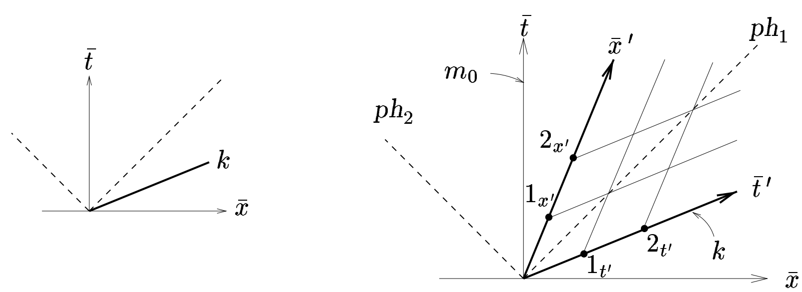
Clearly, in $\mathfrak{M}$ we do have a world-view function $w_{m_0} : {}^2\mathrm{R} \longrightarrow \mathcal{P}(B)$, to begin with. From this world-view function we will construct world-views for new observers. Let us pick randomly $k \in Ib$ such that $v_{m_0}(k) \neq 1$. Now, we would like to raise k to the level of being an observer. Assume, m_0 sees this:



Our task is to choose the world-view of k such that, among others, **AxE** remains valid, i.e. that k observes all photons moving with speed 1. Following Figures 15–20, let us choose k's world-view like this (the figure shows k's coordinate system as seen by observer m_0):



where $\bar{t}\,'$ and $\bar{x}\,'$ are the time-axis and $\bar{x}$-axis, respectively, of k. We note that if k moves faster than light relative to m_0 (i.e. if $v_{m_0}(k) > 1$) then k's coordinate system (as seen by m_0) is like in the following picture:



Now clearly, k will observe photon ph_1 moving with speed 1 and the same applies to ph_2. Then one can check that for these two particular observers m_0, k our axiom

AxE holds, i.e. both k and m_0 will observe all photons moving with speed 1. One can check that for the extended model

$$\mathfrak{M}' := \langle (B;\ \{m_0, k\}, Ph, Ib), \mathfrak{F}, G;\ \in, W^+ \rangle$$

we have **Ax1**–**Ax4**, **Ax5(Ph)**, **Ax6**, **AxE** still valid. Here, W^+ denotes the extension of W with the world-view function w_k of the <u>new</u> observer k.

To complete the "intuitive" proof, one does the above extension <u>not only</u> with a single $k \in Ib$ but with the class $K = \{ k \in Ib : v_{m_0}(k) \neq 1 \}$ of all potential candidates for being an observer. This will make **Ax5(Obs)** true. We note that the condition $\bar{0} \in tr_m(k)$ was <u>not</u> needed in our construction of w_k.

In passing we note that in the above constructed model faster than light observers exist. It is easy to modify the construction in such a way that faster than light observers will not exist in the modified model. This modification begins with adding to statement $(\star)$ above that the photon-preserving bijective collineation in question takes slow-lines to slow-lines. The rest of the modifications are straightforward, we leave them to the reader.

<u>*END of Intuitive Idea of Proof.*</u>

Let us turn to precise proofs.

Let P be a function that to each $\ell \in \mathsf{Eucl}(2, \mathfrak{R})$ associates a pair of two distinct points lying on ℓ. We will denote $P(\ell)$ by $\langle o_\ell, t_\ell \rangle$. To each such function P, we will define two frame models, $\mathfrak{M}_0^P$ and $\mathfrak{M}_1^P$. These two frame models will be very similar in spirit, but in $\mathfrak{M}_0^P$ we have as few observers as possible, while in $\mathfrak{M}_1^P$ there will be an observer on each line (with angle $\neq 1$, cf. Prop.2.3.3(iii)).

First we define $\mathfrak{M} \stackrel{\text{def}}{=} \mathfrak{M}_0^P \stackrel{\text{def}}{=} \langle (B;\ Obs, Ph, Ib), \mathfrak{F}, G;\ \in, W \rangle$, where

$\mathfrak{F} \stackrel{\text{def}}{=} \mathfrak{R}$, the ordered field of real numbers,

$G \stackrel{\text{def}}{=} \mathsf{Eucl}(2, \mathfrak{R})$, the set of straight lines over $\mathfrak{R}$,

$Obs \stackrel{\text{def}}{=} \{ \ell \in \mathsf{Eucl}(2, \mathfrak{R}) \ :\ ang^2(\ell) < 1 \}$,

$Ph \stackrel{\text{def}}{=} \{ \ell \in \mathsf{Eucl}(2, \mathfrak{R}) \ :\ ang^2(\ell) = 1 \}$,

$B \stackrel{\text{def}}{=} Ib \stackrel{\text{def}}{=} Obs \cup Ph = \{ \ell \in \mathsf{Eucl}(2, \mathfrak{R}) \ :\ ang^2(\ell) \leq 1 \}$.

By the above, **Ax1** and **Ax2** are true in $\mathfrak{M}$. It remains to define W. Let

$$m_0 \stackrel{\text{def}}{=} \bar{t} \stackrel{\text{def}}{=} \mathrm{R} \times \{0\}.$$

First we will define $w_{m_0} : {}^2\mathrm{R} \longrightarrow \mathcal{P}(B)$ and $\mathsf{f}_{km_0} : {}^2\mathrm{R} \longrightarrow {}^2\mathrm{R}$ for all $k \in \mathsf{Eucl}(2, \mathfrak{R}), ang^2(k) \neq 1, k \neq m_0$. To define w_{m_0}, let $p \in {}^2\mathrm{R}$. Then

$$w_{m_0}(p) \stackrel{\text{def}}{=} \{\, \ell \in B \,:\, p \in \ell \,\}\,.$$

By this we have that for all $\ell \in B$,

$$tr_{m_0}(\ell) = \ell,$$

in particular, $tr_{m_0}(m_0) = m_0$. Thus **Ax3**, **Ax4**, **Ax5**, **AxE** are satisfied when m is replaced in them with m_0. See Figure 23.

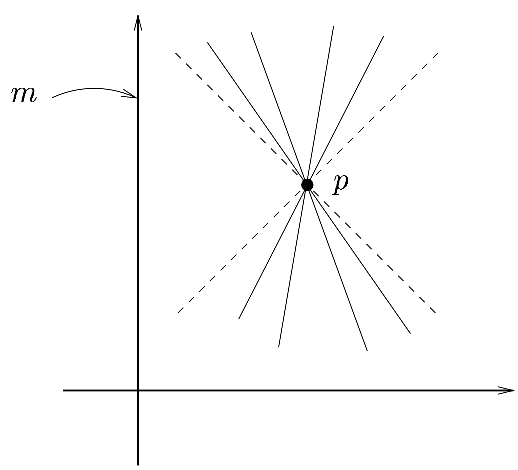


Figure 23: $w_{m_0}(p)$ in $\mathfrak{M}_0^P$.

Let $k \in \mathsf{Eucl}(2, \mathfrak{R}), k \neq m_0, ang^2(k) \neq 1$ be arbitrary. We are going to define f_{km_0}. In the following, we will write f for f_{km_0}.

For any two distinct points $p, q \in {}^nF$, $\overline{pq}$ denotes the Euclidean line containing both p and q.

Sometimes we write (x, y) for the ordered pair $\langle x, y\rangle$. We apologize to the reader for this inconsistency.

Recall that two distinct points, o_k and t_k are given to us by the parameter P of the model $\mathfrak{M} \stackrel{\text{def}}{=} \mathfrak{M}_0^P$. First we define the point s_k as the mirror image of t_k w.r.t. the line ℓ_k such that $o_k \in \ell_k$ and ℓ_k is parallel to the line $\overline{(0,0)(1,1)}$. See Figure 24.

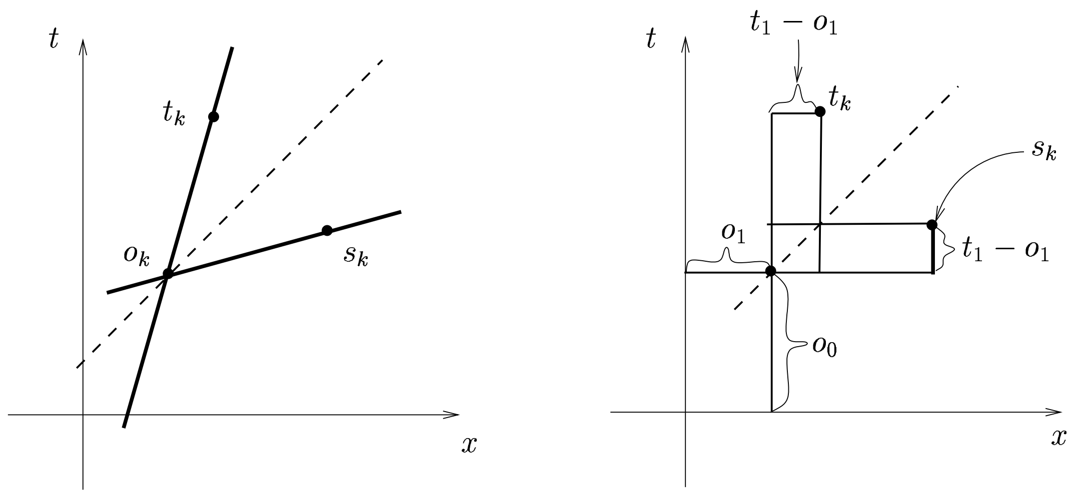


Figure 24: The definition of the point s_k.

In more detail: Let $o_k = (o_0, o_1), t_k = (t_0, t_1)$. We define

$$s_k \stackrel{\text{def}}{=} (o_0 + (t_1 - o_1), o_1 + (t_0 - o_0)).$$

By $ang^2(k) \neq 1$ we have that $s_k \neq t_k$, moreover, $s_k \neq a \cdot t_k$ for all $a \in \mathrm{R}$.

We will define $\mathsf{f} \stackrel{\text{def}}{=} \mathsf{f}_{km_0} : {}^2\mathrm{R} \longrightarrow {}^2\mathrm{R}$ to be the affine transformation[106] that takes $(0,0), (1,0), (0,1)$ to o_k, t_k, s_k respectively. See Figure 25.

In more detail,

$$\mathsf{f}_{km_0}(a, d) \stackrel{\text{def}}{=} a \cdot (t_k - o_k) + d \cdot (s_k - o_k) + o_k.$$

(Here we used that t_k, s_k, o_k are also vectors.) See Figure 26.

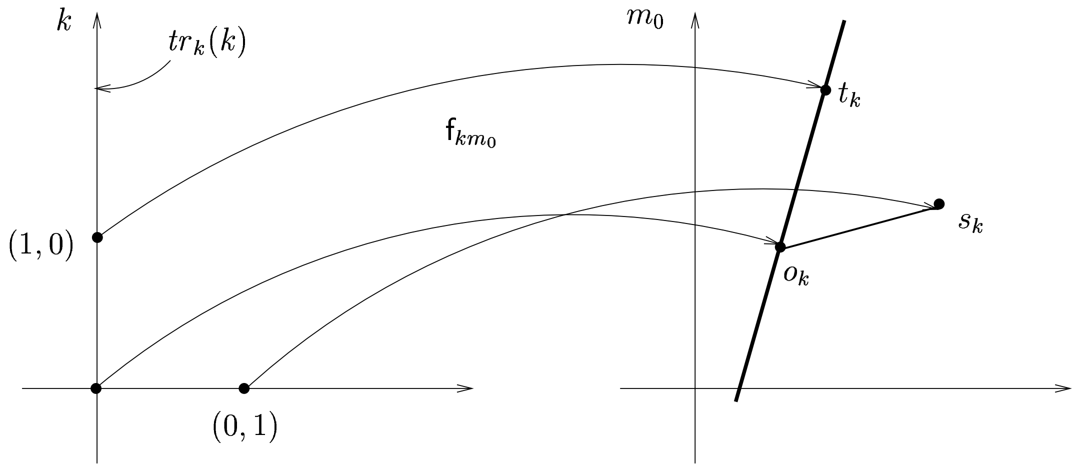


Figure 25: The world-view transformation f_{km_0}.

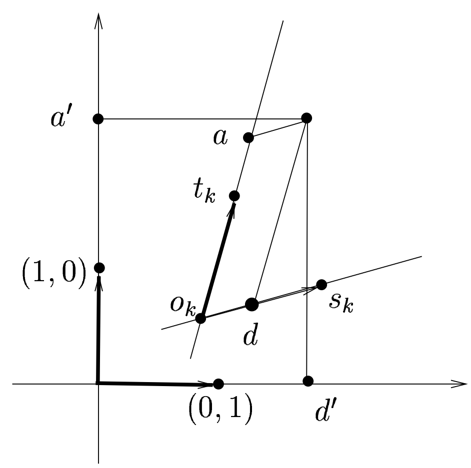


Figure 26: The world-view transformation f_{km_0}.

Intuitively, take a point $p = (a, d)$ in ${}^2\mathrm{R}$, and let $\mathsf{f}_{km_0}(a, d) = (a', d')$. Then a', d' are the coordinates of p in the coordinate system with basis $\{(1, 0), (0, 1)\}$, while a, d are the coordinates of p in the coordinate system with basis $\{(t_k - o_k), (s_k - o_k)\}$, see Figure 26.

By this, f_{km_0} is defined for all $k \in \mathsf{Eucl}(2, \mathfrak{R}), k \neq m_0, ang^2(k) \neq 1$. We now define

$$w_k \stackrel{\text{def}}{=} \mathsf{f}_{km_0} \circ w_{m_0} \text{ for all } k \in Obs \setminus \{m_0\}, \text{ and}$$

$$W \stackrel{\text{def}}{=} \{ \langle m, p, b \rangle \; : \; m \in Obs, b \in w_m(p) \} .$$

By this, the model $\mathfrak{M} \stackrel{\text{def}}{=} \mathfrak{M}_0^P \stackrel{\text{def}}{=} \langle B, \ldots, W \rangle$ has been defined. $\mathfrak{M}_0^P$ is a frame model.

THEOREM 2.4.1 $\mathfrak{M}_0^P \models \mathbf{Basax}(2)$.

Proof. Let $\mathfrak{M} \stackrel{\text{def}}{=} \mathfrak{M}_0^P$. We have already observed that $\mathfrak{M} \models \mathbf{Ax1}, \mathbf{Ax2}$, and that $\mathbf{Ax3} - \mathbf{Ax5}, \mathbf{AxE}$ hold for the fixed observer $m_0 \in Obs$. Let $k \in Obs \setminus \{m_0\}$ be arbitrary. Denote

$$\mathsf{f} \stackrel{\text{def}}{=} \mathsf{f}_{km_0},$$

$$\mathsf{Eucl} \stackrel{\text{def}}{=} \mathsf{Eucl}(2, \mathfrak{R}).$$

We will check that the following $(i) - (v)$ hold:

(i) $\mathsf{f} : {}^2\mathrm{R} \longrightarrow {}^2\mathrm{R}$ is a bijection.

(ii) f takes straight lines to straight lines, i.e. $\mathsf{f}[\ell] \in \mathsf{Eucl}$ for all $\ell \in \mathsf{Eucl}$.

(iii) f takes $\bar{t}$ to k, i.e. $\mathsf{f}[\bar{t}\,] = k$.

(iv) f maps a slow-line to a slow-line, i.e. $(\forall \ell \in Obs)\mathsf{f}[\ell] \in Obs$.

(v) f maps photon-lines onto photon-lines, i.e. $\forall \ell (\ell \in Ph \;\Leftrightarrow\; \mathsf{f}[\ell] \in Ph)$.

[106] For the definition of an affine transformation see §2.9. We will not need the definition here.

Indeed, (i)-(ii) hold because f is a so called affine transformation[107]. (iii) holds then because $o_k, t_k \in k$, and $\mathsf{f}(0,0) = o_k$, $\mathsf{f}(1,0) = t_k$, $m_0 = \overline{(0,0)(0,1)}$, $k = \overline{o_k t_k}$. Since we defined s_k to be the mirror image of t_k w.r.t. ℓ_k, we have that $\mathsf{f}(1,1) = (t_k - o_k) + (s_k - o_k) + o_k$ lies on the line ℓ_k. Thus $\mathsf{f}[\overline{(0,0)(1,1)}] = \ell_k$. Similarly, $ang^2(\mathsf{f}[\overline{(0,0)(1,-1)}]) = 1$. See Figure 27. In dimension 2 (i.e. in $\mathsf{Eucl}(2, \mathbf{F})$), there are exactly two photon-lines going through each point, and we have seen that f takes the two photon-lines going through $(0,0)$ to the two photon-lines going through $\mathsf{f}(0,0)$.

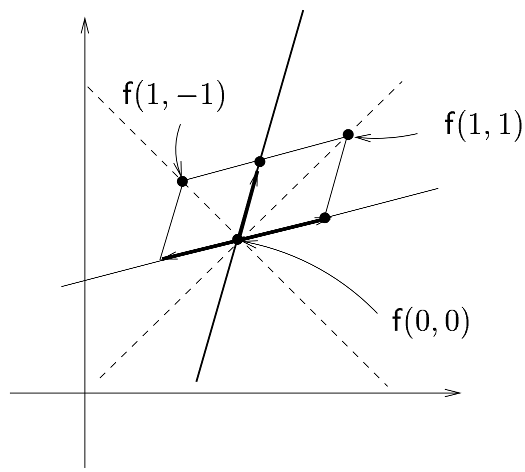


Figure 27: f takes photon-lines to photon-lines.

By (i),(ii) we have that f takes parallel lines to parallel ones. This proves (v). To see that (iv) holds, use a similar argument, and use Figure 28.

We have checked that (i)-(v) hold. Now, in $\mathfrak{M}$ we have that for all $\ell \in B$

(2) $$\ell = tr_k(\mathsf{f}[\ell]).$$

Indeed, let $\ell \in B$. Then $\mathsf{f}[\ell] \in B$, and $\mathsf{f}[\ell] = tr_{m_0}\mathsf{f}[\ell]$. Thus $\ell = \mathsf{f}_{m_0 k}[\mathsf{f}[\ell]] = \mathsf{f}_{m_0 k}[tr_{m_0}\mathsf{f}[\ell]] = tr_k(\mathsf{f}[\ell])$, by Proposition2.3.3(vii). Since f is a bijection, by (ii) we have that both f and f^{-1} preserve Eucl. Using this, together with (ii)-(v), (2) and

[107] This is so because $t_k \neq a \cdot s_k$ for all $a \in \mathrm{R}$, since $ang^2(k) \neq 1$.

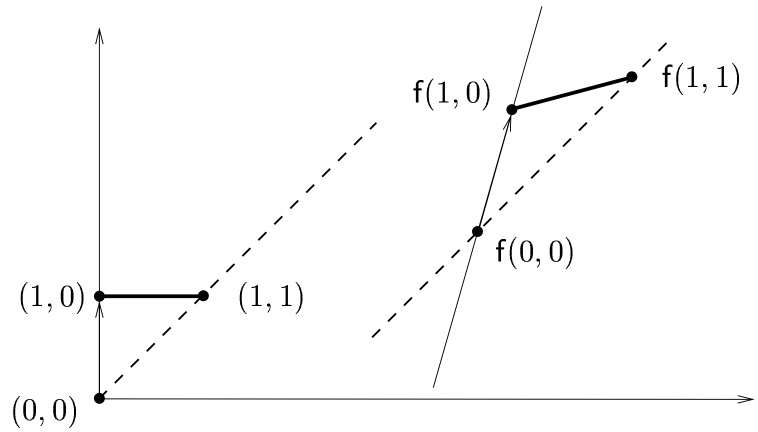


Figure 28: f takes slow-lines to slow-lines.

the fact that $\mathbf{Ax3} - \mathbf{Ax5}, \mathbf{AxE}$ hold for m_0, we get that $\mathbf{Ax3} - \mathbf{Ax5}, \mathbf{AxE}$ hold for k, too. From (i) we get that $Rng(w_k) = Rng(w_{m_0})$. Since k was arbitrary, this proves $\mathfrak{M} \models \mathbf{Basax}(2)$. ∎

Now we define the other model $\mathfrak{M}_1^P$. The definition of $\mathfrak{M}_1^P$ is completely analogous to that of $\mathfrak{M}_0^P$, the only difference is that we allow all lines (with angle $\neq 1$) to be observers. In detail: let

$$Obs_1 \stackrel{\text{def}}{=} \{\, \ell \in \mathsf{Eucl}(2, \mathfrak{R}) \ :\ ang^2(\ell) \neq 1 \,\},$$

$$B_1 \stackrel{\text{def}}{=} Ib_1 \stackrel{\text{def}}{=} Obs_1 \cup Ph = \mathsf{Eucl}(2, \mathfrak{R}).$$

Then $m_0 \in Obs \subseteq Obs_1$. We define

$$w'_{m_0}(p) \stackrel{\text{def}}{=} \{\, \ell \in B_1 \ :\ p \in \ell \,\},$$

$$w'_k \stackrel{\text{def}}{=} \mathsf{f}_{km_0} \circ w'_{m_0},$$

$$W' \stackrel{\text{def}}{=} \{\, \langle m, p, b \rangle \ :\ b \in w'_m(p) \,\},$$

$$\mathfrak{M}_1^P \stackrel{\text{def}}{=} \langle (B_1;\ Obs_1, Ph_1, Ib_1), \mathfrak{R}, G;\ \in, W' \rangle.$$

THEOREM 2.4.2 $\mathfrak{M}_1^P \models \mathbf{Basax}(2)$.

The proof is analogous to that of Theorem 2.4.1, we omit it.

In Figure 29, $\mathfrak{M}_1, \mathfrak{M}_2, \mathfrak{M}_3$ represent possible models of **Basax**(2). There $\bar{t}, \bar{t}', \bar{t}''$ are the time axes of observers $k, k', k'' \in Obs$.

Consider e.g. the picture representing $\mathfrak{M}_1$ (first picture of Figure 29). What the picture *really* represents is the *world-view of a particular observer k* and also how k sees k', k'' etc. In the picture $\bar{t}', \bar{t}''$ represent the life-lines of observers k', k''. Further $1_{t'} = \mathsf{f}_{k'k}(1_t)$, $1_{t''} = \mathsf{f}_{k''k}(1_t)$ and $1_{x'} = \mathsf{f}_{k'k}(1_x)$ etc. Intuitively, $1_{t'}$ is the time-unit vector of k' as seen by k, while $1_{x'}$ is the x-unit vector of k' as seen by k. We do *not* claim that the world-view of observer k' would be similar. Actually it is not. The *only* thing we claim is that there is an observer k of $\mathfrak{M}_1$ whose world-view is as represented in the picture. The same convention applies to the pictures representing $\mathfrak{M}_2$ and $\mathfrak{M}_3$.

Figure 29 represents possible choices for the parameter P of the model $\mathfrak{M}^P$ introduced on p.80. Recall that $\langle o_\ell, t_\ell \rangle = P(\ell)$. In Figure 29 $\ell \in \{\bar{t}, \bar{t}', \bar{t}'', \text{ etc.}\}$. In the figure we chose $o_\ell := \bar{0}$, further $t_\ell := 1_t$ etc.

Let us turn now to the *ideas* we wanted to represent in these pictures.

In the third picture, the curves connecting $1_t, 1_{t'}, 1_{t''}$ etc. are hyperbolas. In §2.8 we will introduce a symmetry axiom called **Ax(symm)**. We note that $\mathfrak{M}_3 \models$ **Ax(symm)**, while $\mathfrak{M}_i \not\models$ **Ax(symm)** for $i < 3$. Roughly speaking **Ax(symm)** says that "I see you the same way as you see me". Thus in $\mathfrak{M}_3$ all observers see the other observers' unit vectors as m sees it (while as we mentioned, in $\mathfrak{M}_1$ and in $\mathfrak{M}_2$ this is not so). We also note that $\mathfrak{M}_3$ corresponds to the *usual* (or classically standard) so-called Minkowskian models of relativity, while $\mathfrak{M}_1, \mathfrak{M}_2$ are "non-Minkowskian" (for the definition of a Minkowskian model see Definition 3.8.42 on p.331).

A common feature of $\mathfrak{M}_1$–$\mathfrak{M}_3$ in Figure 29 is that, for m fixed,

$$v_m(k) \longmapsto |\mathsf{f}_{mk}(1_t) - \mathsf{f}_{mk}(\bar{0})|$$

is (i) a function (of $v_m(k) \in F$) and this function is (ii) continuous. These properties will re-emerge as potential axioms in a later version of this work. Although these properties do *not* follow from **Basax**, we will not put too much emphasis on studying models which do not satisfy (i) or (ii). *Analogous* properties will show up in §4.4 as potential axioms.

We will return to the pictures in Figure 29 in §6.7.

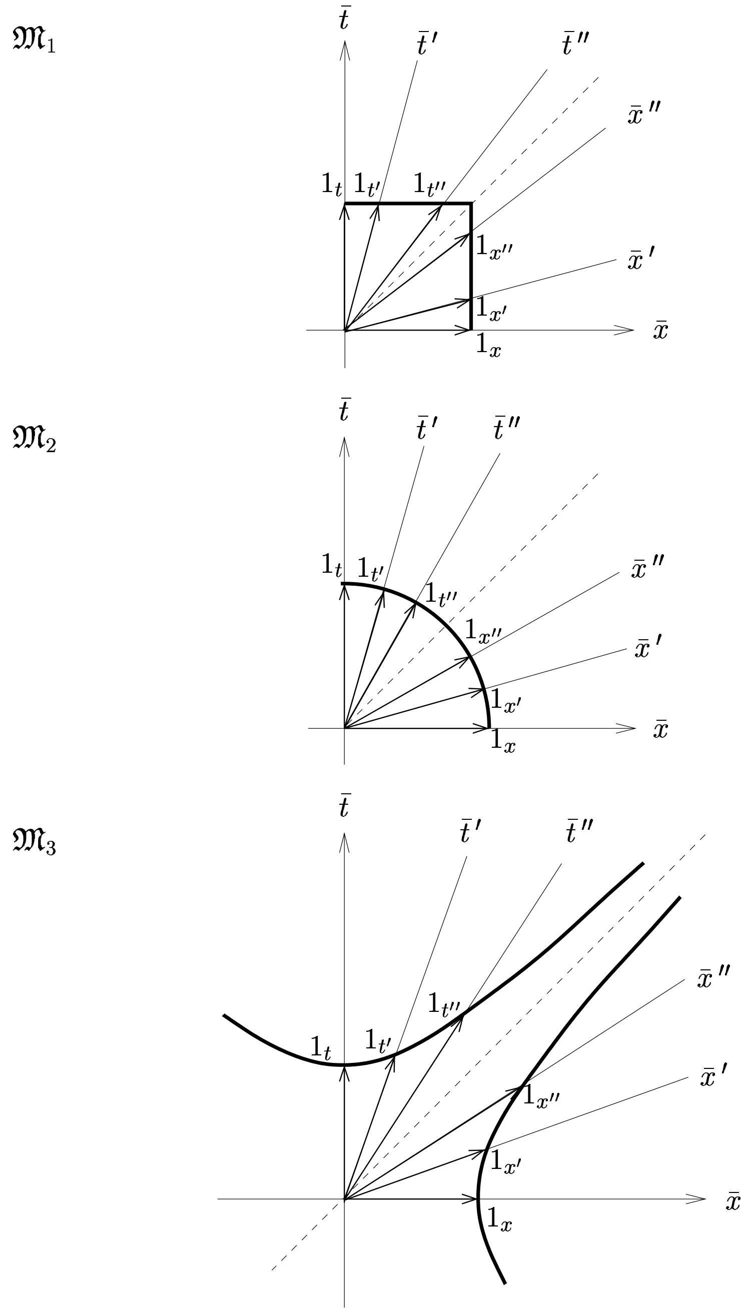


Figure 29: Possible models of **Basax**. Possible choices for the parameter P (on p.80): $o_\ell = \bar{0}$ and t_ℓ are represented in the picture, for slow-lines ℓ going through $\bar{0}$.

Minkowski-circles, Minkowski-spheres.

Assume $n = 2$. The drawings in Figure 29 are called Minkowski-circles. They are often useful in representing models by simple drawings.

Definition 2.4.3 Minkowski-sphere Let $n \geq 2$, $\mathfrak{M}$ be a frame model, and $m \in Obs$. Then the *Minkowski-sphere* MS around m is defined as

$$\mathsf{MS} \stackrel{\text{def}}{=} \mathsf{MS}(\mathfrak{M}, m) \stackrel{\text{def}}{=} \{p : (\exists k \in Obs)(\exists i < n)(\mathsf{f}_{mk}(\bar{0}) = \bar{0} \ \text{ and } \ p \in \{\mathsf{f}_{mk}[\{1_i, -1_i\}]\}) \}.$$

$\triangleleft$

For very nice models (e.g. the ones studied in §2.8) MS forms a kind of surface such that one can imagine that this surface is a boundary[108] of a connected region like the inside of a ball (or a cube, or something like these). This is indeed the case with the three models in Figure 29 (p.88). In two dimensions, instead of "spheres" we speak of *Minkowski-circles*. What we said above about the Minkowski-spheres in dimension n, sounds like the following for $n = 2$. In nice 2-dimensional models, MS as defined above looks like a nice curve (like a circle, or a square etc) such that one can imagine that MS is the boundary of a connected subset of the plane like the circle is the boundary of a "disc". This is the case in all three drawings in Figure 29. Classically, in standard relativity theory, only the figure associated to $\mathfrak{M}_3$ was called a Minkowski-circle. (The reason for this is that only $\mathfrak{M}_3$ satisfies the symmetry axiom to be introduced in §2.8.) However, here we generalize this concept to arbitrary frame models. As we said, in nice models, $\mathsf{MS}(\mathfrak{M}, m)$ looks like a curve surrounding (or forming the boundary of) some connected area. However, in many less "well behaved" models $\mathsf{MS}(\mathfrak{M}, m)$ is just a set of points and does not even form a curve. Later we will introduce an axiom called $\mathbf{Ax(\|)}$. Typically, if $\mathbf{Ax(\|)}$ fails, then MS tends to become more like a random set of points than a curve. With this, we stop the discussion of Minkowski-spheres and Minkowski-circles, but from time to time they will serve us as pleasant devices for visualizing certain nice, well behaved models.

In passing we note that in the case of $n = 2$, it is more often the case that $\mathsf{MS}(\mathfrak{M}, m)$ is like a curve surrounding a well defined area, while if $n > 2$ then this is more rare.[109] Basically, if $n > 2$, and $(\forall m \in Obs)[\mathsf{MS}(\mathfrak{M}, m)$ is a surface surrounding a connected and well defined area], then the extra axiom **Ax(symm)** to be defined later is true in our model $\mathfrak{M}$, and then the Minkowski-sphere becomes practically the same what is called such in the classical literature (cf. e.g. "Minkowski-metric" in Friedman [90]). On the other hand, for $n = 2$ this is far from being true as is illustrated e.g. by Figure 29.

[108]if we disregard the points on the life-lines of photons crossing the origin.

[109]We mean this with "surface" in place of "curve", of course.

2.5 The three "paradigmatic" theorems of relativity

What the average layperson usually knows about relativity is that

(I) moving clocks slow down,

(II) moving spaceships shrink (cf. Figure 30), and

(III) moving clocks get out of synchronism, i.e. the clock in the nose of the spaceship is late (shows less time) when compared with the clock in the rear, see Figure 31.

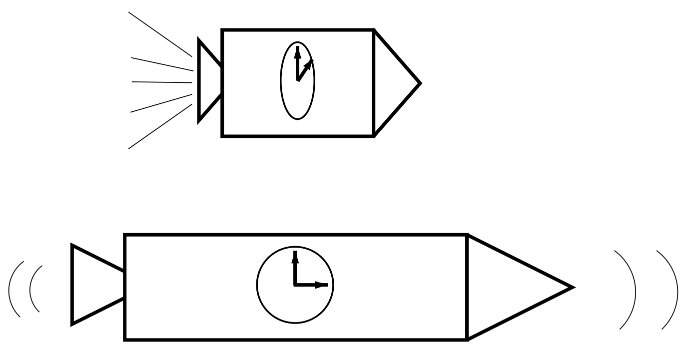


Figure 30: Moving clocks slow down and moving spaceships shrink.

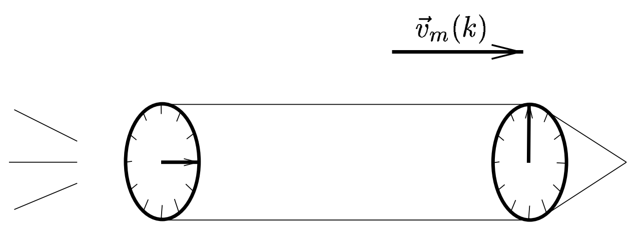


Figure 31: Moving clocks get out of synchronism.

In all of (I)–(III) above the spaceship is represented by an observer k, "we" who look at the spaceship are represented by observer m, and all of (I)–(III) are understood in the world-view of m. Below we formalize (I)–(III) as our *"paradigmatic" theorems*.[110] We will prove them from **Basax**. In §4 when investigating weaker (or subtler) versions of **Basax** (e.g. the Reichenbachian version with non-standard simultaneities) we will systematically re-visit our paradigmatic theorems to see if they are still true. It will turn out that these paradigmatic theorems can be proved from surprisingly weak axioms. Cf. §4, and especially section §4.8 which is devoted to paradigmatic effects. In § 2.8 we will see that our paradigmatic theorems (I)-(III) hold in a stronger and simpler form in the stronger axiom system $\mathbf{Basax} + \mathbf{Ax(symm)}$.

Our next axiom, $\mathbf{Ax}(\sqrt{\ })$, is of a technical nature. Namely, sometimes we will need to assume that square roots of positive (greater than 0) elements exist in the ordered field reduct $\mathfrak{F}$ of the frame model $\mathfrak{M}$ we are speaking about.

$\mathbf{Ax}(\sqrt{\ })$ $\quad (\forall 0 < x \in F)(\exists y \in F)\ y^2 = x.$

If $\mathfrak{F} \models \mathbf{Ax}(\sqrt{\ })$ then we say that $\mathfrak{F}$ is *Euclidean*. Clearly, $\mathfrak{R} \models \mathbf{Ax}(\sqrt{\ })$. For any $0 < x \in F$, $\sqrt{x}$ denotes that positive y for which $y^2 = x$. For brevity, by an *Euclidean field* we mean an Euclidean ordered field.

CONVENTION 2.5.1 Let **Th** be a set of formulas of our frame language. Let $\mathbf{Ax}_1$, $\mathbf{Ax}_2$ be further formulas. Then

$$\mathbf{Th} + \mathbf{Ax}_1 + \mathbf{Ax}_2 \quad \text{denotes} \quad \mathbf{Th} \cup \{\mathbf{Ax}_1, \mathbf{Ax}_2\}.$$

Similar convention applies to other combinations like $\mathbf{Th} + \mathbf{Ax}_1$. (This notational convention is taken from axiomatic set theory.)

$\lhd$

The intuitive meaning of Thm.2.5.2 below is the following. Item (i) of the theorem states that observer m thinks that k's clocks are late at time-instance 1. As a generalization of this, (ii) says the same for many time instances $\lambda \in F$ namely for those λ's which are not "infinitely big" or "infinitely small".

[110] In passing we note that the official names for effects (I) and (II) are "time dilation" and "length contraction" cf. d'Inverno [75, §§3.3, 3.4].

THEOREM 2.5.2 (Clocks slow down.)
Assume **Basax** + **Ax**($\sqrt{\ }$). *Then (i)–(iii) below hold.*

(i) *There are observers m and k such that m "thinks" that k's clocks run slow; formally:*

$$(\exists m, k \in Obs)\ |\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > 1\,,$$

see Figure 32. Moreover;

(ii) $(\exists m, k \in Obs)$
$$(\forall \lambda \in F)\ \Big((\exists 0 < j \in \omega)\ 1/j < |\lambda| < j \quad \Rightarrow \quad |\mathsf{f}_{km}(\lambda \cdot 1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > |\lambda|\Big).$$[111]

(iii) *Assume $m, k \in Obs$ and $0 \neq v_m(k) < 1$. Then either m thinks that k's clocks run slow <u>or</u> k thinks that m's clocks run slow (cf. Figure 33); formally:*

$$(\exists m', k' \in \{m, k\})\ |\mathsf{f}_{k'm'}(1_t)_t - \mathsf{f}_{k'm'}(\bar{0})_t| > 1.$$

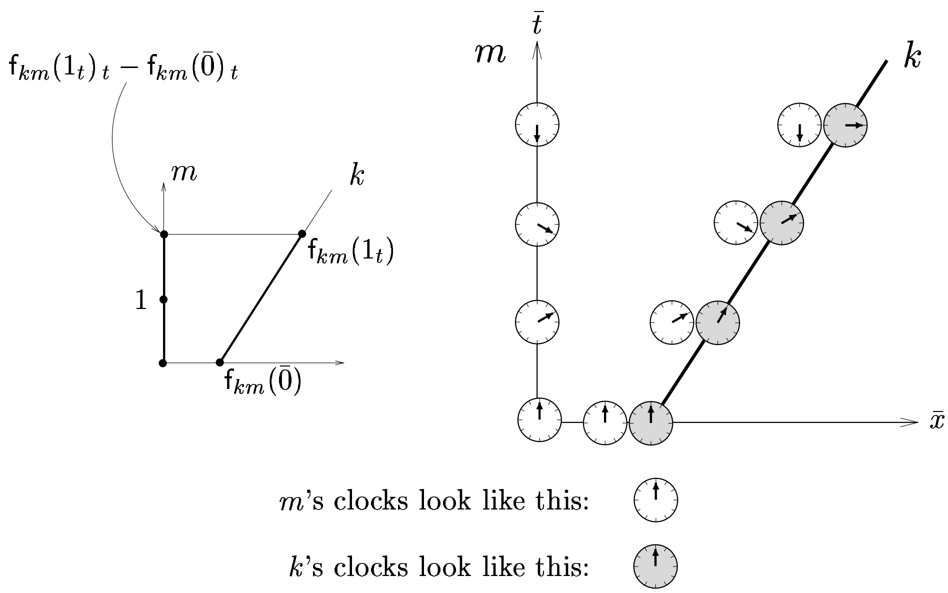


Figure 32: m thinks that k's clocks run slow.

[111]We note that for every ordered field the set ω of the natural numbers can be considered as a subset of the ordered field, or in more careful wording ω is embeddable into the ordered field in a natural way. Further we note that if $\mathfrak{F}$ is Archimedian (cf. footnote 88 on p.66) then (ii) above is true in the following simpler form: (ii)' $(\exists m, k \in Obs)(\forall p \in \bar{t})\ |\mathsf{f}_{km}(p)_t - \mathsf{f}_{km}(\bar{0})_t| > |p_t|$. An analogous remark applies to Thm.2.5.3(ii).

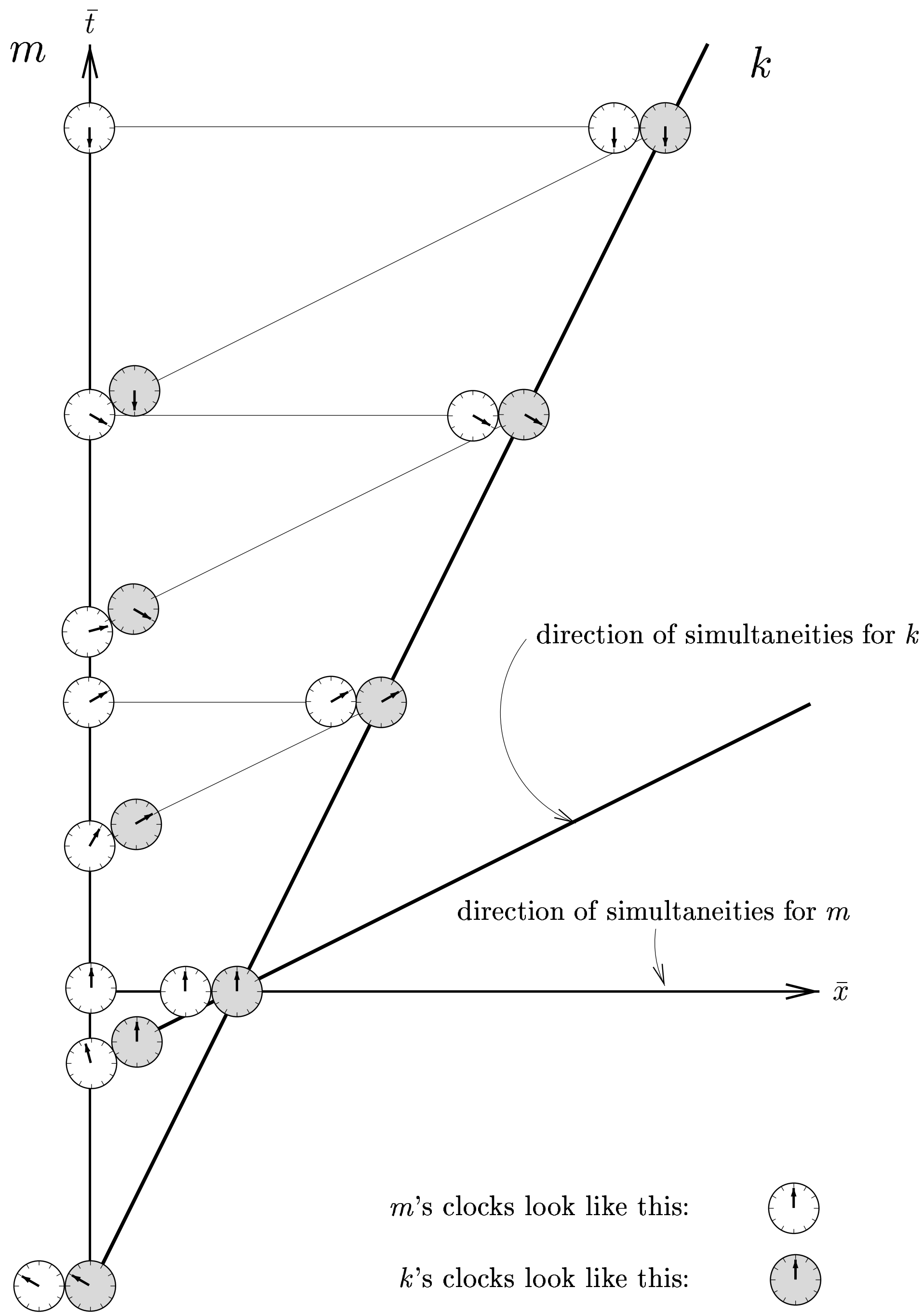


Figure 33: Assume that for m, k's clocks do not run slow. Then k will think that m's clocks run slow.

On the proof: The main idea of the proof of (iii) is illustrated in Figure 33. (i) is a corollary of (iii). (ii) can be proved from Thm.2.5.3(i) below as follows:
By Thm.2.5.3(i), there are $m, k \in Obs$ such that

$$(3) \qquad |\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > 2\,.$$

We have that every automorphism of $\mathbf{F}$ is order preserving, i.e. every automorphism of $\mathbf{F}$ is an automorphism of $\mathfrak{F}$ since $\mathfrak{F}$ is Euclidean. So, by Prop.3.1.4 (p.162), we have that

$$(4) \qquad (\forall \lambda \in F)\ \Big(|\mathsf{f}_{km}(\lambda \cdot 1_t)_t - \mathsf{f}_{km}(\bar{0})_t| \ = \ \varphi(|\lambda|) \cdot |\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t|\Big),$$

for some automorphism φ of $\mathfrak{F}$.

For every automorphism φ of $\mathfrak{F}$ we have

$$(5) \qquad (\forall \lambda \in F)\ \left((\exists 0 < j \in \omega)\ \frac{1}{j} < |\lambda| < j \ \Rightarrow\ \varphi(|\lambda|) > \frac{|\lambda|}{2}\right),$$

because of the following. Let $\lambda \in F$ such that $(\exists j \in \omega)\ 1/j < |\lambda| < j$. Between $|\lambda|/2$ and $|\lambda|$ there is a rational number, say x. Let such an x be fixed. Every automorphism (of $\mathfrak{F}$) is the identity function on the rational numbers. Therefore by $|\lambda|/2 < x < |\lambda|$, we have $|\lambda|/2 < x = \varphi(x) < \varphi(|\lambda|)$. So (5) holds. Let $\lambda \in F$ such that there is $0 < j \in \omega$ with $1/j < \lambda < j$. Then

$$\begin{array}{rcll} |\mathsf{f}_{km}(\lambda \cdot 1_t)_t - \mathsf{f}_{km}(\bar{0})| & = & \varphi(|\lambda|) \cdot |\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t| & \text{by (4)} \\ & > & 2\,\varphi(|\lambda|) & \text{by (3)} \\ & > & |\lambda| & \text{by (5).} \end{array}$$

This completes the proof of (ii). A more detailed proof will be filled in at a later stage of development. ∎

THEOREM 2.5.3 (Clocks can run very slow.)
Assume $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ })$. *Let* $\varrho \in \omega$ *be arbitrary.*

(i) *Then there are observers m, k such that m thinks that k's clocks run more than ϱ-times slower than m's; formally:*

$$(\exists m, k \in Obs)\ |\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > \varrho\,.$$

Moreover:

(ii) $(\exists m, k \in Obs)$

$$(\forall \lambda \in F)\ \Big((\exists 0 < j \in \omega)\ 1/j < |\lambda| < j \ \Rightarrow\ |\mathsf{f}_{km}(\lambda \cdot 1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > \varrho \cdot |\lambda|\Big).$$

On the proof: We include Figure 34 as a hint for the idea of the proof of (i). (ii) follows from (i) similarly as item (ii) of Thm.2.5.2 did: Now, for a $\varrho \in \omega$ we choose $m, k \in Obs$ such that $|\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > 2\varrho$. Then analogously to the proof of Thm.2.5.2(ii) one can prove that for this choice of m, k the "main body" of (ii) holds. A detailed proof will be filled in later. ∎

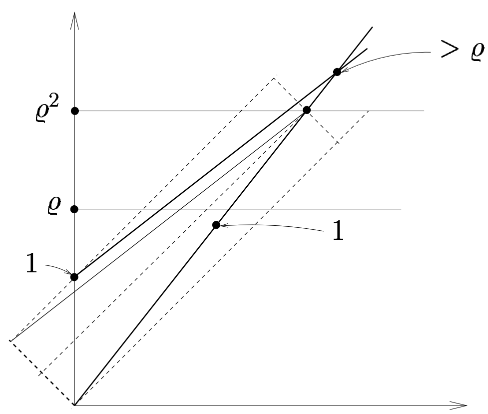


Figure 34: Hint for the idea of proof of Thm.2.5.3(i).

<u>*Let us turn to clocks getting out of synchronism*</u> ("effect" (III) on our "paradigmatic" list). First we need some definitions.

Definition 2.5.4 Let $\mathfrak{M}$ be a frame model. Events $e, e_1 \in \mathcal{P}(B)$ are said to be <u>*simultaneous*</u> for observer $m \in Obs$ iff

$$e, e_1 \in Rng(w_m) \quad \wedge \quad (\forall p \in w_m^{-1}(e))(\forall q \in w_m^{-1}(e_1))\ p_t = q_t\,.$$ [112]

$\triangleleft$

[112] To improve readability we write $w_m^{-1}(e)$ instead of $w_m^{-1}[\{e\}]$, where $m \in Obs$ and $e \in \mathcal{P}(B)$.

THEOREM 2.5.5 (Clocks get out of synchronism.)
Assume **Basax**. *Let $m, k \in Obs$ be such that $v_m(k) \neq 0$. Then (i) and (ii) below hold.*

(i) *There are events $e, e_1 \in \mathcal{P}(B)$ which are simultaneous for m, but are not simultaneous for k.*

(ii) *Assume that k moves in direction $\bar{x}$ as seen by m, formally: $tr_m(k) \subseteq \mathsf{Plane}(\bar{t}, \bar{x})$. Then*

$$(\forall p, q \in {}^nF)\ \Big((p_t = q_t \quad \wedge \quad p_x \neq q_x) \quad \Rightarrow \quad \mathsf{f}_{mk}(p)_t \neq \mathsf{f}_{mk}(q)_t\Big),$$

cf. Figure 35. I.e. if m thinks that $e = w_m(p)$ and $e_1 = w_m(q)$ are simultaneous but their x-coordinates are different, then k will think that e and e_1 are not simultaneous.

Intuitively, let us imagine that k is traveling on a spaceship and is being observed by m. Then m will think that clocks in the nose and the rear of k's spaceship are not synchronous, cf. Figures 31, 37. (They do not show the same time.)

The **proof** will be filled in later. ∎

We note that Thm.2.5.5 can be refined in the style of Thm.2.5.7 below.

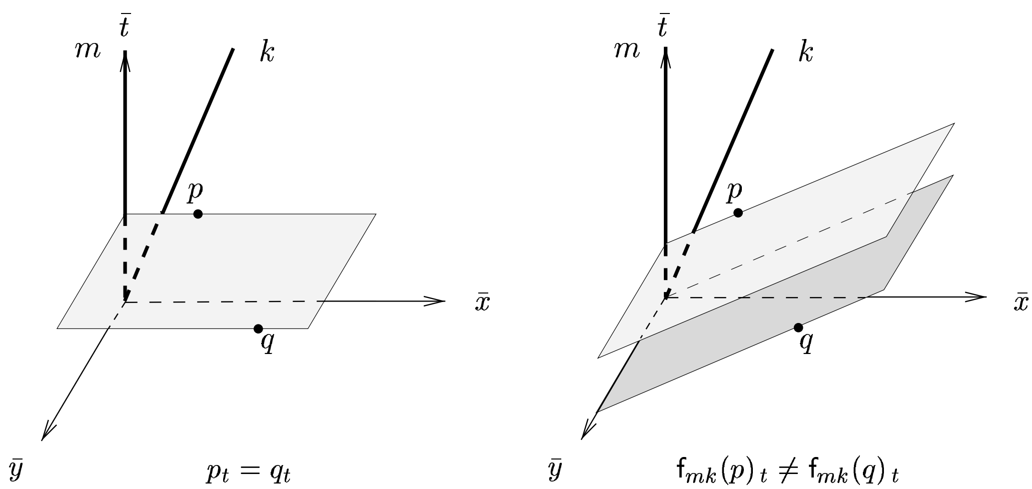


Figure 35: Events $w_m(p)$ and $w_m(q)$ are simultaneous for observer m, but they are not simultaneous for observer k.

THEOREM 2.5.6 (Clocks do not get out of synchronism in direction orthogonal to movement.)
From the point of view of synchronism, "nothing" happens in the spatial direction orthogonal to the direction of movement (cf. Figure 36);[113] *formally: Assume* **Basax**. *Let* $m, k \in Obs$. *Then (i) and (ii) below hold.*

(i) *Assume* m *sees that* k *does not move in direction* $\bar{y}$, *i.e.*
$$(\forall p, q \in tr_m(k))\ p_y = q_y\,.$$
Then,
$$(\forall p, q \in {}^nF)\ \Big((\forall i \in n)(i \neq 2 \Rightarrow p_i = q_i) \quad \Rightarrow \quad \mathsf{f}_{mk}(p)_t = \mathsf{f}_{mk}(q)_t\Big)\,.$$
In particular
$$p, q \in \bar{y} \qquad \Rightarrow \qquad \mathsf{f}_{mk}(p)_t = \mathsf{f}_{mk}(q)_t\,.$$
That is, simultaneous events observed by m *as separated only in a direction* $\bar{y}$ *orthogonal to the direction of movement* <u>*remain*</u> *simultaneous for the moving observer* k.

The following is an almost equivalent re-formulation of (i).[114]

(ii) *Assume that* k *moves in direction* $\bar{x}$ *as seen by* m, *formally:* $tr_m(k) \subseteq \mathsf{Plane}(\bar{t}, \bar{x})$. *Then*
$$(\forall p, q \in {}^nF)\ \Big((p_t = q_t \ \wedge\ p_x = q_x) \quad \Rightarrow \quad \mathsf{f}_{mk}(p)_t = \mathsf{f}_{mk}(q)_t\Big)\,.$$

The **proof** will be filled in later. ∎

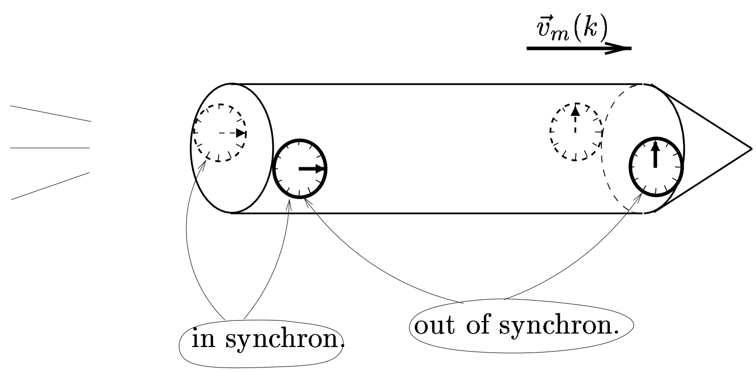


Figure 36: Clocks do not get out of synchronism orthogonal to movement. Imagine the little clocks glued to the hull of the spaceship.

[113]More precisely, if two clocks are separated ***only*** in a spatial direction which is orthogonal to the direction of movement then they do ***not*** get out of synchronism.

Let us return to discussing the "out of synchronism" effects in Thm.2.5.5. Throughout the rest of this section (in §2.5) $\underline{\textbf{Basax} + \textbf{Ax}(\sqrt{\ })\ \textit{is assumed}}$ (unless otherwise specified), therefore we do not indicate this.

Let $\mathfrak{M}$ be a frame model, and let $m, k \in Obs$ such that $tr_m(k) \in \mathsf{Eucl}$. Let us recall that the velocity of k as seen by m is denoted by $\vec{v}_m(k)$, cf. p.48, and it is a "space vector", i.e. an element of ${}^{n-1}F$.

In connection with the next two theorems we note the following. Since we assumed **Basax**, for every $m, k \in Obs$, $tr_m(k)$ can be considered as a function $tr_m(k) : F \longrightarrow {}^{n-1}F$, if $v_m(k) \neq \infty$; therefore $tr_m(k)(0)$ is well defined.[115]

THEOREM 2.5.7 (The clock in the nose of the spaceship is late.) *Let $m, k, k_1 \in Obs$. Assume, $tr_k(k_1)$ is parallel with $\bar{t}$,[116] $0 \neq v_m(k) < 1$, and that time passes forwards for k as seen by m, formally: $(\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t) > 0$. Intuitively, k represents the rear of the "spaceship", while k_1 represents the nose of the "spaceship". Assume further that this "spaceship" is moving forwards as seen by m;* [117] *formally:*

$$\Big(tr_m(k_1)(0) - tr_m(k)(0)\Big) = \lambda \cdot \vec{v}_m(k) \qquad \textit{for some positive } \lambda \in F.$$

(i) *Then m thinks that the clock in the nose of the spaceship is late w.r.t. the clock in the rear of the spaceship (see Figure 37); formally:*

$$(\forall p \in tr_m(k))(\forall q \in tr_m(k_1)) \ \Big(p_t = q_t \ \Rightarrow \ \mathsf{f}_{mk}(p)_t > \mathsf{f}_{mk}(q)_t\Big).$$

(ii) *Let $m, k, k_1 \in Obs$ satisfy all the conditions of the present theorem. Assume further that the length of the "spaceship" as seen by k is 1. For simplicity we formalize this condition as $1_x \in tr_k(k_1)$. Then*

$$(\forall p \in tr_m(k))(\forall q \in tr_m(k_1)) \ \Big(p_t = q_t \ \Rightarrow \ (\mathsf{f}_{mk}(p)_t - \mathsf{f}_{mk}(q)_t) < 1\Big).$$

Intuitively, this says that assuming the length of the spaceship as seen by k is 1 then the difference between the two clock readings (in the rear and in the nose) as seen by m is always smaller than 1.

[114] If we assume $\textbf{Ax}(\sqrt{\ })$ together with the auxiliary axiom $\textbf{Ax}(\mathit{Triv})$ which will be introduced in §2.8, then (i) and (ii) become equivalent.

[115] Cf. Fact 2.2.4.

[116] This means that k_1 is in rest w.r.t. k, i.e. we use the relationship parallelism between lines in the sense of Euclidean geometry.

[117] Intuitively, this means that m sees the spaceship moving in the direction of its nose.

(iii) *For any $\lambda \in F$ with $0 < \lambda < 1$, there are m, k, k_1 satisfying all the conditions of the present theorem, including the condition in (ii) saying that the length of the "spaceship" as seen by k is 1, such that*

$$(\forall p \in tr_m(k))(\forall q \in tr_m(k_1)) \left(p_t = q_t \;\Rightarrow\; (\mathsf{f}_{mk}(p)_t - \mathsf{f}_{mk}(q)_t) > \lambda\right).$$

Intuitively, the asynchronism between the two clock readings can get arbitrarily close to 1*.*

We will fill in the **poof** later. ∎

Item (iii) of the above theorem describes how much "asynchronism" we can obtain as a relativistic effect.

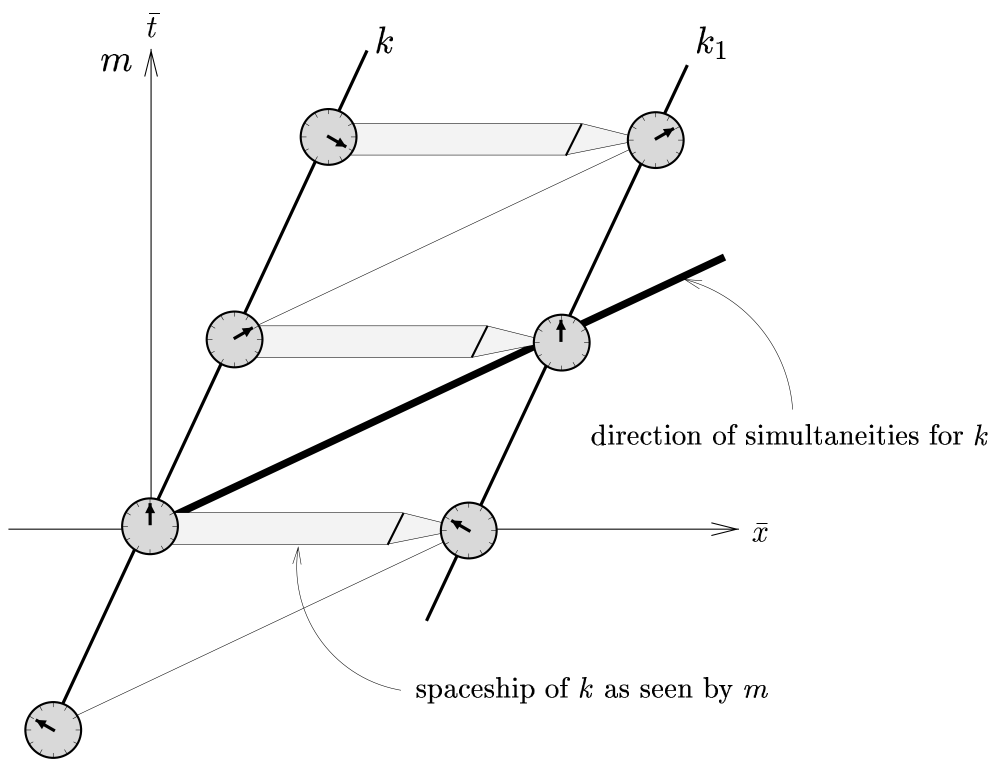


Figure 37: Clock (of k) in the nose of the spaceship is late w.r.t. the clock in the rear, when viewed by m. (The length of this spaceship is more than 1 as seen by k.)

Next, we turn to discussing how meter-rods shrink, i.e. to paradigmatic effect (II). (Strangely enough, one has to represent meter-rods and their shrinking slightly differently than it was the case with clocks.)

Notation 2.5.8 Assume $\mathfrak{F}$ is Euclidean, i.e. that $\mathfrak{F} \models \mathbf{Ax}(\sqrt{\ })$. Let $p \in {}^nF$. Then $|p|$ denotes the Euclidean length of vector p, i.e.

$$|p| \stackrel{\text{def}}{=} \sqrt{p_0^2 + p_1^2 + \ldots + p_{n-1}^2}\,.$$

$\triangleleft$

THEOREM 2.5.9 (Spaceships shrink.)
There are observers $m, k, k_1 \in Obs$ with $\bar{0} \in tr_m(k)$, and with $tr_k(k_1)$ parallel to $\bar{t}$,[118] such that for $p := tr_m(k_1)(0)$ and $q := tr_k(k_1)(0)$, we have $|p| < |q|$.[119]

The last statement $|p| < |q|$ can be interpreted as saying that m thinks that the purely spatial distance between observers k and k_1 is shorter than it is observed by k, see Figure 38.

Intuitively, if k represents the rear of the "spaceship" and k_1 represents the nose of the "spaceship", then this spaceship is shorter for m than for k.

The **proof** will be filled in later. ∎

Remark 2.5.10 An improved version of Thm.2.5.9 could be formulated analogously to Thm.2.5.3 saying that meter-rods can get arbitrarily short (if they are parallel with the direction of movement).

$\triangleleft$

Remark 2.5.11 Analogously with Thm.2.5.6, we could formulate a theorem saying that meter-rods orthogonal to the direction of movement do not get shorter (at least not as a consequence of relativistic effects). But for this we would need extra conditions formulated in §2.8 "A symmetry axiom". Cf. Item 2.8.12 (p.133).

$\triangleleft$

[118] This means that k_1 is in rest w.r.t. k, i.e. we use the relationship parallelism between lines in the sense of Euclidean geometry.
[119] Here again $|p|$ is the Euclidean length of $p \in {}^{n-1}F$.

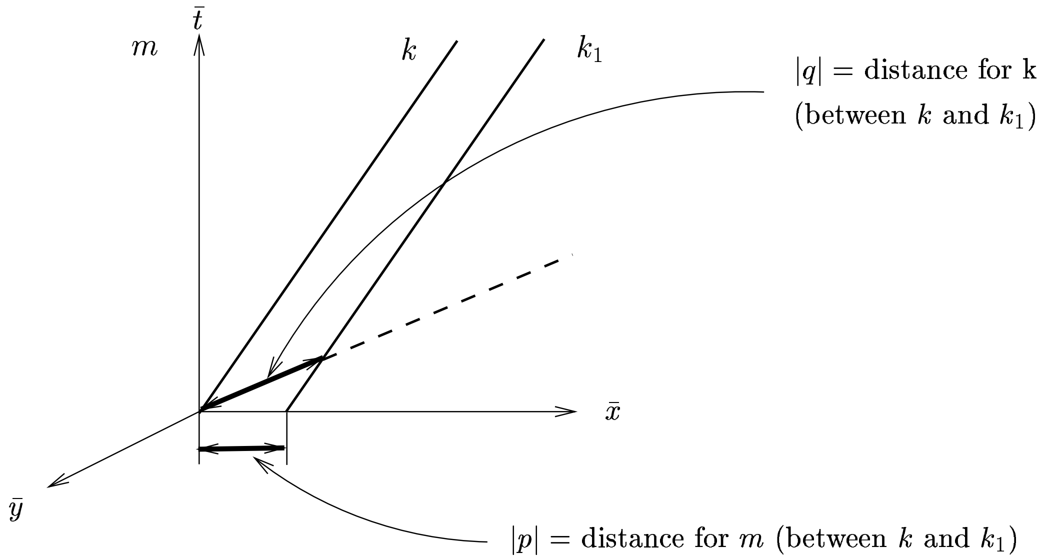


Figure 38: Illustration for Thm.2.5.9.

Remark 2.5.12 Throughout the present remark we assume that the ordered field reduct $\mathfrak{F}$ of our model $\mathfrak{M}$ has no nontrivial automorphisms.[120]

(i) Thm.2.5.9 above could be interpreted and modified intuitively as follows: There are observers m and k such that m sees k moving in direction $\bar{x}$, and those meter-rods of k which are pointing in direction $\bar{x}$ *as seen by m,* are shorter when observed by m than as observed by k. In short: m thinks that k's meter-rods pointing in direction $\bar{x}$, as seen by m, shrink. We will use this intuitive language in the rest of the remark without formalizing it. The reader is invited to formalize it.

(ii) Assume m sees k moving in direction $\bar{x}$ slower than light and with nonzero speed.

(a) Let us concentrate on meter-rods pointing in direction $\bar{x}$ *as seen by m*.[121] Let us call them x-meter-rods. Then either m will think that k's x-meter-

[120] This assumption can be eliminated on the expense of restricting discussion to meter-rods of rational length as seen by that observer whose meter-rods they are.

[121] I.e. even if the meter-rod is k's one we check whether m sees it pointing in direction $\bar{x}$.

rods shrink or k will think that m's x-meter-rods shrink or both.[122]

(b) Let us concentrate on meter-rods pointing in direction $\bar{y}$ as seen by m.[123] Let us call them y-meter-rods. Then if m thinks that k's y-meter-rods shrink, then k will think that m's y-meter-rods grow.

(iii) Let $m, k \in Obs$. Assume $v_m(k) < 1$. Then one of them thinks that all meter-rods of the other shrink or remain unchanged. Those meter-rods shrink the most which point in the direction of movement. Further, those can remain unchanged which are orthogonal to the direction of movement.

(iv) Let us return to clocks getting out of synchronism in connection with item (iii), cf. Theorems 2.5.5–2.5.7. Let $m, k \in Obs$. Consider pairs of clocks which are synchronous for k and the distance between two clocks in a pair is 1 for k. Then that pair will get out of synchronism most the connecting line of which is parallel to the direction of motion of k (as seen by m). $\triangleleft$

The following theorem says that on a moving spaceship (i) either clocks slow down or meter rods (pointing in the direction of movement) shrink (or both, of course) and (ii) clocks in the rear and the nose of ship get out of synchronism.

THEOREM 2.5.13 (Clocks slow down or meter rods shrink.)
Assume $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ })$. *Let* $m, k \in Obs$, $0 < v_m(k) < 1$. *Then (i), (ii) below hold.*

(i) *Either the clocks of k run slow or meter rods of k parallel with $\vec{v}_m(k)$ shrink (as seen by m).*

(ii) *Clocks in the nose and rear of the ship of k get out of synchronism.*

The **proof** goes by the methods of the earlier similar theorems, and is left to the reader.

Remark 2.5.14 If we omit the condition $\mathbf{Ax}(\sqrt{\ })$ from Theorem 2.5.13 above, then the theorem remains basically true but the formulation gets more complicated, cf. e.g. the formulation of Theorem 2.5.7(iii). Further, in Theorem 2.5.13, $\mathbf{Ax}(\sqrt{\ })$ can be replaced with the weaker assumption that f_{mk} is *betweenness-preserving*[124]. $\triangleleft$

[122]Here the emphasis is on that it is consistent with **Basax** that both m and k think that the other's x-meter-rods shrink.

[123]I.e. even if the meter-rod is k's one we check whether m sees it pointing in direction $\bar{y}$.

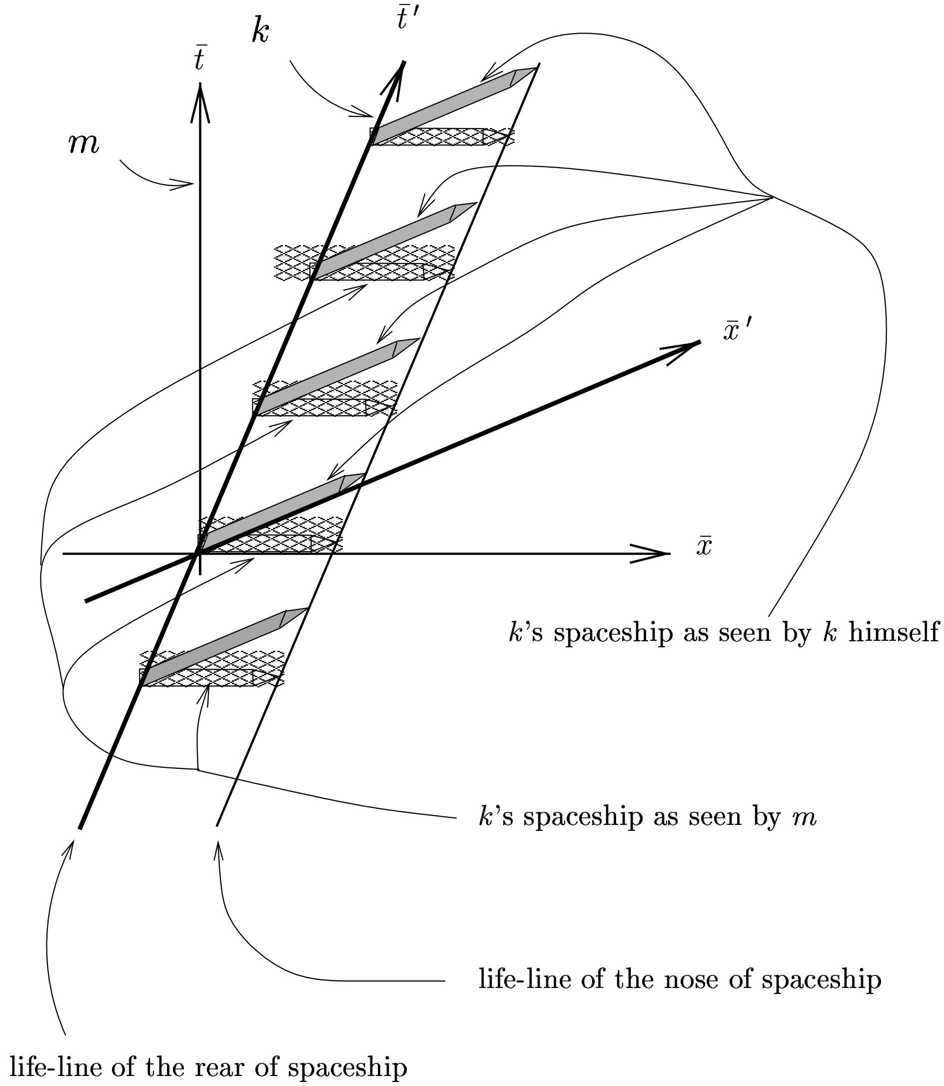
[124]The ternary relation Betw of betweenness will be defined in footnote 405 on p.492 in §4. Intuitively, for $p, q, r \in {}^nF$ the relation $\mathsf{Betw}(p, q, r)$ holds if p, q, r are collinear and q is between p and r.

The relativistic effects (I)–(III) discussed so far will lead e.g. to the famous <u>*twin paradox*</u>. However, for that we will need a symmetry axiom **Ax(symm)** discussed in section "A symmetry axiom" (§2.8). So, we will return to the twin paradox in §2.8, cf. Thm.2.8.18 (p.140). An even more satisfactory discussion of this paradox can be given by looking at accelerated observers, hence we will return to the "twin" in §8 ("Accelerated observers") again.

We will see in §2.8 how **Ax(symm)** (introduced in §2.8) influences the paradigmatic effects (I)–(III). E.g. we will see that these effects (for example the effect of clocks slowing down) admit a simpler and stronger formulation in **Basax** + **Ax(symm)** than in pure **Basax**. Cf. Items 2.8.7–2.8.12.

As we already said, in a later chapter we will return to seeing how paradigmatic effects (I)–(III) (discussed in items 2.5.2– 2.5.11) change if we use more subtle (than **Basax**) axiom systems, cf. §4, in particular section §4.8, and Figure 223 on p.653.

Our next figure illustrates the meter-rod shrinking effect, i.e. items 2.5.9–2.5.12. To be more intuitive, we draw spaceships instead of meter-rods. The figure represents how observers m and k see k's spaceship.



Here $\bar{t}\,'$ and $\bar{x}\,'$ denote the respective coordinate axes of observer k.

Let us see how the above picture (illustration of paradigmatic effect (II), i.e. of items 2.5.9–2.5.12) looks like if k moves faster than light relative to m. By §2.4, if $n = 2$ then faster than light observers are possible. But this leads us to the subject of our section 2.7. Therefore our next similar picture comes in section 2.7.

2.6 Are all three paradigmatic effects necessary?

A central question that motivates our interest in the models of **Basax** is the following:

($\star$) Are all three paradigmatic effects discussed in section 2.5 necessary consequences of special relativity? If not, are they independent of one another?

It is question ($\star$) that triggers our interest in the class $\mathsf{Mod}(\mathbf{Basax})$, i.e. in the question how different the models of **Basax** can be from each other. [125]

Turning to the question itself, the following idea naturally comes to one's mind. Are all the paradigmatic effects, items (I) to (III) on page 90, necessary consequences of **Basax**? If not, in which combinations can they occur? Cf. also the pictures on p. 88.

First, concerning effect (III) (moving clocks get out of synchronism), the answer is simple. It *is* a necessary consequence of **Basax**, by Theorem 2.5.7. That is, whenever $v_m(k) > 0$, m thinks that the clocks in the nose and in the rear of k's spaceship are out of synchronism (provided, of course, that k thinks they are synchronized). This is so in every model of **Basax** and for every $m, k \in Obs$.

On the other hand, to answer the question as far as the other effects are concerned, we must pose it more precisely. Let us fix an observer m_0 in a model $\mathfrak{M}$ of **Basax**. We shall think in m_0's world view. For example, "k moves" means that k moves relative to m_0. Now, we shall seek for the answers to our question in a systematic manner (cf. items 1-3 below).

1. As we have already pointed out, paradigmatic effect (III) must be true in m_0's world view.

2. Clocks do *not* necessarily slow down on moving spaceships (i.e. effect (I) is not necessary). More formally, there are a model $\mathfrak{M} \models \mathbf{Basax}$ and an observer $m_0 \in Obs^{\mathfrak{M}}$ such that

   $(\forall k \in Obs)[m_0$ thinks that k's clocks tick with exactly the same rate as his clocks$]$. [126]

   Such a model $\mathfrak{M}$ (with a distinguished m_0) is represented in Figure 29 (p. 88) under the name $\mathfrak{M}_1$. That model is 2-dimensional, but it can be extended to 3

[125] This interest will lead us to the investigations in §3.6 (Models of **Basax**), as well as to the study of non-elementarily equivalent models of **Basax** in e.g. Theorem 3.8.18(ii) on p. 303. Cf. also Remark 3.6.15 on p. 271.

[126] We emphasize again that m_0 thinks that *all observers* have clocks running with the correct rate.

or 4 dimensions, too. However, this generalization from $n = 2$ to $n \geq 3$ is not completely straightforward. We do not go into the details here, but the key idea is described in §3.2. We should mention one difference between the cases $n = 2$ and $n = 3$. In the case $n = 2$ we have a so-called Minkowski-sphere around the origin which, assuming $\mathsf{f}_{m_0 k}(\bar{0}) = \bar{0}$, tells us for each k how long its unit-vectors are (i.e. where $\mathsf{f}_{m_0 k}(1_i)$ is). This sphere works uniformly for all choices of $k \in Obs$ (assuming $\mathsf{f}_{m_0 k}(\bar{0}) = \bar{0}$). By contrast, in the case of $n = 3$ all we know is that all the points $\mathsf{f}_{m_0 k}(1_i)$ are in a horizontal plane as depicted in Figure 39. However, after 1_t^k is determined by this plane for each choice of

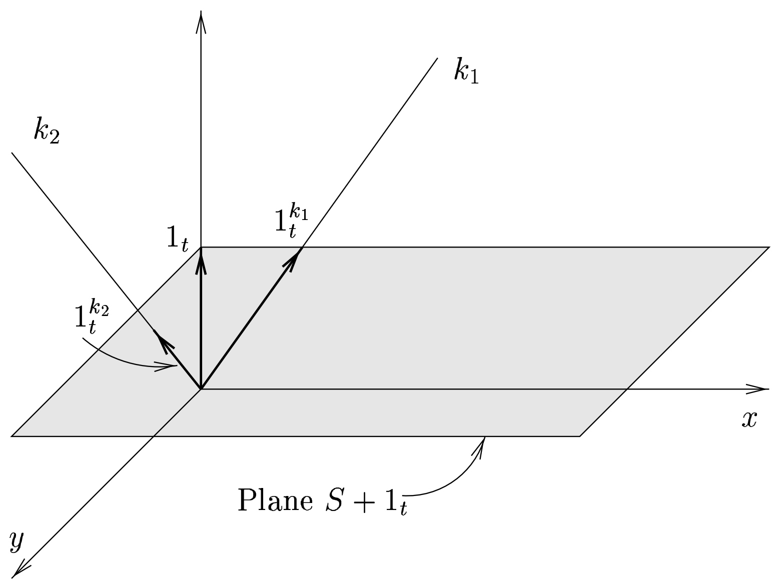


Figure 39: Illustration for item 2.

k, we still have to fix the rest of k's unit vectors as it is done in §3.2. (Where it is shown that choosing 1_t^k *arbitrarily*, the other unit vectors can be fixed so that the axioms of **Basax** are validated.)

In the second part of this section, when discussing the independence of the particular paradigmatic effects, we shall concentrate on the case $n = 2$; but all the results can be generalized to $n \geq 3$, analogously to the generalization indicated in item 2 above. We invite the interested reader to figure out what the answers look like for $n = 3$ first, and later to all $n \geq 3$.

Let us return to discussing what happens if m_0 thinks that all moving clocks tick with the correct rate (i.e., no clocks slow down or run fast). The present answer to the question ($\star$) (namely, that effect (I) is not necessary) applies if we are allowed to fix $\mathfrak{M}$ *and* m_0. However, in the same model $\mathfrak{M}$ there will be an observer m_1 who thinks that moving clocks *do* slow down. Indeed, if $v_{m_0}(k) > 0$, then k will think that m_0's clocks do slow down (and they slow down more than would be necessary if we did not force k's clocks not to slow down for m_0; this will be implicitly seen in §2.8).

3. Similarly to item 2, moving spaceships *need not* shrink. That is, there are $\mathfrak{M} \models$ **Basax** and $m_0 \in Obs^{\mathfrak{M}}$ such that in m_0's world view moving spaceships do not shrink. Formally,

$$(\forall k \in Obs)[m_0 \text{ thinks that } k\text{'s meter-rods are of the correct length}].$$

(The reader is invited to formalize this statement in our frame language.) The model proving this claim is remotely similar to $\mathfrak{M}_3$ in Figure 29 (p. 88), but the functions that are parts of the Minkowski circle must grow faster in the model. The reader is invited to construct (and draw) such a model (based on the world view of some m_0 in which *no* spaceship shrinks). For this exercise it might be useful to consult Figure 38 on p. 101 proving that spaceships "usually" do shrink. On the other hand, see picture 40.

By items 2 and 3 above we received permissive answers to our question concerning the removability (or changeability) of the paradigmatic effects. The second part of our question asked how *independent* effects (I)-(III) are of one another.

The paradigmatic effects (I)-(III) are *not* independent. We have already seen that effect (III) (violation of synchronism) is necessary. Further, assume $v_{m_0}(k) > 0$, and consider m_0's world view. The following holds:

$$(k\text{'s clocks do } \textit{not} \text{ slow down}) \Rightarrow (k\text{'s spaceship shrinks}).$$

Similarly,

$$(k\text{'s spaceship does } \textit{not} \text{ shrink}) \Rightarrow (k\text{'s clocks slow down}).$$

These statements are stated in Theorem 2.5.13. Actually in Thm. 2.5.13, $\mathbf{Ax}(\sqrt{\ })$ was assumed, but we conjecture that it is not needed here, because we are stating only that

($\star$) (k's clocks always show the correct time) $\Rightarrow$ (one of k's meter rods[127] shrinks),

[127] Namely, the one pointing in the direction of k's movement.

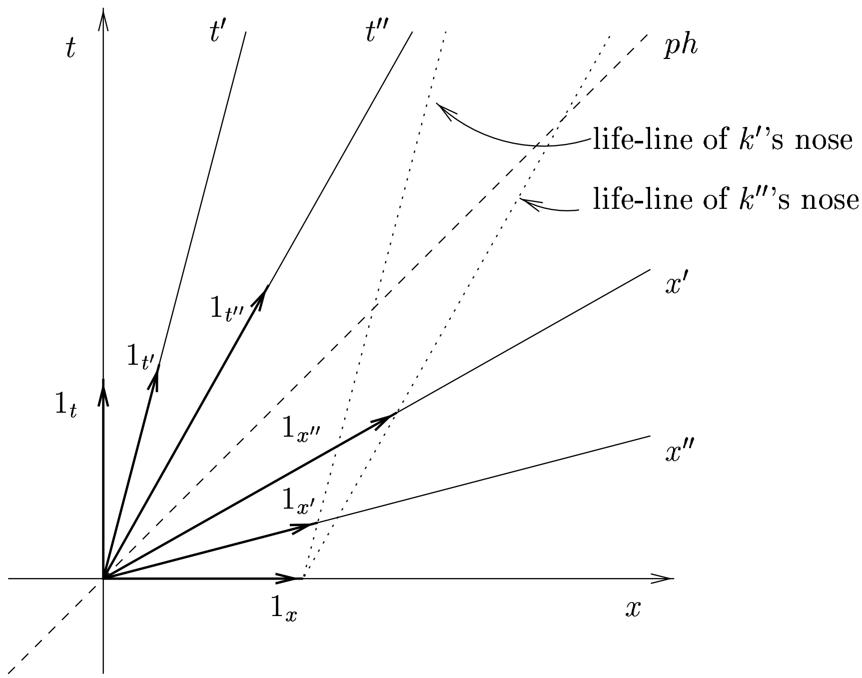


Figure 40: Illustration for item 3.

that is,

$$(\forall p \in \bar{t})p_t = (\mathsf{f}_{km}(p))_t \Rightarrow \text{(one of } k\text{'s meter rods shrinks)}.$$

It seems to us that **Basax** $\models (\star)$, but we have not checked this claim carefully.

To sum up: On the one hand, we can get rid of effect (I), but *then* we must have (II) and (III).[128] On the other hand, we can get rid of effect (II), but *then* we must have (I) and (III). So, for a possible observer m_0 in a possible model $\mathfrak{M} \models$ **Basax**, the following combinations can be realized:

(A) All three effects (I), (II) and (III) are experienced by m_0.

(B) Effects (I) and (III) prevail, but (II) does not.

(C) Effect (I) is not present, but (II) and (III) are.

There are no other possibilities. Thus we have at least three essentially different classes of models for **Basax**,[129] and this fact triggers our interest in asking how

[128]Moreover, we pay for not having (I) by having (II) to a higher extent.

[129]Say, $\mathfrak{M}'$ is such that all observers are of type (A), $\mathfrak{M}''$ is such that it has both (A) and (B) type observers but none of type (C), and $\mathfrak{M}'''$ has observers of type (A) and (C), but none of type

many, and what sorts of, *non-elementarily equivalent* models **Basax** has. The reason why we talk about non-elementarily equivalent models is that this expression means that the models in question are not only different (i.e., non-isomorphic), but they are actually distinguishable by a formula in our frame language like (the formalized versions of) (A), (B) and (C) are. Actually, the above mentioned three classes of models are distinguishable by thought experiments too, which might be a stronger notion of distinguishability (than the one using formulae). It would be interesting to see how many classes of models of **Basax** are distinguishable by means of thought experiments, but for this purpose we would need to define which formulas of our frame language count as thought experiments. We do not deal with this issue here.[130]

Later, in §2.8, we shall introduce a natural axiom $\mathbf{Ax}(\|)$ saying that observers not moving relative to each other see the world *essentially* the same way. We mention this because the presently discussed issue concerning the connection between the paradigmatic effects is even more interesting in $\mathbf{Basax}+\mathbf{Ax}(\|)$ than in pure **Basax**. Therefore we mention that the answer to our question remains exactly the same, i.e. cases (A) to (C) are all the possibilities, for $\mathbf{Basax} + \mathbf{Ax}(\|)$ too. Actually, most of those axioms to be introduced that we will call *auxiliary axioms* (cf. §3.8 on **BaCo**) leave the answer to the present question *unchanged* (i.e., (A), (B), (C) remain possible). For example, for $\mathbf{Basax}+\mathbf{Ax}(\sqrt{\ })+\mathbf{Ax}(\|)+\mathbf{Ax}(\mathit{Triv})+\mathbf{Ax}(\uparrow)+\mathbf{Ax}(\mathbf{ext}) + \mathbf{Ax}(\mathbf{rc})$ the situation is the same as outlined above for **Basax**.[131] We could even add the continuity Principle ($\star$) from §5 (on $\mathbf{Bax}(\mathbf{nop})$) to the axioms without changing this result. However, the symmetry axioms, e.g. **Ax(symm)** to be introduced soon, in §2.8, will change this picture.

(B). It takes some time to check that $\mathfrak{M}''$ and $\mathfrak{M}'''$ exist, but they do. We omit the proof of this claim.

[130]Cf. also Remark 3.6.15 for further philosophical reasons for studying non-elementarily equivalent models of **Basax**.

[131]The mentioned extra axioms can be found in §3.8.

2.7 Faster than light in two dimensions

In §2.4 we saw that if $n = 2$ then faster than light (FTL) observers are possible, more formally the existence of FTL observers is consistent with **Basax**(2). In the present section we will briefly discuss how these FTL observers behave, and what their world looks like. For completeness we note that **Basax**(n) with $n > 2$ excludes FTL observers, cf. Thm.3.4.1 (p.203). However, there are refinements of **Basax** which allow FTL observers for $n > 2$ too. Cf. e.g. the axiom system **Relphax** and Thm.3.4.22 (p.223).

Our first theorem states that, under assuming **Basax**+**Ax**($\sqrt{\ }$), if m sees k moving FTL, then k sees m moving FTL, too.

THEOREM 2.7.1 *Assume* **Basax**(n) + **Ax**($\sqrt{\ }$). *Let* $m, k \in Obs$. *Then*

$$v_m(k) < 1 \;\Leftrightarrow\; v_k(m) < 1\,, \qquad \textit{and therefore}$$
$$v_m(k) > 1 \;\Leftrightarrow\; v_k(m) > 1\,.$$

The **proof** will be filled in later. ∎

The next theorem says the following. Assume that $\mathfrak{M}$ is a frame model of **Basax**(2) + **Ax**($\sqrt{\ }$) and that there is an FTL observer in $\mathfrak{M}$. Then the observers are classified into two disjoint "worlds", and inhabitants of the same world see each other as "normal" slower than light observers, while observers coming from different worlds see each other as FTL ones.

THEOREM 2.7.2 *Assume* $\mathfrak{M} \in \mathsf{Mod}(\mathbf{Basax}(2) + \mathbf{Ax}(\sqrt{\ }))$. *Let* $STL \subseteq Obs \times Obs$ *be a binary relation defined as follows.*

$$(\forall m, k \in Obs) \quad m\ STL\ k \;\overset{\text{def}}{\Longleftrightarrow}\; v_m(k) < 1\,.$$

(m STL k means that k moves slower than light as seen by m).
Then (i) and (ii) below hold.

(i) *STL is an equivalence relation on the set of observers Obs.*

(ii) *Assume that in $\mathfrak{M}$ there is an FTL observer, i.e.* $(\exists m, k \in Obs)\ v_m(k) > 1$. *Then the equivalence relation STL has exactly two equivalence classes.*

The **proof** will be filled in later. ∎

The following theorem says that Thm.2.7.2 above does not remain true if we omit the condition $\mathbf{Ax}(\sqrt{\ })$.

THEOREM 2.7.3 *There are* $\mathfrak{M} \in \mathsf{Mod}(\mathbf{Basax}(2))$ *and* $m, k \in Obs^{\mathfrak{M}}$ *such that* $v_m(k) < 1$ *and* $v_k(m) > 1$.
Intuitively, observer k *sees* m *moving FTL, while* k *does not see* m *moving FTL.*

Proof: Let $\mathfrak{F} = \langle \mathbf{F}, \leq \rangle$ be an arbitrary ordered field such that its field reduct $\mathbf{F}$ has an automorphism, call it φ, which is not order preserving, i.e. φ is not an automorphism of $\mathfrak{F}$. Let $\mathfrak{M} = \langle (B;\ Obs, Ph, Ib), \mathfrak{F}, G;\ \in, W \rangle$ be a frame model constructed as the model $\mathfrak{M}_1^P$ was constructed in §2.4 (for arbitrary P), i.e. $Obs := \mathsf{Eucl}(2, \mathfrak{F})$ etc; the only difference (between our $\mathfrak{M}$ and $\mathfrak{M}_1^P$ in §2.4) is that $\mathfrak{M}$ is constructed over $\mathfrak{F}$ while $\mathfrak{M}_1^P$ was constructed over the ordered field of reals $\mathfrak{R}$.

There are FTL observers in $\mathfrak{M}$, but also

$$(6) \qquad (\forall m, k \in Obs)\ (v_m(k) < 1 \ \Leftrightarrow\ v_k(m) < 1)$$

holds in $\mathfrak{M}$. We will use the non-order preserving automorphism φ to modify $\mathfrak{M}$ obtaining a model $\mathfrak{M}^+$ (of **Basax**) in which (6) above will fail. Let $\widetilde{\varphi} : {}^2F \longrightarrow {}^2F$ be the function induced by $\varphi : F \longrightarrow F$ as follows: $(\forall x, y \in F)\ \widetilde{\varphi}(\langle x, y \rangle) :\stackrel{\text{def}}{=} \langle \varphi(x), \varphi(y) \rangle$. We note that $\widetilde{\varphi}$ takes straight lines to straight lines, i.e. $(\forall \ell \in \mathsf{Eucl})\ \widetilde{\varphi}[\ell] \in \mathsf{Eucl}$. The modified version $\mathfrak{M}^+$ of $\mathfrak{M}$ will differ from $\mathfrak{M}$ only in the world-view relation. We will denote the world-view relation of $\mathfrak{M}^+$ by W^+. Let $m \in Obs$ be arbitrary, but fixed. Then for every $k \in Obs$ we define the world-view function of k as follows:

$$w_k^+ :\stackrel{\text{def}}{=} \begin{cases} \widetilde{\varphi} \circ w_k & \text{if } k = m \\ w_k & \text{otherwise.} \end{cases}$$

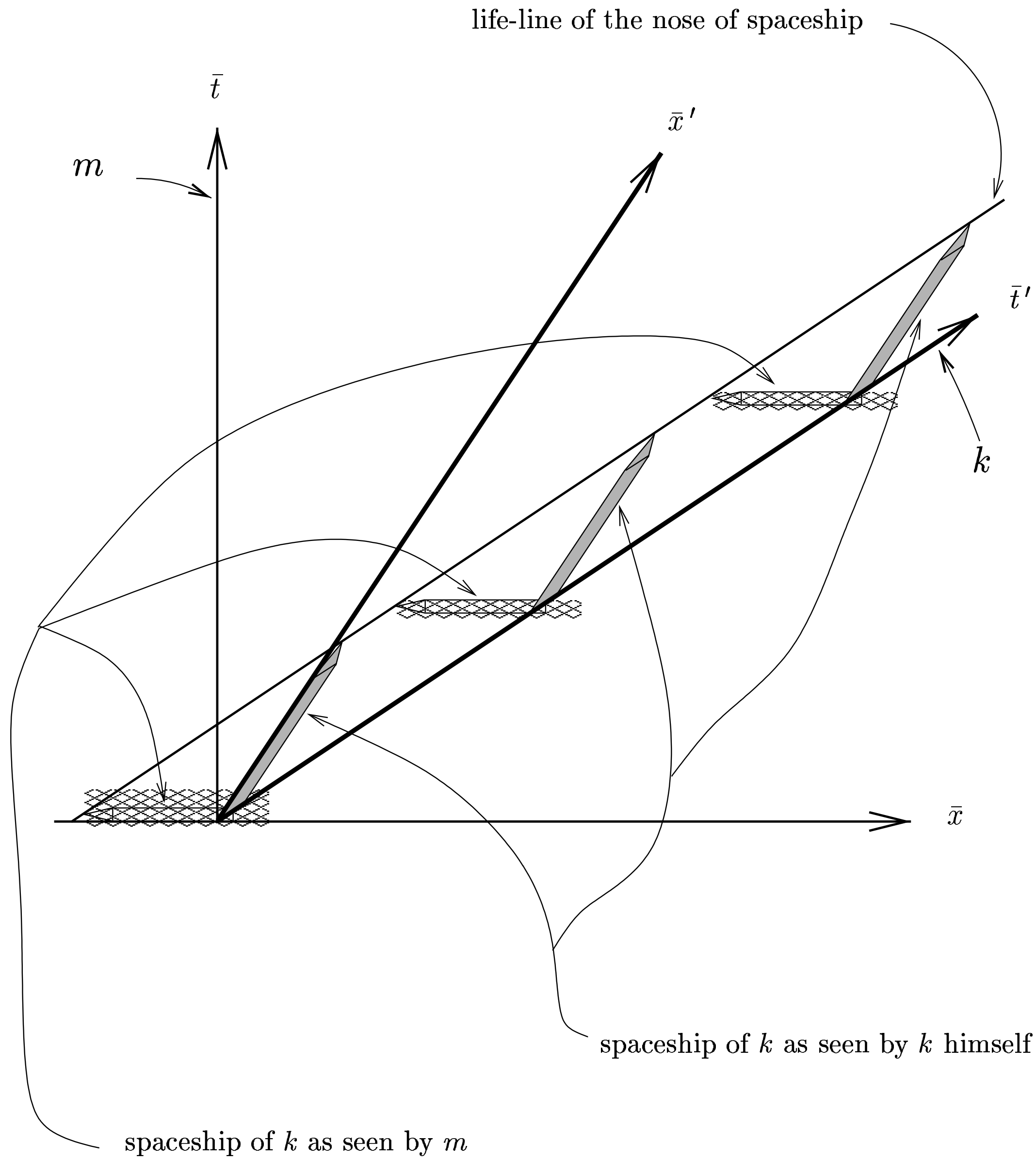
We define W^+ from w_k^+'s the obvious way. By the above $\mathfrak{M}^+$ is defined. Checking that $\mathfrak{M}^+ \models \mathbf{Basax}$ is left to the reader.

Now we show that (6) above fails in $\mathfrak{M}$. Let $k \in Obs := \mathsf{Eucl}$ be such that $k \notin \mathsf{SlowEucl}$, but $\widetilde{\varphi}[k_1] = k$, for some $k_1 \in \mathsf{SlowEucl}$.[132] Fix this k_1. The "speed" in $\mathfrak{M}^+$ is denoted by v^+ while the "speed" in $\mathfrak{M}$ is denoted, as usual, by v. Now, it is easy to see that $v_m^+(k) = v_m(k_1)$ and $v_k^+(m) = v_k(m)$ since $w_m^+ = \widetilde{\varphi} \circ w_m$ and $w_k^+ = w_k$. Hence $v_m^+(k) < 1$ and $v_k^+(m) > 1$. ∎

Let us return to the subject matter (and style) of the picture on p.104. In particular let us see how FTL (w.r.t. each other) observers see each other's spaceships. After that, we will also draw clocks and even some "visual effects".

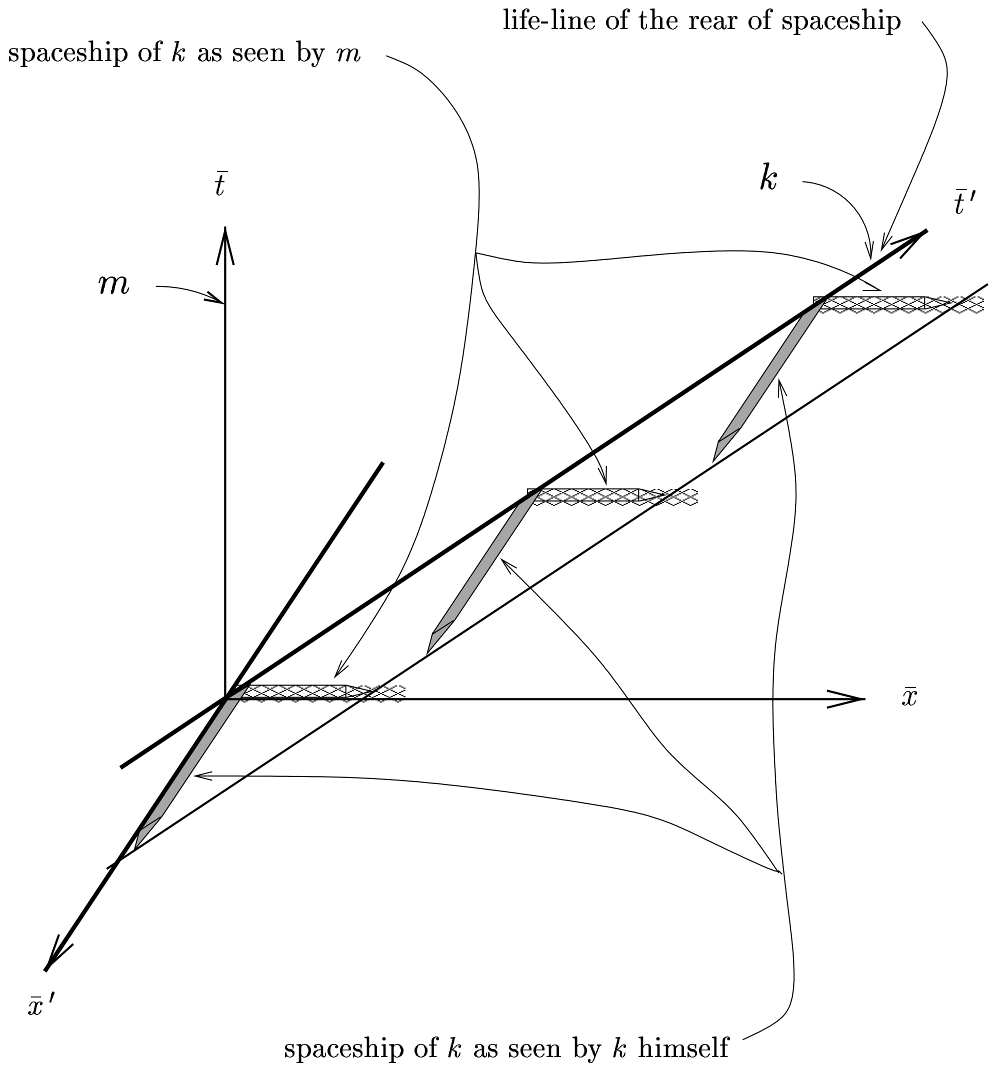
[132] Such a k exists since φ is non-order preserving.

Throughout the rest of this section (§2.7) **Basax**(2) + **Ax**($\sqrt{\ }$) *is assumed*, unless otherwise specified.

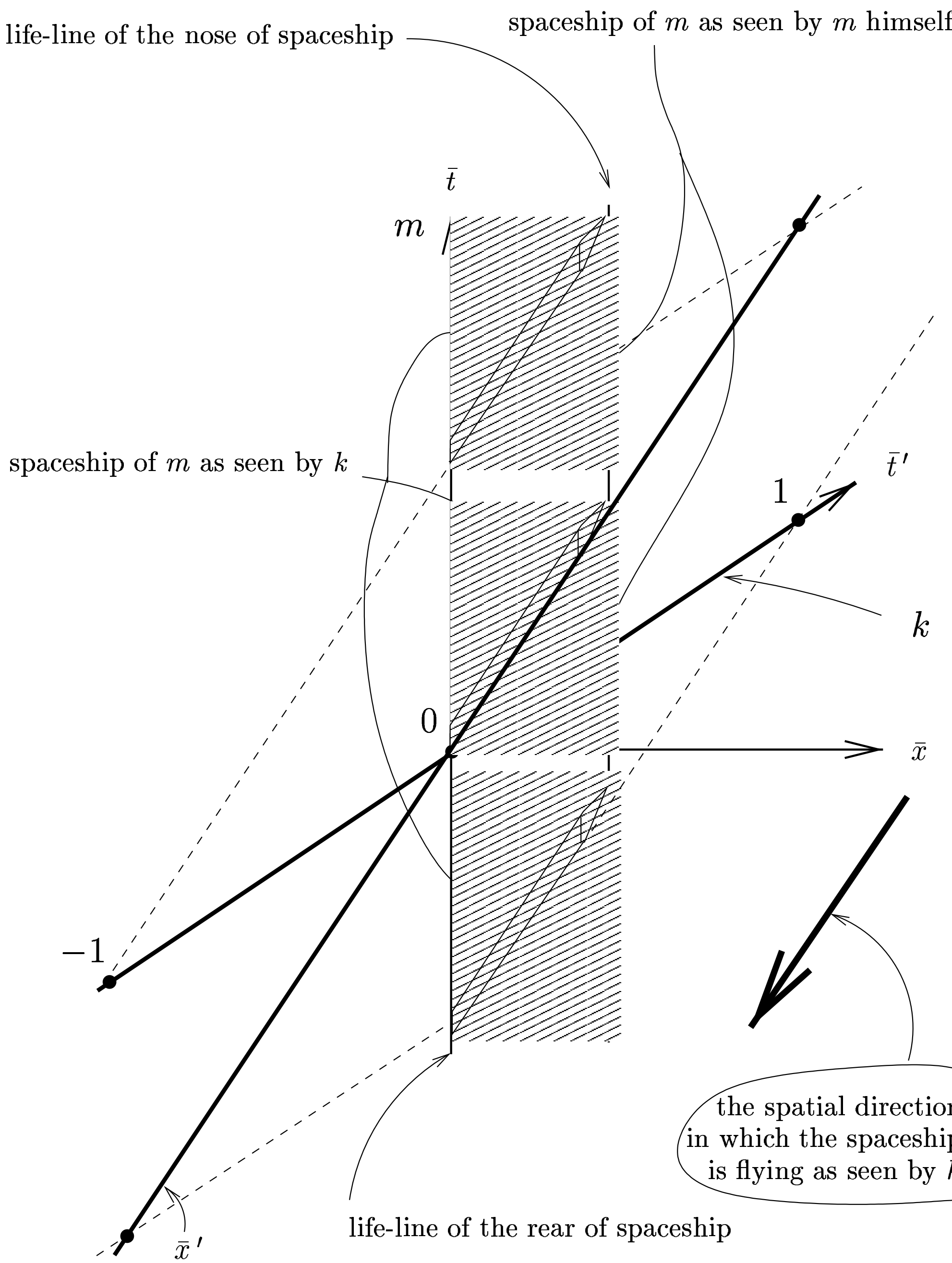


We are in **Basax**(2)+**Ax**($\sqrt{\ }$). The above picture shows meter-rods (represented as spaceships) of FTL observer k as seen by observer m. Notice that the spaceship is "flying" backwards as seen by m (i.e. moving in the direction of its rear). However, if k points the nose of his spaceship in the opposite direction then m will see the spaceship of k "flying" forwards. Formally, this amounts to reversing the direction of the $\bar{x}$-axis of k. This is illustrated in the following picture.

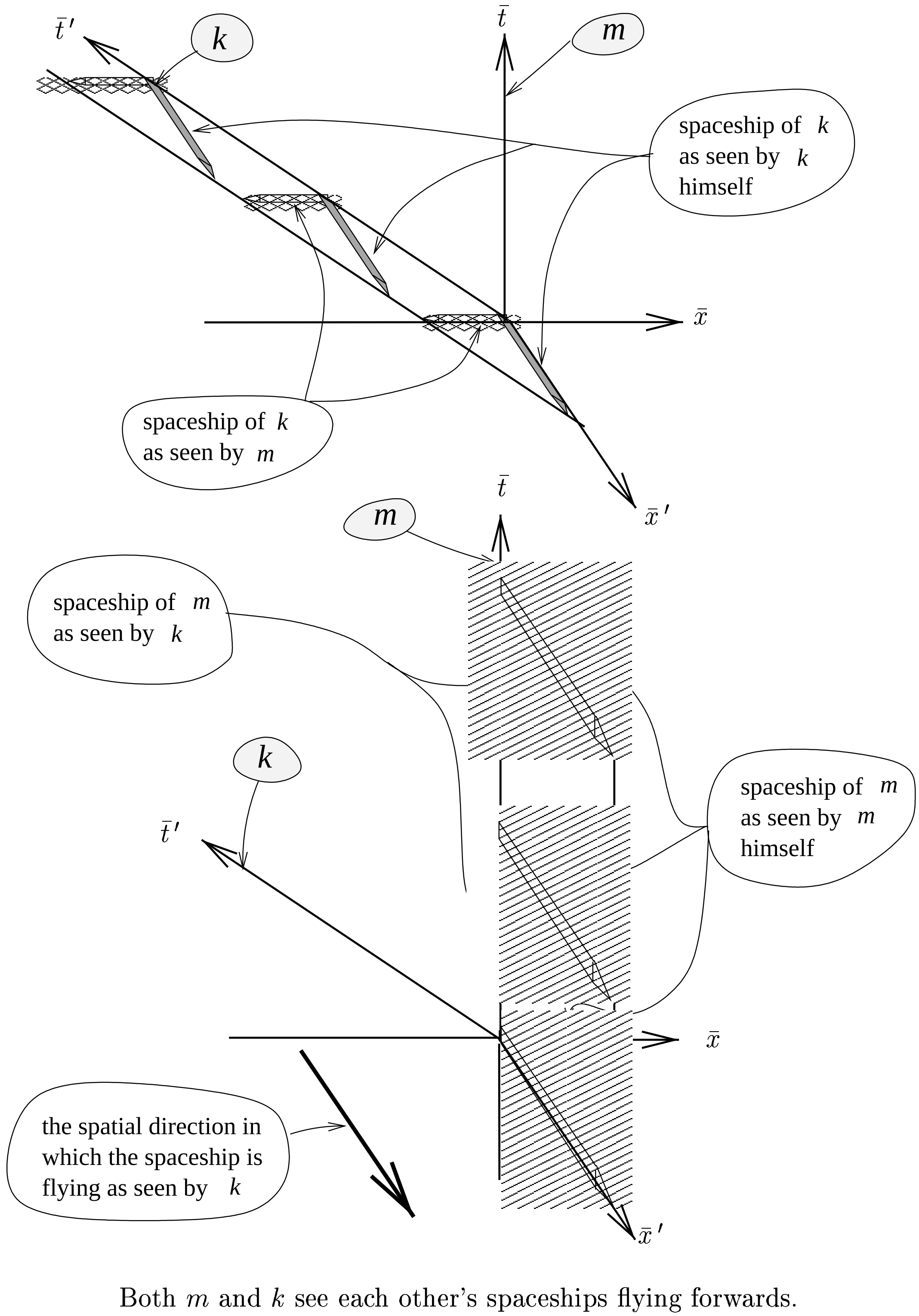
Convention: Throughout we assume that a spaceship always points its nose in the positive direction of its $\bar{x}$-axis.



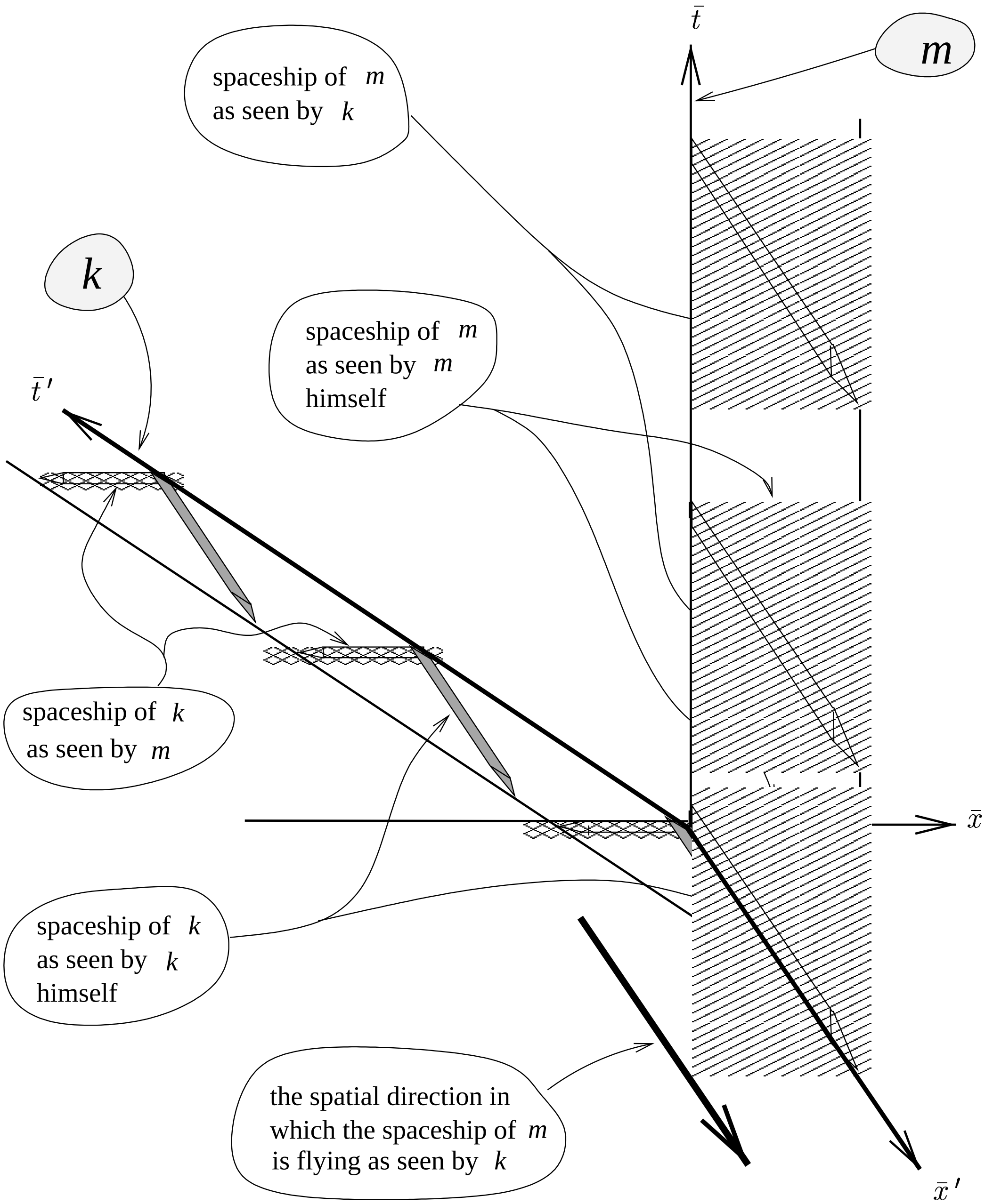
This was the spaceship of k; in turn the next picture illustrates how k sees the spaceship of m.



k sees that m's spaceship is flying backwards (i.e. moving in the direction of its rear). However, if m points the nose of his spaceship in the opposite direction then both k and m will see each other's spaceships flying forwards (i.e. moving in the directions of their noses). This is illustrated on the next two pages (pp.115–116).



Both m and k see each other's spaceships flying forwards.



Both m and k see each other's spaceships flying forwards.

The pictures above show that it is possible to have a kind of symmetry when considering the spatial directions in which the spaceships are flying, that is there are observers m and k such that k moves FTL relative to m, and both m and k see each other's spaceships flying forwards, i.e. the spaceships are moving in direction of their noses. However, this symmetry is <u>*not a perfect symmetry*</u>, namely for every $m, k \in Obs$ with $v_m(k) > 1$ the following holds: If m thinks that k's spaceship is flying forwards in the negative direction of $\bar{x}$ then k will think either that k's spaceship is flying forwards in the positive direction of $\bar{x}\,'$ or it is flying backwards (i.e. it is flying in the direction of its rear). If m thinks that k's spaceship is flying forwards in the positive direction of $\bar{x}$ then k will think either that k's spaceship is flying forwards in the negative direction of $\bar{x}\,'$ or it is flying backwards. Similar arguments apply <u>*in the cases*</u> when m thinks that k is flying backwards. We leave the details of this to the reader.

As a contrast if $v_m(k) < 1$ and if time passes forwards for k as seen by m then the above lamented "perfect symmetry" is achievable, namely if m sees k flying forwards in the negative direction of $\bar{x}$ then k will see that m is doing the same.

Let us switch from regarding spaceships to regarding their clocks. In Thm.2.7.4 and in Figure 41 below we will see that it is impossible to have symmetry when considering the directions of "flows of time" (i.e. considering whether the observed clocks run forwards or backwards).

THEOREM 2.7.4 *Assume* $\mathbf{Basax}(n) + \mathbf{Ax}(\sqrt{\ })$.[133] *Let* $m, k \in Obs$. *Assume* k *moves FTL relative to* m, *i.e.* $v_m(k) \geq 1$. *Then the following hold. If* m *thinks that* k*'s clock runs forwards then* k *will think that* m*'s clock runs backwards. However, if* m *thinks that* k*'s clock runs backwards then* k *will think that* m*'s clock runs forwards.*[134] *Summing up,* m *and* k *see each other's clocks differently. Formally:*

$$\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t > 0 \quad \Leftrightarrow \quad \mathsf{f}_{mk}(1_t)_t - \mathsf{f}_{mk}(\bar{0})_t < 0 .$$

On the proof: By Prop.2.3.3(iii), we may assume $v_m(k) > 1$. Then for $n = 2$, the idea of the proof is illustrated in Figure 41. For $n > 2$, one either checks that the idea represented in Figure 41 works; or equivalently one may use the no FTL theorems in §3.4.1. The detailed proof will be filled in a later version. ∎

[133] We note that a variant of this theorem remains true without $\mathbf{Ax}(\sqrt{\ })$, i.e. in pure **Basax**.

[134] Sometimes we quote this theorem as if it stated "... k's clocks run backwards ...". In these quotations we have in mind the clock in the rear of k's spaceship together with the clock in the nose of the spaceship etc. and that is why we write in the plural k's clocks instead of just k's clock as the theorem above says.

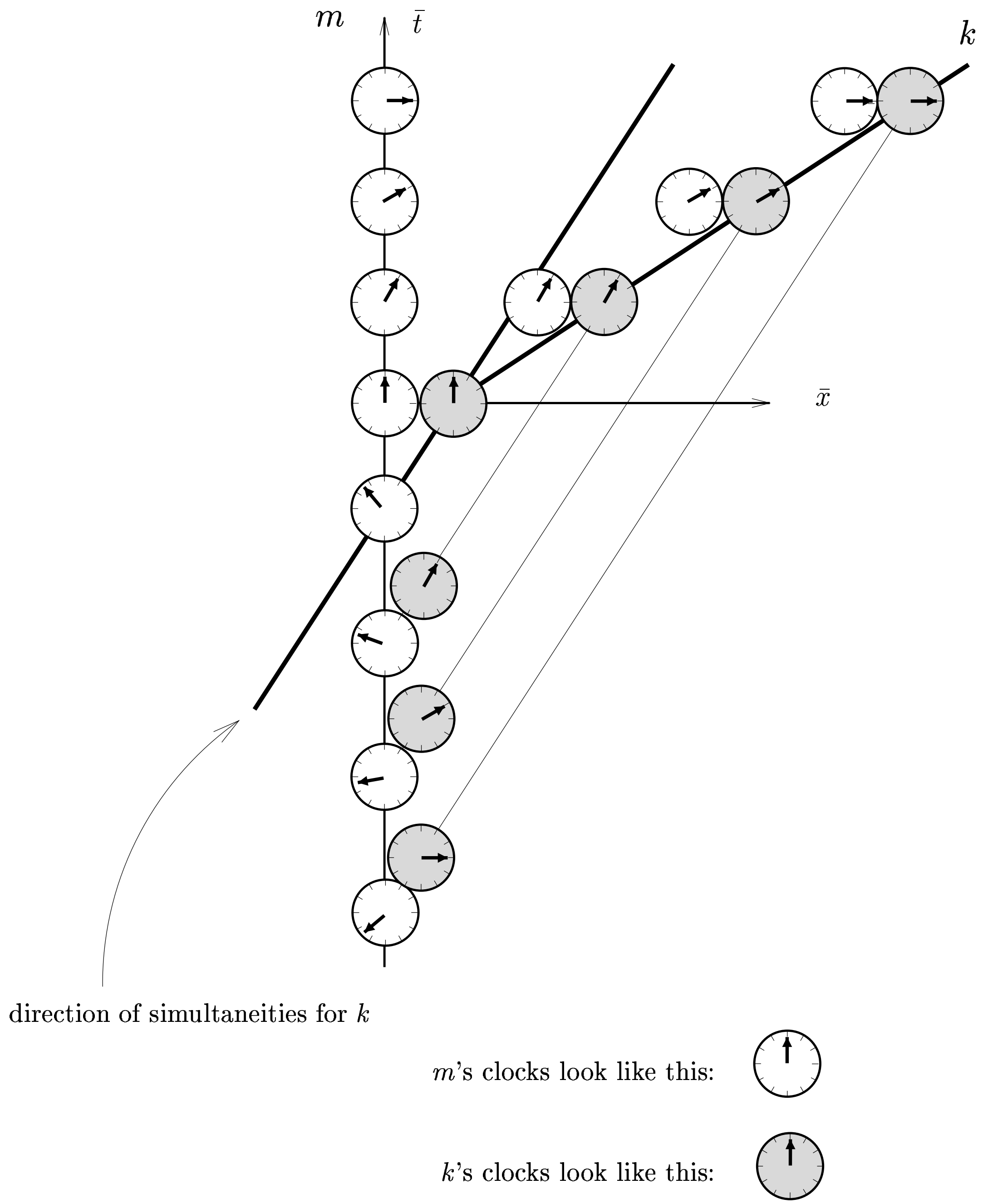


Figure 41: Assume that m thinks that k's clock runs forwards. Then k will think that m's clock runs backwards.

THEOREM 2.7.5 (FTL excludes “perfect symmetry”.)
Assume **Basax**(n). *Let* $m, k \in Obs$. *If* $v_m(k) \geq 1$ *then* $\mathsf{f}_{mk} \neq \mathsf{f}_{km}$.

Proof: Under assuming $\mathbf{Ax}(\sqrt{\ })$ the theorem is an immediate corollary of Thm.2.7.4 above. The proof without assuming $\mathbf{Ax}(\sqrt{\ })$ will be filled in later. ∎

Thm.2.7.5 above leads us to the following considerations. In §3.9 (“Symmetry axioms”) we will introduce two symmetry principles $\mathbf{Ax}\triangle\mathbf{1}$ and $\mathbf{Ax}\Box\mathbf{1}$ which can be regarded as special cases of Einstein's Special Principle of Relativity.[135]

COROLLARY 2.7.6
$\mathbf{Basax}(2) + \mathbf{Ax(eqtime)} + \mathbf{Ax}\triangle\mathbf{1} \models$ “$\nexists$ **FTL observers**”, where **Ax(eqtime)** will be defined in §2.8 on p.127.

◁

The above is an immediate corollary of Thm.2.7.5.

Conjecture 2.7.7 *Assume* **Basax**(2). *Then FTL observers are consistent with* $\mathbf{Ax}\Box\mathbf{1}$.

◁

In connections with items 2.7.5–2.7.7 above cf. Thm.2.8.2 on p.127 and Thm.3.8.11 on p.300.

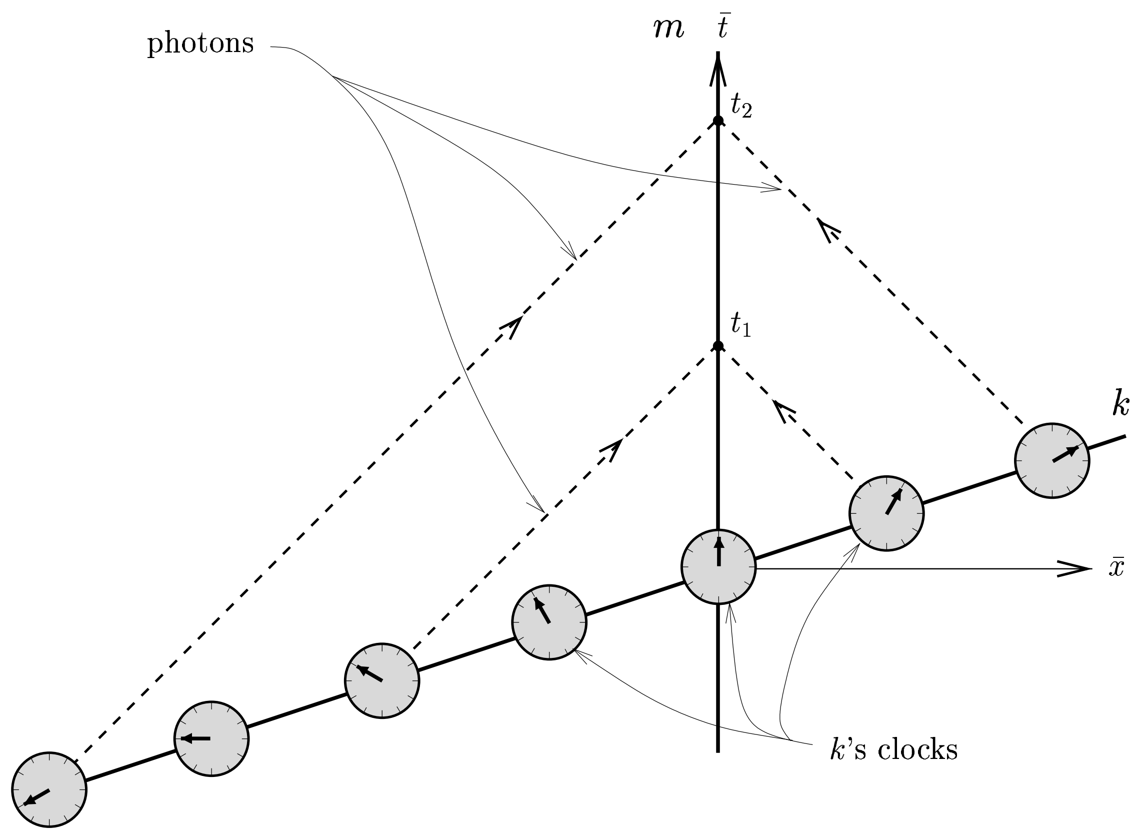
Visual effects in FTL

Now we turn to discussing, briefly, what we call “visual effects”, i.e. what observers see in the *usual* sense of seeing, i.e. via photons arriving “at their eyes”. (Recall that in the rest of this work we use the word seeing in a completely *different* sense namely we use it as “coordinatizing”.)

Let us recall that in the present section (§2.7) $\mathbf{Basax}(2) + \mathbf{Ax}(\sqrt{\ })$ is assumed. Let $m, k \in Obs$ such that k moves FTL relative to m. The way m *optically* (i.e. via photons) sees k is represented in the following picture. Studying the picture reveals the following. The experience of m concerning k is this: For a while there is nothing just empty space. (In particular no k is around). Then suddenly, out of nothing *two* copies of k appear and they start moving away from each other. Both copies of k are moving *away* from m one copy is getting *younger* and younger, while the other one is getting older (as normal). I.e. m optically sees the clocks of one copy of k ticking forwards while sees the other clock ticking backwards.

[135] $\mathbf{Ax}\triangle\mathbf{1}$ is $(\forall m, k \in Obs)(\exists k' \in Obs)(tr_m(k) = tr_m(k') \wedge \mathsf{f}_{mk'} = \mathsf{f}_{k'm}$ and $\mathbf{Ax}\Box\mathbf{1}$ is $(\forall m, k, m' \in Obs)(\exists k' \in Obs)\mathsf{f}_{mk} = \mathsf{f}_{m'k'}$.

Because of the no FTL theorem in §3.4.1 one might have the impression that the just described phenomenon can happen only in two dimensions. However, in $\mathbf{Basax}(n) + \mathbf{Ax}(\sqrt{\ })$ with $n > 2$, we still can have this with the only difference that k is *not* an observer but only a body *having an "inner clock"*. Such FTL bodies *are* permitted by $\mathbf{Basax}(n)$, actually they are more or less the same what the literature calls *tachyons*. Bodies with inner clocks will be discussed in a future chapter of this work (which chapter does already exist in unfinished hand-written form).[136] (This might remotely remind one of the case when a particle and its anti-particle is created from nothing.)

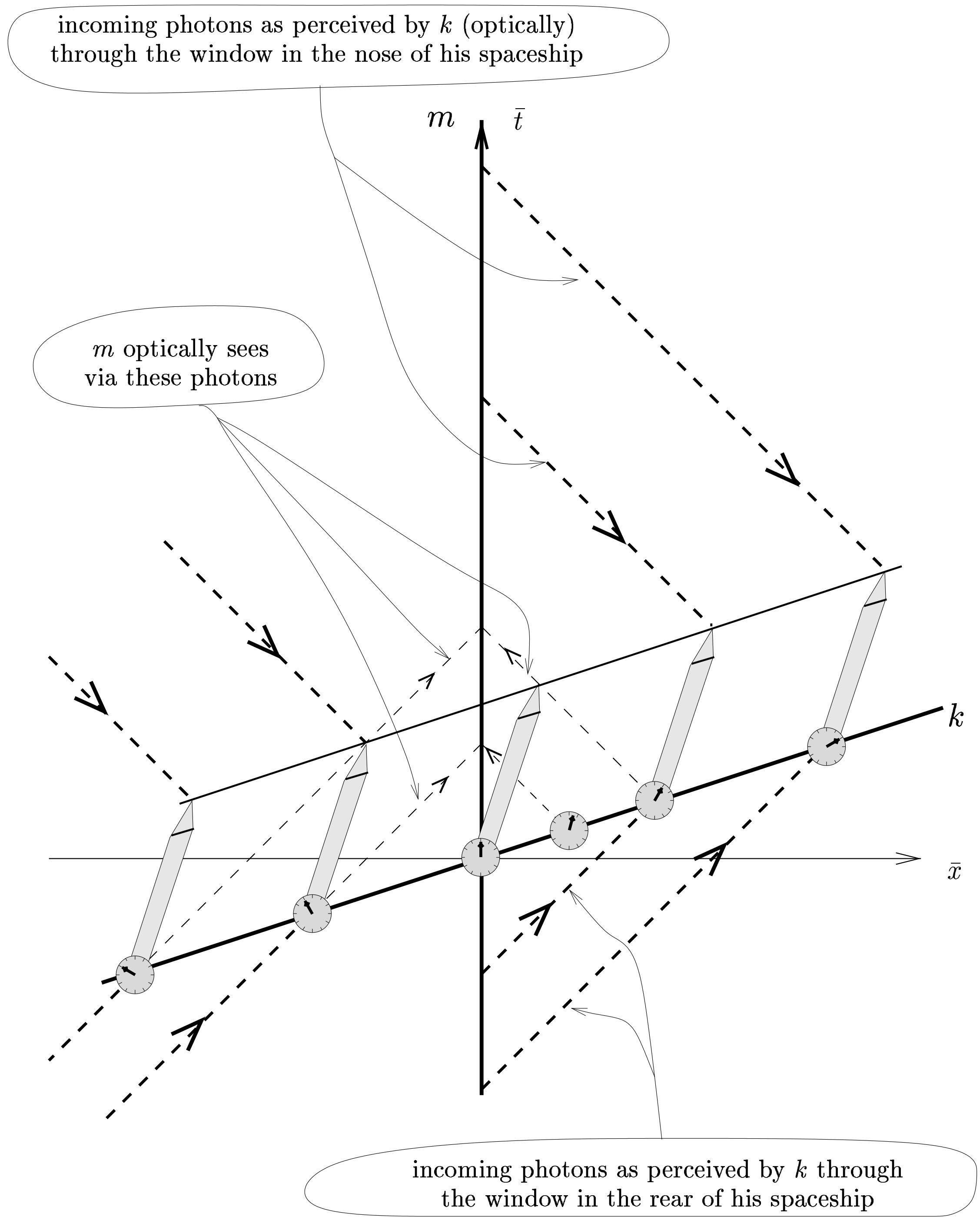


[136] Actually FTL bodies with inner clocks have been extensievly discussed at the seminars which served as a basis for the present lecture notes.

In the next picture, we illustrate how k sees (the rest of the world) *visually*/*optically* and how m sees (the rest of the world) visually. E.g. k interprets photons such that they are "coming" this $\substack{\searrow \\ \nearrow}$ way (relative to the world-view or coordinate-system of m), while m will interpret them "moving" this $\nearrow\nwarrow$ way.

The picture represents a somewhat surprising aspect of FTL visual effects. Namely m sees (optically) k as we already described in connection with the previous picture. Let us turn to the visual (optical) effects experienced by k as coordinatized by m: The photons which k *will interpret as entering* k's spaceship through the front window, "carry information" from m's future according to the world-view (or coordinatization) of m. If k looks out through the rear window then he will see m growing younger and younger. Of course through the front window k will see the "other copy" of m getting older.

Further there are life-lines of photons in this picture the "causal direction" of which is the opposite for k as for m. For k it is this $\searrow$ while for m is this $\nwarrow$. (It is no coincidence that in the present work we did not talk about "causality". We think that a truly *logical* theory of causality is not evident how to create, cf. Russel [229]. As we said, in a future version of this work there will be a chapter about bodies having an inner clock. After that chapter we will be in a better position for reasoning about causality.)


incoming photons as perceived by k (optically) through the window in the nose of his spaceship
m
$\bar{t}$
m optically sees via these photons
k
$\bar{x}$
incoming photons as perceived by k through the window in the rear of his spaceship

2.8 Some symmetry axioms and the twin paradox

In order to discuss the twin paradox, we will need some kind of symmetry axioms.[137] In this section we study the possibility of adding certain symmetry axioms to **Basax**. An example for a symmetry axiom is **Ax(symm)** to be introduced soon. We consider **Ax(symm)** as a possible formalization of (an instance or a fragment of) Einstein's Special Principle of Relativity (SPR), cf. Friedman [90, p.149] principle (R) therein. After introducing **Ax(symm)** and investigating its effect on the paradigmatic effects, we discuss the twin paradox.

Axiom **Ax(symm)** below is an "optional" postulate; sometimes we add it to **Basax** (or other theories of special relativity introduced later in this study) and sometimes we do not. Its usage is somewhat analogous with the Axiom of Choice (AC) in set theory, where people are interested *both* in set theory without AC and also with AC. Moreover, **Ax(symm)** is of a *different nature* than the other axioms introduced up to this point. It expresses a sort of methodological (or aesthetics-motivated) principle: by making all observers similar (in a certain sense) we commit ourselves for describing the world as simply as possible. In this respect **Ax(symm)** will serve as a kind of "Occam's razor" in our analysis. To distinguish aesthetics-motivated axioms like our symmetry principles (e.g. **Ax(symm)**) from experiments-motivated ones (like e.g. **AxE**), statements like our **Ax(symm)** are often called *principles of parsimony* (i.e. principles of economy of explanation in conformity with Occam's razor), cf. e.g. Friedman [90, p.29 line 23]. So, what we call in the present work symmetry axioms[138] all belong to the kind of axioms called principles of parsimony. For more on the special nature of **Ax(symm)** in connection with Occam's razor etc. we refer to § 2.8.3 on page 138 and to item 4.2.18 (p.464).

We provide deeper discussions of symmetry principles in §§ 3.9, 4.2, 4.7.

In the present section we include a relatively brief discussion of **Ax(symm)** and its effects on theorems (or phenomena) already studied in the preceding parts. E.g. we will discuss how **Ax(symm)** influences paradigmatic effects (I)–(III) discussed in §2.5. We will see that these effects (for example the effect of clocks slowing down) admit a simpler and stronger formulation in **Basax** + **Ax(symm)** than in pure **Basax**, and all three paradigmatic effects are necessary if we assume **Ax(symm)**. After this, and motivated by these theorems, we introduce some other axioms and show that they are equivalent with **Ax(symm)**. We then briefly investigate what

[137] This is so because we will approximate the accelerated twin by several inertial observers, and thus we need a kind of "similar behaviour" of these inertial observers.

[138] Cf. besides the present section §§ 3.9, 4.2, 4.7.

Ax(symm) says about the physical world. After this we show that, (in the presence of **Basax** + **Ax**($\sqrt{\ }$)), **Ax(symm)** implies the twin paradox, more precisely, an approximated version of the twin paradox. We then investigate the twin paradox a little. We conclude this section with introducing one of our central axiom systems, **Specrel**.

First, we postulate a natural symmetry principle **Ax(symm₀)**, and then an auxiliary axiom **Ax(eqtime)**. **Ax(symm)** will be defined to be
Ax(symm₀) + **Ax(eqtime)**.

Ax(symm₀) $(\forall m, k \in Obs)(\exists m', k' \in Obs)$
$$\Big(tr_m(m') = tr_k(k') = \bar{t} \quad \wedge \quad \mathsf{f}_{mk} = \mathsf{f}_{k'm'} \Big).$$

Let us see what **Ax(symm₀)** says intuitively, and why we claim that **Ax(symm₀)** is a natural symmetry postulate about "how the world behaves". Assume m, k are two observers. We would like to state that observers m and k are equivalent in some sense. A natural thing to say in this direction would be saying that "as I see you so do you see me". That is

($\star$) as m sees k so does k see m.

But formally this would mean saying that $\mathsf{f}_{mk} = \mathsf{f}_{km}$ which is a too strong statement, e.g. because k may be "looking in the wrong direction". If the bicyclist sees the train moving forwards, the train inhabitants may see the bicycle moving backwards. Cf. the next sequence of pictures. However, this can be easily mended; instead of ($\star$) we state the following more subtle version ($\star\star$).

($\star\star$) As m sees k so does some sister k' of k see some brother m' of m.

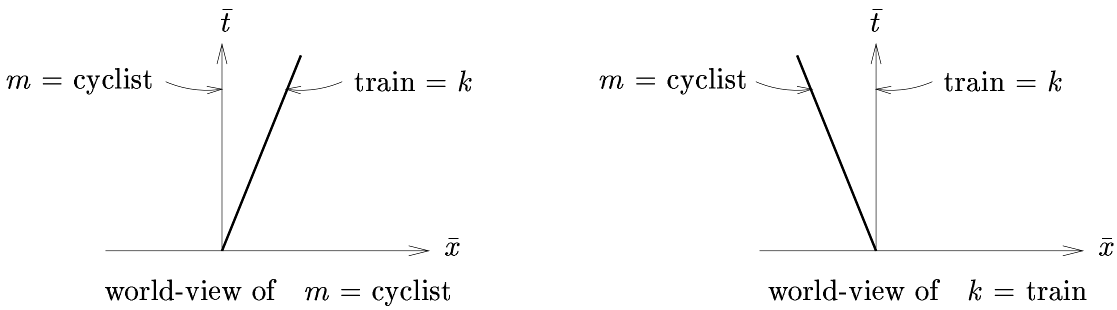
Here saying that k' is a sister of k means that $tr_k(k') = \bar{t}$, i.e. they have the same life-line.[139] Indeed it is exactly the formalized version of ($\star\star$) what is stated in axiom **Ax(symm₀)**.

Perhaps a more natural form of **Ax(symm₀)** would state the existence of brothers m' and k' such that m' and k' see each other exactly the same way, i.e. $\mathsf{f}_{m'k'} = \mathsf{f}_{k'm'}$. If there are no FTL observers (e.g. if $n > 2$), then this is an equivalent

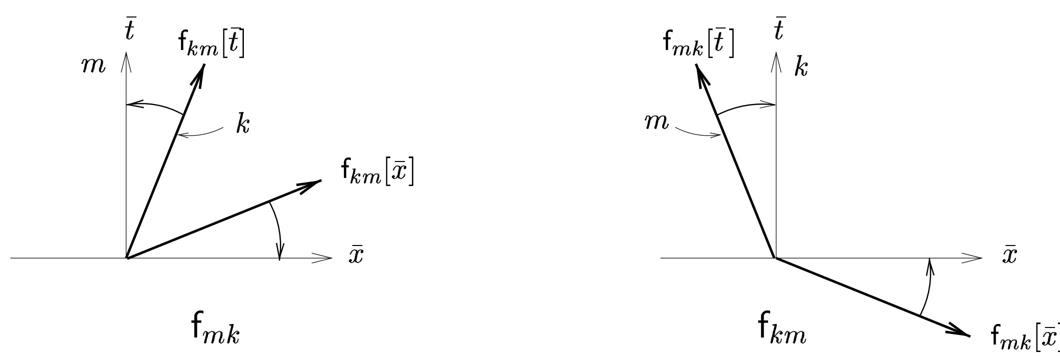
[139] By quantifying over observers having the same life-line (like our quantifiers $\exists k'$, $\exists m'$) we sort of abstracted from "the directions in which our observers are looking" and this is exactly what we needed.

form of **Ax(symm₀)**, see § 3.9. However, this more natural form excludes FTL observers (see Theorem 2.7.5), this is why se stated **Ax(symm$_0$)** in its present form. Cf. Theorem 2.8.2.

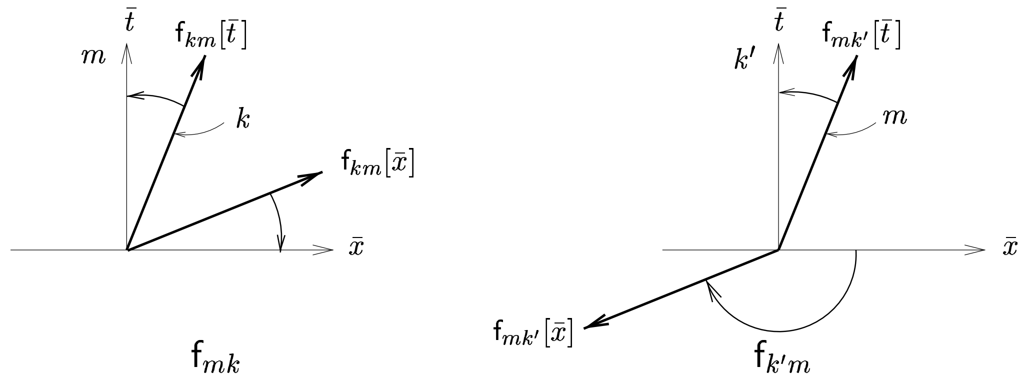
As an illustration for why k' and m' are needed in **Ax(symm$_0$)** we include the following sequence of pictures. (Throughout the discussion of these pictures we assume **Basax**(2).)



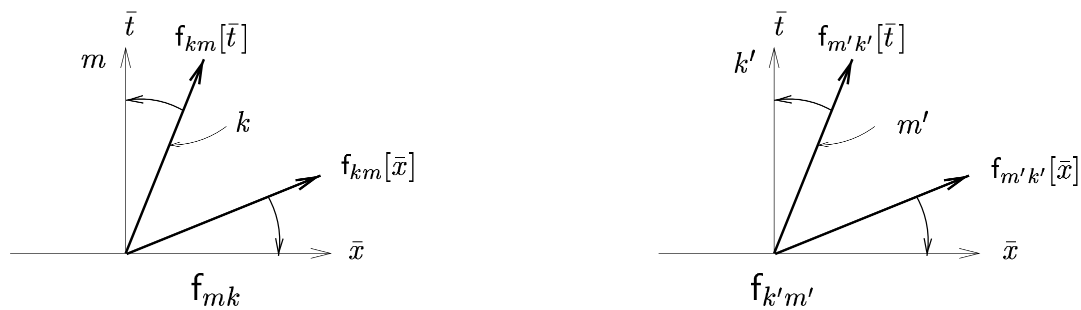
The above represents one possible configuration of the cyclist and the train: The cyclist sees the train moving forwards in the positive $\bar{x}$-direction, while the train people see the cyclist moving backwards in the negative $\bar{x}$-direction. For this configuration the world-view transformations are represented in the following picture.



The above are the world-view transformations corresponding to the configuration represented in the previous picture (the train and the cyclist look in the same direction). Obviously $\mathsf{f}_{mk} \neq \mathsf{f}_{km}$. Let us try to mend this by turning the train around such that the train people will be looking backwards, which in our formalism means that we choose a sister k' of k as illustrated in the picture below.



Well, $\mathsf{f}_{mk} = \mathsf{f}_{k'm}$ is still not satisfied (but we made a step forward, look at $\mathsf{f}_{km}[\bar t\,]$ and $\mathsf{f}_{mk'}[\bar t\,]$). Let us turn the cyclist around too. This means that we take a brother m' of m as illustrated in the picture below.



This picture shows that now we have a chance for $\mathsf{f}_{mk} = \mathsf{f}_{k'm'}$ being true since the lines which are mapped to the axes $\bar t$ and $\bar x$ are mapped to the right places. (In connection with this we note that $v_m(k) = v_{k'}(m')$ follows from Thm.2.8.6 on p.129 which says that under some mild assumptions, $v_m(k) = v_k(m)$. This implies that in our present case $tr_m(k) = tr_{k'}(m')$.) But in addition to this, we need that f_{mk} and $\mathsf{f}_{k'm'}$ agree on these lines, and not just that they take these lines to the same sets. The fact that it is possible to <u>*arrange this*</u> at least in <u>*some*</u> model of **Basax** will be seen in Theorems 2.8.1-2.8.2 below (and in more detail in §3.8.2).

Next, we introduce an auxiliary axiom **Ax(eqtime)**. We call **Ax(eqtime)** *axiom of "equi-time"* because it says that time passes with the same rate for "observer brothers" m and m'.

Ax(eqtime) $(\forall m, m' \in Obs)$
$$\Big(tr_m(m') = \bar{t} \quad \Rightarrow \quad (\forall p, q \in \bar{t}\,)\ |p - q| = |\mathsf{f}_{mm'}(p) - \mathsf{f}_{mm'}(q)|\Big).$$

Concerning **Ax(eqtime)** we note that this is a very natural and convincing axiom, it only says that if two observers *do not move* relative to each other (moreover they are at the same place) then their clocks have the same rate. In other words this means that our paradigmatic effect (I) [140] does *not* show up in the absence of motion. (This is a natural assumption which has always been assumed beginning with ancient Greeks, then by Galileo and Newton and of course by Einstein.)

Let us turn to defining **Ax(symm)**.

$$\mathbf{Ax(symm)} \stackrel{\text{def}}{:=} \mathbf{Ax(symm_0)} + \mathbf{Ax(eqtime)}\,.$$

Let us see first if studying **Basax+Ax(symm)** makes sense at all. We already stated on p.77 that **Basax** is consistent, cf. also §3.5 ("Simple models for **Basax**").

THEOREM 2.8.1 **Basax**(n) + **Ax(symm)** *is consistent, for all* $n > 1$.

On the proof: In the present section we discuss the proof only for $n = 2$. The general case will be proved in §3.8.2. A "computational" proof is given in the proof of Thm.2.8.2 below. As a contrast, the proof in §3.8.2 is a more intuitive, "structuralist" proof. We note that the "standard Minkowskian model" [141] of special relativity validates **Basax** + **Ax(symm)**. ∎

THEOREM 2.8.2

(i) (**Basax**(2) + **Ax(symm)** + **"∃ FTL observers"**) *is consistent i.e.*

(ii) *there is* $\mathfrak{M} \in \mathsf{Mod}(\mathbf{Basax}(2) + \mathbf{Ax(symm)})$ *such that in* $\mathfrak{M}$ *there are FTL observers.*

[140] Moving clocks slow down.
[141] Cf. Def.3.8.42 on p.331.

Proof: We use the construction in §2.4 (pp.80–87). Recall that P was an arbitrary choice function which to each $\ell \in \mathsf{Eucl}(2, \mathfrak{R})$ associated two distinct points o_ℓ, t_ℓ lying on ℓ. Then the model $\mathfrak{M}_1^P = \langle (B;\ Obs, Ph, Ib), \mathfrak{R}, \mathsf{Eucl}(2, \mathfrak{R});\ \in, W\rangle$ was constructed from this P. In $\mathfrak{M}_1^P$ there were FTL observers. Let us modify this construction such that for each observer m we include all the "brothers and sisters" of m as new observers. I.e. let $(\mathfrak{M}_1^P)^+ := \langle (B^+;\ Ph^+, Obs^+, Ib^+), \mathfrak{R}, \mathsf{Eucl}(2, \mathfrak{R});\ \in, W^+\rangle$ be defined as follows:

$Obs^+ :\stackrel{\text{def}}{=} Obs \times F \times \{-1, 1\} \times \{-1, 1\}$,

$Ph^+ :\stackrel{\text{def}}{=} Ph \times F \times \{-1, 1\} \times \{-1, 1\}$.[142]

$B^+ :\stackrel{\text{def}}{=} Ib^+ := Obs^+ \cup Ph^+$,

$W^+ :\stackrel{\text{def}}{=} \Big\{ \Big\langle \langle m, t, i, j\rangle, p, \langle b, t', i', j'\rangle \Big\rangle \in Obs^+ \times {}^2\mathrm{R} \times B^+ \ :\ W(m, \langle p_0 + it, jp_1\rangle, b) \Big\}$.

By the above, $(\mathfrak{M}_1^P)^+$ is defined. We claim that $(\mathfrak{M}_1^P)^+ \models$ **Basax** and in $(\mathfrak{M}_1^P)^+$ there are FTL observers. (Checking these are left to the reader.) Therefore here it is enough to show that P can be chosen such that **Ax(symm)** *too* will be valid in $(\mathfrak{M}_1^P)^+$. To this end (for each $\ell \in \mathsf{Eucl}$) let us choose $o_\ell = \langle o_0, o_1\rangle$ and $t_\ell = \langle t_0, t_1\rangle$ such that the "Minkowski-distance" between them is 1; that is we choose them such that $|(o_0 - t_0)^2 - (o_1 - t_1)^2| = 1$. Checking that this works (e.g. $(\mathfrak{M}_1^P)^+ \models$ **Ax(symm)**) is left to the reader. ∎

The next theorem states that in models of $\mathbf{Basax} + \mathbf{Ax(symm)} + \mathbf{Ax}(\sqrt{\ })$, no field-automorphisms are involved in the world-view transformations.

THEOREM 2.8.3 *Assume* $\mathfrak{M} \models (\mathbf{Basax} + \mathbf{Ax(symm)} + \mathbf{Ax}(\sqrt{\ }))$. *Let* $m, k \in Obs$. *Assume* $\mathsf{f}_{mk}(\bar{0}) = \bar{0}$. *Then* f_{mk} *is a bijective linear transformation of the vector space* ${}^n\mathbf{F}$ *preserving the set of photon-lines.*

Proof: The proof will be filled in later, but it can be easily reconstructed from the proof of Prop.3.8.35 (p.317). ∎

For completeness we note that **Basax** + **Ax(symm)** implies that on Figure 29 the Minkowski-sphere can look like *only* as in the case of $\mathfrak{M}_3$. Cf. the proof of Theorem 2.9.5 on p.155. Further we note that $\mathbf{Basax} + \mathbf{Ax(symm_0)}$ is not sufficient for this. (This was noted by Gergely Székely and Ramon Horváth.)

It is interesting to compare the above theorem about f_{mk}'s with Thm.2.3.12(ii)'. Namely in the above theorem we did not need mentioning field automorphisms

[142] We included all "brothers" of photons only for technical reason.

while we did need them in Thm.2.3.12. There is a similar contrast between the above theorem and Thm.3.1.4 on p.162. In connection with the above theorem we note that, under assuming **Basax** + **Ax(symm)** + **Ax**($\sqrt{\ }$), the world-view transformations f_{mk} are so called Poincaré transformations; and those world-view transformations which preserve $\bar{0}$ are Lorentz transformations, cf. Thm.2.9.5 on p.155.

Remark 2.8.4 Consider the possible models $\mathfrak{M}_1, \mathfrak{M}_2, \mathfrak{M}_3$ of **Basax** represented in Figure 29 (p.88). As we said, of these only $\mathfrak{M}_3$ is a model of **Ax(symm)**. In particular **Ax(symm)** fails both in $\mathfrak{M}_1$ and $\mathfrak{M}_2$. Therefore **Basax** + **Ax(symm)** has radically fewer kinds of models than **Basax** does. The proof is left to the reader. ◁

We think that the answer to the following question should be relatively easy.

Question for future research 2.8.5 Assume $\mathfrak{M} \models$ **Basax** + **Ax(symm)**. Let the set $Wtm^{\mathfrak{M}}$ of the world-view transformations be as in Remark 2.3.9. Is then $\langle Wtm^{\mathfrak{M}}, \circ, ^{-1}, \mathrm{Id}\rangle$ a group? ◁

So far we investigated a symmetry property of the the kind "as I see you so do you see me". According to the following theorem, a simple property of this kind follows already from **Basax** and some mild extra assumption. Compare Theorem 2.7.3 on p.111.

THEOREM 2.8.6 *Assume $\mathfrak{M} \models$ **Basax**. Let $m, k \in Obs$. Then (i)-(iii) below hold.*

(i) *Assume that the field reduct **F** of $\mathfrak{M}$ has no nontrivial automorphisms. Then*

$$v_m(k) = v_k(m)\,.$$

(ii) *Assume that f_{mk} is an affine transformation of ${}^n\mathbf{F}$ (cf. Def.2.9.1 on p.152 for the definition of affine transformations). Then*

$$v_m(k) = v_k(m)\,.$$

(iii) *Assume that $\mathfrak{M} \models$ **Ax(symm)** + **Ax**($\sqrt{\ }$). Then*

$$v_m(k) = v_k(m)\,.$$

On the proof: In the present version we include the proof only of (i) and (ii), and only for $n = 2$. To prove (ii), let $\mathfrak{M} \models \mathbf{Basax}(2)$. Let $m, k \in Obs$ such that f_{mk} is an affine transformation. Without loss of generality we may assume that $\bar{0} \in tr_m(k)$ and that $v_m(k) \neq 0$. Throughout the proof the reader is asked to consult Figure 42.

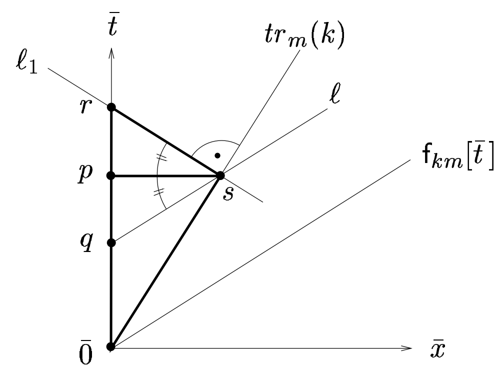


Figure 42: Illustration for the proof of $v_m(k) = v_k(m)$.

Let $s \in tr_m(k)$ be arbitrary such that $s \neq \bar{0}$. Let ℓ be the mirror image of $tr_m(k)$ w.r.t. a photon-line containing s. Let $p := \langle s_t, 0 \rangle$. Let ℓ_1 be the mirror image of ℓ w.r.t. the line $\overline{ps}$ (determined by points p and s). Let $q := \ell \cap \bar{t}$ and $r := \ell_1 \cap \bar{t}$. Note that segments sq and sr have the same length, i.e. $|s - q| = |s - r|$. The following two sentences will not be formalized. Angle $\bar{0}sr$ is a right angle by elementary geometry, so we have that triangles $\bar{0}ps$ and $\bar{0}sr$ are similar. ℓ is parallel with $\mathsf{f}_{km}[\bar{t}]$ by Thm.2.3.12 on p.65. So, we have

$$v_k(m) = \left(\frac{\text{length of segment } sq}{\text{length of segment } s\bar{0}}\right)^2 \quad \left(= \frac{|s-q|^2}{|s|^2}\right) \quad \text{since } \mathsf{f}_{km} \text{ is an affine transformation preserving photon-lines}$$

$$= \left(\frac{\text{length of segment } sr}{\text{length of segment } s\bar{0}}\right)^2 \quad \left(= \frac{|s-r|^2}{|s|^2}\right) \quad \text{since } sq \text{ and } sr \text{ have the same length}$$

$$= \left(\frac{\text{length of segment } ps}{\text{length of segment } p\bar{0}}\right)^2 \quad \left(= \frac{s_x^2}{s_t^2}\right) \quad \text{since triangles } \bar{0}ps \text{ and } \bar{0}sr \text{ are similar}$$

$$= v_m(k).$$

Item (i) follows from item (ii) and Thm.2.3.12(iii) on p.65.

The rest of the proof of this theorem (the case $n > 2$ and also (iii)) will be included in a later version. ∎

2.8.1 Ax(symm) and the paradigmatic effects

Let us turn to seeing how **Ax(symm)** simplifies the "picture" of special relativity, e.g. what it "says" about our paradigmatic effects (I)–(III) (p.90).

THEOREM 2.8.7 (Clocks slow down.)
Assume **Basax** + **Ax(symm)** + **Ax**($\sqrt{\ }$). *Let* $m, k \in Obs$, *with* $0 < v_m(k) \leq 1$. *Then:*

(i) m *thinks that* k*'s clocks run slow, i.e.*

$$|\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > 1\,; \text{ moreover}$$

(ii) $(\forall 0 \neq \lambda \in F)\ |\mathsf{f}_{km}(\lambda \cdot 1_t)_t - \mathsf{f}_{km}(\bar{0})_t| > |\lambda|$.

Proof: Item (i) of this theorem is a corollary of Thm.2.5.2(iii) (p.92) and Thm.2.8.9 below. Item (ii) of Thm.2.8.7 follows from item (i) and Thm.2.8.3 below (p.128). ∎

In connection with the above theorem cf. Theorem 2.5.2. The novelty in Theorem 2.8.7 is that it says that all observers' clocks slow down in the presence of **Ax(symm)**, while without **Ax(symm)** we only know that some clocks slow down. This also means that both m thinks that k's clocks slow down and k thinks that m's clocks slow down. This is counterintuitive to the thinking we got used to in our Newtonian world where if k thinks that m's clocks run slow, then m will think that k's clocks run fast. That both can think that the other's clocks run slow is possible because they do not percieve the same events as simultaneous, i.e. because of paradigmatic effect (III). In connection with this see Figure 50.

Analogous statement can be made about paradigmatic effect (II), i.e. about shrinking of meter-rods, cf. the following theorem. In connection with the next theorem we note the following: If we assume **Basax**, then for every $m, k \in Obs$, such that $v_m(k) \neq \infty$, $tr_m(k)$ can be considered as a function $tr_m(k) : F \longrightarrow {}^{n-1}F$, therefore $tr_m(k)(0)$ is well defined (cf. Fact 2.2.4).

THEOREM 2.8.8 (Meter-rods shrink.)
Assume **Basax** + **Ax(symm)** + **Ax**($\sqrt{\ }$). *Let* $m, k \in Obs$, *with* $0 < v_m(k) \leq 1$. *Then (i) and (ii) below hold.*

(i) *Meter-rods of k parallel with the direction of movement of k shrink when observed by m. I.e. m will think that k's meter-rods are shorter than k thinks, formally:*

For simplicity assume that $\bar{0} \in tr_m(k) \subseteq \mathsf{Plane}(\bar{t}, \bar{x})$. *Let* $k_1 \in Obs$ *with* $\bar{t} \neq tr_k(k_1) \subseteq \mathsf{Plane}(\bar{t}, \bar{x})$ *such that* $tr_k(k_1)$ *is parallel with* $\bar{t}$.[143] *Let* $p := tr_m(k_1)(0)$ *and* $q := tr_k(k_1)(0)$.

Then $|p| < |q|$. *Cf. Figure 38 on p.101.*

(ii) *Those meter rods of k which are not orthogonal to the direction of movement shrink when observed by m. Formally:*

For simplicity assume again that $\bar{0} \in tr_m(k) \subseteq \mathsf{Plane}(\bar{t}, \bar{x})$. *Let* $k_1 \in Obs$ *such that* $tr_k(k_1)$ *is parallel with* $\bar{t}$ *and* $tr_k(k_1)(0)_0 \neq 0$.

Then $|p| < |q|$.

The **proof** will be filled in later. ∎

In connection with the above theorem cf. Items 2.5.9–2.5.12 (pp.100–101).

THEOREM 2.8.9 (Clocks slow down exactly the same way.)
Assume **Basax**+**Ax(symm)**. *Assume* $m, k \in Obs$. *Then m sees k's clocks slowing down to exactly the same degree as k sees m's clocks doing the same; formally:*

(i) $|\mathsf{f}_{km}(1_t)_t - \mathsf{f}_{km}(\bar{0})_t| = |\mathsf{f}_{mk}(1_t)_t - \mathsf{f}_{mk}(\bar{0})_t|$, *moreover:*

(ii) $(\forall p \in \bar{t})\ |\mathsf{f}_{km}(p)_t - \mathsf{f}_{km}(\bar{0})_t| = |\mathsf{f}_{mk}(p)_t - \mathsf{f}_{mk}(\bar{0})_t|$.

Proof: We will fill in the proof later, but it can be easily reconstructed from the proof of Prop.3.8.34 (p.317). ∎

Remark 2.8.10 An analogous statement can be made about the effect of shrinking meter-rods as follows.
Assume **Basax** + **Ax(symm)** + **Ax**($\sqrt{\ }$). Then (i) and (ii) below hold.

(i) Assume further $m, k \in Obs$ are in standard configuration. Let us concentrate on meter-rods *parallel with the direction of movement*. Then m will see k's meter-rods shrink exactly with the same ratio as k sees m's meter-rods shrink.

[143]Intuitively k and k_1 together represent a meter-rod of k.

(ii) Assume k moves in direction $\bar{x}$ when observed by m (i.e. $tr_m(k) \subseteq \mathsf{Plane}(\bar{t},\bar{x})$). Let us concentrate on meter-rods which are parallel with direction $\bar{x}$ when observed by m. I.e. even if the meter-rod is k's one we check whether m sees it parallel with the plane $\mathsf{Plane}(\bar{t},\bar{x})$ determined by axes $\bar{t}$ and $\bar{x}$.

Then m will see k's meter-rods shrinking with the same ratio as k sees m's meter-rods. $\lhd$

Roughly, the following theorem implies that meter-rods orthogonal to the direction of movement do not shrink or grow, assuming **Basax** + **Ax(symm)**, cf. Corollary 2.8.12.

THEOREM 2.8.11 *Assume* **Basax** + **Ax(symm)** + **Ax**($\sqrt{\ }$). *Let* $m, k \in Obs$. *Let* e, e_1 *be two events which are simultaneous for both* m *and* k. *Then the spatial distance between* e *and* e_1 *is the same for* m *as for* k*; formally:*

$$(\forall p, q \in {}^nF)\ [\,(p_t = q_t\ \wedge\ \mathsf{f}_{mk}(p)_t = \mathsf{f}_{mk}(q)_t) \quad \Rightarrow \quad |p-q| = |\mathsf{f}_{mk}(p) - \mathsf{f}_{mk}(q)|\,]\,.$$

The **proof** will be filled in later. ∎

The following is a corollary of Thm.2.8.11 above and Thm.2.5.6 (p.97) which says that (under assuming **Basax**) if two clocks are separated <u>*only*</u> in the spatial direction which is orthogonal to the direction of movement they do not get out of synchronism.

COROLLARY 2.8.12 (Meter-rods orthogonal to movement do not shrink.)
Assume **Basax** + **Ax(symm)** + **Ax**($\sqrt{\ }$). Then meter-rods orthogonal to the direction of movement do <u>*not*</u> get shorter; formally: Let $m, k \in Obs$. Assume m sees that k does not move in direction $\bar{y}$, that is

$$(\forall p, q \in tr_m(k))\ p_y = q_y\,.$$

Then,

$$(\forall p, q \in {}^nF)\ \Big((\forall i \in n)(i \neq 2\ \Rightarrow\ p_i = q_i) \quad \Rightarrow \quad |p-q| = |\mathsf{f}_{mk}(p) - \mathsf{f}_{mk}(q)|\Big)\,.$$

In particular

$$p, q \in \bar{y} \qquad \Rightarrow \qquad |p-q| = |\mathsf{f}_{mk}(p) - \mathsf{f}_{mk}(q)|\,.$$

$\lhd$

2.8.2 Equivalent forms of Ax(symm)

In this part we show that certain forms of the paradigmatic effects are actually equivalent with **Ax(symm)** (in **Basax** under mild conditions). Different equivalent forms of **Ax(symm)** will be given in § 3.9.

Thm.2.8.9 motivates the axiom **Ax(syt$_0$)** below. We note that the "name" **Ax(syt$_0$)** intends to refer to "symmetry of time". Intuitively, **Ax(syt$_0$)** says that

"as I see your clocks slowing down (because of your speed relative to me) so do you see my clocks (because of my speed relative to you) slowing down".

In the formulation of **Ax(syt$_0$)** below the assumption $tr_m(k) \neq \emptyset$ is superfluous at the present point, because **Basax** $\models tr_m(k) \neq \emptyset$. However in later sections this assumption will become useful.[144]

Ax(syt$_0$) $$(\forall m, k \in Obs)\Big(tr_m(k) \neq \emptyset \quad \Rightarrow \\ (\forall p \in \bar{t}\,)\ |\mathsf{f}_{mk}(p)_t - \mathsf{f}_{mk}(\bar{0})_t| = |\mathsf{f}_{km}(p)_t - \mathsf{f}_{km}(\bar{0})_t|\Big).$$

In terms of the just defined **Ax(syt$_0$)**, Thm.2.8.9 says that

$$\textbf{Basax} \models \textbf{Ax(symm)} \rightarrow \textbf{Ax(syt}_\mathbf{0}\textbf{)}.$$

In Thm.2.8.13 below we will see that, under mild assumptions, the implication holds in the other direction too, i.e. **Ax(syt$_0$)** is an equivalent form of **Ax(symm)** (in th presence of **Basax**). To formulate Thm.2.8.13 we introduce auxiliary axioms **Ax(*Triv*)** and **Ax(*Triv$_t$*)**. First we introduce the notion of an isometry and the set *Triv* of trivial transformations.

Triv denotes the set of all mappings of ^nF into intself which preserve Euclidean distance, take $\bar{t}$ to a line parallel with it, and so that the order of points does not change on $\bar{t}$. Formally: Let $\mathfrak{F}$ be an ordered field. Then $f : {}^nF \longrightarrow {}^nF$ is said to be an *isometry* iff it preserves the square of Euclidean distances, i.e. $(\forall p, q \in {}^nF)$

$$(p_0 - q_0)^2 + (p_1 - q_1)^2 + \ldots + (p_{n-1} - q_{n-1})^2 = \\ (f(p)_0 - f(q)_0)^2 + (f(p)_1 - f(q)_1)^2 + \ldots + (f(p)_{n-1} - f(q)_{n-1})^2,$$

[144]We would like to remind the reader that we mentioned that when generalizing our axioms toward general relativity, **Ax6** and **Ax3** will be weakened and therefore $tr_m(k) = \emptyset$ will be possible for some choices of $m, k \in Obs$.

cf. also Def.3.9.3 on p.349. Let ${}^{({}^nF)}\,{}^nF$ denote the set of all functions mapping ${}^nF$ into itself. Then

$$Triv \stackrel{\text{def}}{=} Triv(n,\mathfrak{F}) \stackrel{\text{def}}{=} \{\, f \in {}^{({}^nF)}\,{}^nF \ :\ f \text{ is an isometry, } f[\bar{t}] \parallel \bar{t},\ f(1_t)_t - f(\bar{0})_t > 0 \,\}\,.$$

As we will explain in §3.5 in more detail, the transformations in *Triv* involve *no "relativistic effects"*, one could say that they are very non-relativistic or, so to speak, trivial. To illustrate this, assume $f(\bar{0}) = \bar{0}$. Then $f \in Triv$ if and only if f is identity on $\bar{t}$ and f maps the space-part S to itself (i.e. $f[S] = S$) so that it preserves Euclidean distance on S.

Ax(*Triv*) $\quad (\forall m \in Obs)(\forall f \in Triv)(\exists k \in Obs)\ \mathsf{f}_{mk} = f.$

> **Ax(*Triv*)** says that every observer can "re-coordinatize" his world-view by any trivial transformation. As is explained in §3.9, **Ax(*Triv*)** is first-order, because each isometry is an affine transformation, and so quantifying over elements of *Triv* can be replaced with quantifying over elements of $\mathfrak{F}$.
>
> **Ax(*Triv$_t$*)** below is a weaker form of **Ax(*Triv*)**.

Ax(*Triv$_t$*) $\quad (\forall m \in Obs)(\forall f \in Triv)\ \Big(f[\bar{t}] = \bar{t} \quad \Rightarrow \quad (\exists k \in Obs)\ \mathsf{f}_{mk} = f\Big).$

THEOREM 2.8.13 *Assume $n > 2$. Then*

$$\mathbf{Basax} + \mathbf{Ax}(\surd\,) + \mathbf{Ax}(Triv_t) \ \models\ \mathbf{Ax(symm)} \leftrightarrow \mathbf{Ax(syt_0)}.$$

Proof: This will be proved as Prop.3.9.47(ii) (p.391), cf. also [174]. ∎

We consider **Ax(*Triv*)** and **Ax(*Triv$_t$*)** as some of our auxiliary axioms.[145] A similar auxiliary axiom is **Ax(||)** to be introduced below.

---

[145] The axioms we call auxiliary are of a status that we assume them without any hesitation whenever we need them. I.e. we consider them as true in the "real world" and we omit them from some of our theories only to make these theories look prettier. To be on the safe side, we note that **Ax(*Triv*)** and **Ax(||)** will "not survive" the transition from special to general relativity. They both will need some refining already in our chapter on accelerated observers. The following form **Ax**(*Triv$_{t0}$*) of **Ax(*Triv$_t$*)** will remain "true": $(\forall f \in Triv)[f(\bar{0}) = \bar{0} \Rightarrow (\exists k \in Obs)\mathsf{f}_{mk} = f]$. In the case when we will allow only uniformly accelerated observers, also **Ax(*Triv$_t$*)** will remain "true". Here "true" means "usable" or consistent with our intentions.

Ax(||) $\quad (\forall m, k \in Obs)\Big(tr_m(k) \parallel \bar{t} \quad \Rightarrow \quad (\mathsf{f}_{mk} \text{ is an isometry})\Big).$

Intuitively, assuming **Ax4**, axiom **Ax(||)** says that if you do not move relative to me then we will agree on which events are simultaneous, which occurred at the same place and we agree on both spatial distances and temporal distances between events. Hence **Ax(||)+Ax4** implies that none of the paradigmatic effects shows up in the absence of motion. **Ax(||)** is a stronger version of **Ax(eqtime)**. In passing we note that later (in §6) we will introduce an axiom called **Ax(eqm)** which (under mild assumptions) will be a stronger version of **Ax(||)**.

The proposition below says that, assuming **Basax** + **Ax(*Triv*)**, the auxiliary axioms **Ax(||)** and **Ax(eqtime)** are equivalent.

PROPOSITION 2.8.14 $\quad$ **Basax** + **Ax(*Triv*)** $\models$ **Ax(||)** $\leftrightarrow$ **Ax(eqtime)**.

The **proof** will be filled in later. ∎

The following proposition says that, assuming **Basax**+**Ax($\sqrt{\ }$)**, both **Ax(syt$_0$)** and **Ax(symm)** imply **Ax(||)**.

PROPOSITION 2.8.15

(i) **Basax** + **Ax($\sqrt{\ }$)** + **Ax(syt$_0$)** $\models$ **Ax(||)**.

(ii) **Basax** + **Ax($\sqrt{\ }$)** + **Ax(symm)** $\models$ **Ax(||)**.

Proof: Item (i) will be proved as Prop.3.9.44 (p.389). Item (ii) follows from Thm.2.8.9 and item (i). ∎

Since **Ax(||)** is a stronger form of **Ax(eqtime)** the above proposition implies that

$$\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ }) + \mathbf{Ax(syt_0)} \models \mathbf{Ax(eqtime)}.$$

Thm.2.8.11 motivates the following potential axiom, which we call the *axiom of "equi-space"*.

Ax(eqspace) $\quad (\forall m, k \in Obs)(\forall p, q \in {}^nF)$

$$\Big((p_t = q_t \ \wedge\ \mathsf{f}_{mk}(p)_t = \mathsf{f}_{mk}(q)_t) \quad \Rightarrow \quad |p - q| = |\mathsf{f}_{mk}(p) - \mathsf{f}_{mk}(q)| \Big).$$

Intuitively, **Ax(eqspace)** says that if two events are simultaneous both for m and k, then the spatial distance between those two events is the same for m as for k. Theorem 2.8.16 below explains why we consider **Ax(eqspace)** as one of our symmetry axioms.

In terms of the just defined **Ax(eqspace)**, Thm.2.8.11 says that

$$(\mathbf{Basax} + \mathbf{Ax(symm)} + \mathbf{Ax}(\sqrt{})) \models \mathbf{Ax(eqspace)}\,.$$

In the next theorem we will see that (under assuming $n > 2$ and $\mathbf{Ax}(\mathit{Triv}_t)$) the implication holds in the other direction, too.

THEOREM 2.8.16 *Assume $n > 2$. Then*

$$(\mathbf{Basax} + \mathbf{Ax}(\sqrt{}) + \mathbf{Ax}(\mathit{Triv}_t)) \models \mathbf{Ax(symm)} \leftrightarrow \mathbf{Ax(eqspace)}\,.$$

The **proof** will be filled in later, and it can be found in [174]. ∎

Next we introduce another natural symmetry axiom. Intuitively, **Ax(speedtime)** below says that the rate with which moving clocks slow down depends only on the relative velocity $\vec{v}_m(k)$ with which one observer sees the other moving. Roughly, the idea is the following. The relativistic effects are caused by relative motion (of k relative to m). Motion is completely determined[146] by velocity $\vec{v}_m(k)$ (of k relative to m). Therefore one concludes that relativistic effects (involving f_{mk}) should be determined by $\vec{v}_m(k)$. (At least if we disregard acceleration). For technical reasons the axiom is formulated in terms of speeds instead of velocities. Interestingly, we will see that this axiom turns out to be one of the symmetry axioms analogous with **Ax(symm)** and $\mathbf{Ax(syt_0)}$, cf. Thm.2.8.17.

Ax(speedtime) $(\forall m, k, m', k' \in \mathit{Obs}) \Big(v_m(k) = v_{m'}(k') \quad \Rightarrow$

$$(\forall p \in \bar{t}\,)\, |\mathsf{f}_{mk}(p)_t - \mathsf{f}_{mk}(\bar{0})_t| = |\mathsf{f}_{m'k'}(p)_t - \mathsf{f}_{m'k'}(\bar{0})_t| \Big).$$

Ax(speedtime) also turns out to be equivalent with an instance of (or fragment of) Einstein's SPR.

The theorem below says that (under mild assumptions) the symmetry axioms introduced in this section are equivalent with each other. For similar equivalence theorems we refer the reader to §§ 3.9, 4.7.

[146] If we disregard acceleration and things like that.

THEOREM 2.8.17 *Assume $n > 2$. Then (i) and (ii) below hold.*

(i) $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ }) + \mathbf{Ax}(\mathit{Triv}_t) \models$

$$\mathbf{Ax(symm)} \leftrightarrow \mathbf{Ax(speedtime)} \leftrightarrow \mathbf{Ax(syt_0)} \leftrightarrow \mathbf{Ax(eqspace)};$$

where the "transitive notation" $\psi_1 \leftrightarrow \psi_2 \leftrightarrow \psi_3$ intends to abbreviate $(\psi_1 \leftrightarrow \psi_2)$ & $(\psi_2 \leftrightarrow \psi_3)$. Similarly for the case when we have four formulas say $\psi_1, \ldots, \psi_4$.

(ii) $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ }) \models \mathbf{Ax(speedtime)} \leftrightarrow \mathbf{Ax(syt_0)} \leftrightarrow \mathbf{Ax(eqspace)}$.

The **proof** will be filled in later, but cf. §3.9 and [174]. ∎

2.8.3 Is Ax(symm) objective or subjective?

Instead of **Ax(symm)** let us discuss its corollary $\mathbf{Ax(syt_{00})}$ formulated below, because this simplifies the discussion. However, the whole discussion extends to **Ax(symm)** too.

$\mathbf{Ax(syt_{00})}$ $\quad (\forall m, k \in \mathit{Obs})\,[\mathsf{f}_{mk}(\bar{0}) = \bar{0} \;\Rightarrow\; |\mathsf{f}_{mk}(1_t)_t| = |\mathsf{f}_{km}(1_t)_t|]$.

That is, "if I see your clocks slowing down (as a consequence of your motion relative to me) so will you see my clocks slowing down (as a consequence of my motion relative to you)".

Meditating over the meaning of $\mathbf{Ax(syt_{00})}$ leads to the following question. $\mathbf{Ax(syt_{00})}$ can be made true (or false) by choosing the *units of measurement k uses.* (The same applies to **Ax(symm)**). But choosing units of measurement is something subjective. Assume that m lives on the Earth while k lives in a space-ship originating from another galaxy. Then they can see each other all right, but how can they compare their meter-rods (or their clocks) i.e. how can they agree on using the same units of measurement. Suppose, they are in radio communication. If they cannot compare their units of measurement via radio communication, then perhaps there is no thought-experiment for them to check wether $\mathbf{Ax(syt_{00})}$ is true, which could render this axiom either meaningless or to be a matter of agreement for convenience. In other words, then $\mathbf{Ax(syt_{00})}$ would become kind of subjective (i.e. something which does not say too much about what the world is really like, but instead which is about how we choose to describe the world).

The following intuitive argument says that this danger is not present i.e. that $\mathbf{Ax(syt_{00})}$ and **Ax(symm)** are objective, i.e. they are *checkable* by some *thought experiment*. This goes as follows. Electrons, and hydrogen atoms are the same in all parts of our Universe, according to the best of our knowledge.[147] Observers m and k can agree through their radio contact that they will use the hydrogen atom for defining their units of measurement (both for space and for time). I.e. they can agree to use the *same* units of measurement. After this, it is only a matter of patience to work out a thought experiment for checking whether $\mathbf{Ax(syt_{00})}$ holds for m and k. A similar argument applies to $\mathbf{Ax(symm)}$ in place of $\mathbf{Ax(syt_{00})}$. Therefore, we can conclude that $\mathbf{Ax(syt_{00})}$ and $\mathbf{Ax(symm)}$ are meaningful (objective) axioms about what the world is like (and not only "linguistic toys" like, say, absolute time).

The above considerations (using hydrogen atoms for matching units of measurement) will come up in §4 where we will look into axiom systems weaker than **Basax** and the question will come up whether the difference between the weak system[148] and **Basax** is testable by thought experiments (i.e. is objective) or not.

2.8.4 The twin paradox

The twin paradox $\mathbf{(TwP)}$ was formulated on p.38. However, that formulation cannot be used in **Basax** because it uses non-inertial (i.e. accelerated) observers. Below we will introduce a variant of $\mathbf{(TwP)}$ in which we will simulate an accelerated observer by several inertial ones. To formulate (our present version[149] of) the twin paradox, we will need the binary relation STL of being slower than light between observers to be recalled from Thm.2.7.2. Let $\mathfrak{M}$ be a frame model and $m, k \in Obs^{\mathfrak{M}}$. Then[150]

$$m \; STL \; k \quad \overset{\text{def}}{\Longleftrightarrow} \quad v_m(k) < 1 \, .$$

[147] At this point we acknowledge that we brought a new axiom into our picture of the world. But we think that should be all right as far as one acknowledges it. (I.e. what we are saying about $\mathbf{Ax(syt_{00})}$ is not based on pure logic only.)

[148] in which for different observers the speed of light might be different

[149] In reality this is only an "approximation" of the original paradox, and this is called "clock paradox" in d'Inverno [75]p.24.

[150] The relation STL will be re-defined in Def.4.2.6 (p.460), where the definitions of $\mathbf{Ax(TwP)}$ will be changed too. The reason for this is that in the first 219 pages of this work we assume that the speed of light is 1 (for short, $c = 1$). We relied on this convention in formalizing the definition of STL above. However in later theories (beginning with **Bax**, p.219) our assumptions on the speed of light will become more subtle (than $c = 1$). Hence some of the formal definitions made during the "$c = 1$ era" will need adjustment around the beginning of §4. Therefore the present definitions of STL and $\mathbf{Ax(TwP)}$ live (i.e. they apply) on pp. 139–422. The future definitions of STL and $\mathbf{Ax(TwP)}$ live from p.422 to the end of the present work. However, in the presence of **Basax**, the new and old definitions of STL and $\mathbf{Ax(TwP)}$ will have the same meaning.

In Thm.2.7.2 we saw that STL is an equivalence relation on the set of observers Obs (assuming $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ }))$.

Let us turn to formulating our present version of the twin paradox. Although the formula below might look long at first sight, its intuitive content is simple cf. Figure 43. The key idea is the following. Originally in **(TwP)** we had two twin brothers m and k. Of these, m was inertial while k was accelerated. As we already said, since now (in the present section) we do _not_ have accelerated observers, we will have to _simulate_ (or approximate) brother (i.e. observer) k by two "auxiliary" inertial observers k_1 and k_2.

We will return to discussing the role of STL in **Ax(TwP)**, soon.

Ax(TwP) $(\forall m, k_1, k_2 \in Obs)(\forall p, q, r \in {}^nF)$

$$\Big([\, m\ STL\ k_1 \quad \wedge \quad m\ STL\ k_2 \quad \wedge \quad p_t < q_t < r_t \quad \wedge$$

$$\{p\} = tr_m(m) \cap tr_m(k_1) \ \wedge\ \{q\} = tr_m(k_1) \cap tr_m(k_2) \ \wedge\ \{r\} = tr_m(m) \cap tr_m(k_2)\,] \ \Rightarrow$$

$$|p_t - r_t| > |\mathsf{f}_{mk_1}(p)_t - \mathsf{f}_{mk_1}(q)_t| + |\mathsf{f}_{mk_2}(q)_t - \mathsf{f}_{mk_2}(r)_t| \Big),$$

see Figure 43. [151]

THEOREM 2.8.18 $(\mathbf{Basax} + \mathbf{Ax(symm)} + \mathbf{Ax}(\sqrt{\ })) \models \mathbf{Ax(TwP)}$.

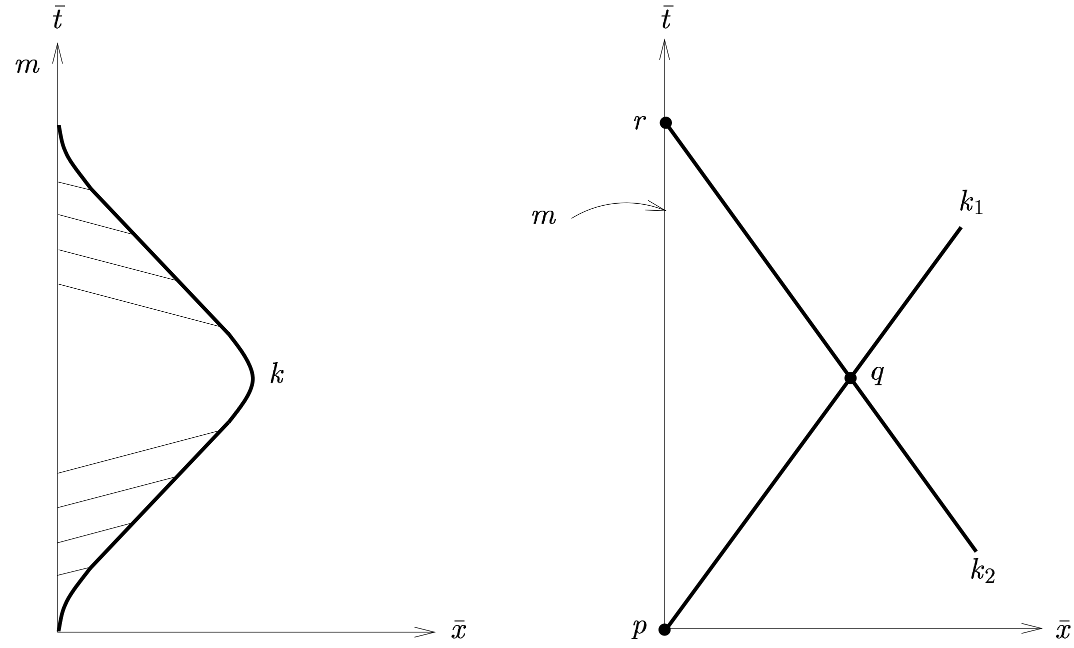
The **proof** will be filled in later. ∎

Remark 2.8.19 (On the role of STL in **Ax(TwP)**)

(i) STL is needed in **Ax(TwP)** (in connection with Thm.2.8.18 above) in the following sense. Assume $\mathfrak{M} \models (\mathbf{Basax} + \mathbf{Ax(symm)} + \mathbf{Ax}(\sqrt{\ }))$ and $m, k \in Obs^{\mathfrak{M}}$ such that $v_m(k) \geq 1$. Then there are $k_1, k_2 \in Obs$ and $p, q, r \in {}^nF$ such that all the assumptions about m, k_1, k_2, p, q, r of **Ax(TwP)** hold with the exception of "STL", and the conclusion of **Ax(TwP)** fails. In other words, if $(\exists m, k)\ v_m(k) \geq 1$ in $\mathfrak{M}$ then that version of **Ax(TwP)** from which the STL condition is deleted fails in $\mathfrak{M}$.

(ii) The condition "STL" can be replaced in **Ax(TwP)** with the following perhaps more natural condition: All three observers m, k_1, k_2 think that event $w_m(q)$ was "temporally between" events $w_m(p)$ and $w_m(r)$. We leave the complete formalization of this version of **Ax(TwP)** to the reader. ◁

[151] As we indicated, the definition of **Ax(TwP)** will be changed on p.460.



This is what we approximate. $|p_t - r_t| > |\mathsf{f}_{mk_1}(p)_t - \mathsf{f}_{mk_1}(q)_t| + |\mathsf{f}_{mk_2}(q)_t - \mathsf{f}_{mk_2}(r)_t|$

Figure 43: Twin paradox. (In this figure we choose the speed of light to be 2 instead of 1 for better representation of the effects we want to illustrate.) The slanted lines in the left-hand picture represent simultaneities of observer k.

Figure 43 shows how the inertial brother, m, observes his accelerated twin brother, k. Let us see how the accelerated brother, k, observes his inertial twin brother, m. Below (and when looking at Figures 43-48) it is important to keep in mind that "m observes k" means that m represents k's life-line in m's coordinate system. Hence "observing" means "coordinatizing" and not visually seeing via photons. Hence "observing" does not involve any visual effect like the doppler effect. As we said before, this convention applies throughout the present work.

In **Ax(TwP)** we approximated k by two inertial observers, k_1 and k_2. We can imagine that k travels with k_1 until k_1 meets k_2, when k "jumps over" to k_2's spaceship. We then put together k's worldview from k_1's and k_2's such that k's worldview agrees with k_1's worldview until they meet k_2, and from that time on k's worldview agrees with k_2's worldview. We also assume that the clocks of k_1 and k_2 are such that they show the same time at their encounter (i.e. we assume that $\mathsf{f}_{mk_1}(q) = \mathsf{f}_{mk_2}(q)$ where $\{q\} = tr_m(k_1) \cap tr_m(k_2)$).

From now on we assume $\mathbf{Basax} + \mathbf{Ax(symm)} + \mathbf{Ax}(\sqrt{\ })$. We will use properties of the world-view transformations in models of $\mathbf{Basax} + \mathbf{Ax(symm)} + \mathbf{Ax}(\sqrt{\ })$ that we proved in this section, see e.g. Theorems 2.8.7-2.8.8.

Figure 44 shows how k observes m when k is approximated by k_1 and k_2 as on Figure 43. Recall that on Figure 43, $\vec{v}_m(k_1) = -\vec{v}_m(k_2)$ and k_1 and k_2 meet at q, i.e. $tr_m(k_1) \cap tr_m(k_2) = \{q\}$, and $p = \bar{0}$. I.e., according to Figure 43, m observes k receding with speed v until time q_t, when k turns back and begins to approach with the same speed v. As illustrated on Figure 44, k will observe m to recede with the *same* speed v until time $\mathsf{f}_{mk_1}(q)_t$, when m turns back (as observed by k) and begins to approach with speed v. This is very similar to how m observes k, except that m, as observed by k, turns back sooner than k does so as observed by m, because by paradigmatic effect (I) (moving clocks slow down) we have that $\mathsf{f}_{mk_1}(q)_t < q_t$. I.e., m needs less time for the journey as k observes it than k needs for the journey as m observes it (this is the twin paradox). This also implies that the *distance* m covered according to k is *less than* the distance k covered according to m.

Let us analyze further (from a different point of view) how k observes m. Assume that when k departs, m is standing there waving goodbye, then goes home, has breakfast, and then comes back to the departing spot again to meet his brother k. Now, k will observe m waving goodbye and starting to go home in slow motion (i.e. all of m's processes are slower than usual), then before m reached home, according to k's worldview, suddenly he is already coming back again (in slow motion) to meet him at the departing spot. In turn, m will observe his twin brother in slow motion all the time, and he will observe all events that happened with k on his journey.

In more technical terminology, using Figures 43 and 44: $e_0 = w_m(\bar{0})$ is the event of k's departing, and $w_m(q)$ is the event of k's turning back on his journey. Let event e_1 in m's life be simultaneous with $w_m(q)$ according to k_1 (i.e. $m \in e_1$ and $\mathsf{f}_{mk_1}(q)_t = w_{k_1}^{-1}(e_1)_t$). See Figure 44. Similarly, let event e_2 in m's life be simultaneous with $w_m(q)$ according to k_2. We can see that e_2 happens much later in m's life than e_1 and that k does not observe the events in m's life that happen between e_1 and e_2. In our story, e_1 is an event in m's life when he is on his way home after waving goodbye to his twin, and e_2 is an event in m's life when he already is on his way back to meet his twin brother upon his return. k observes m in slow motion because k observes that m's clock slows down, and so for k more time passes

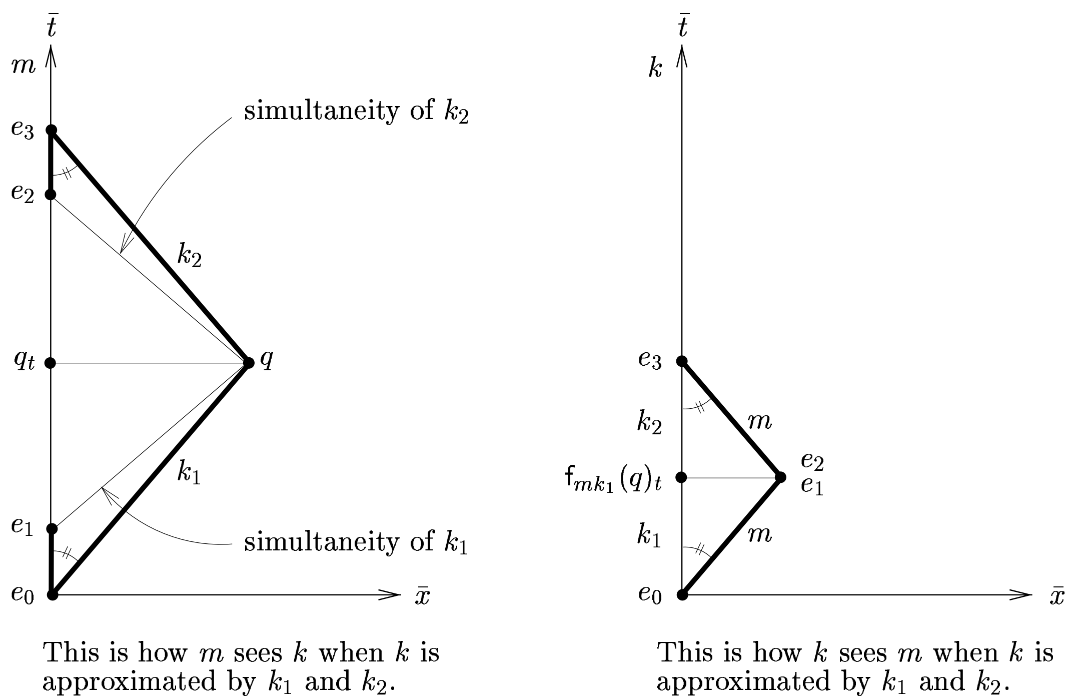


Figure 44: Twin paradox approximated by two inertial observers of the same speed.

between the events e_0 and e_1 than for m.[152]

On the other hand, as we said, m will observe his twin brother in slow motion all the time, and he will observe all events that happened with k on his journey. What is the reason for this strong assymetry between the twins? The reason is that m is an inertial observer while k is not; k's worldview is put together from the worldviews of two different inertial observers, and at the "pasting point" (i.e. at the event when k turns back) there are strange effects, e.g. a large part of m's life-line gets "cut out" (k observes m suddenly at a much later point in m's life). As a side-effect of approximating k by only two inertial observers, "at point q" k experiences infinite acceleration (which in turn naturally causes funny effects). Soon we will approximate k by more and more inertial observers. Then the "irrelevant" parts of the funny effects will gradually fade away while the "relevant" parts of the effects will stay with us (cf. e.g. Figure 47).

If we approximate k by two inertial observers differently than on Figure 43, e.g. if k comes back more slowly than he was travelling outward, then k will observe m at the turning point suddenly placed at a bigger distance, as on Figure 45. But if k_1 and k_2 have the same speed, then this "instantaneous displacement" will not occur.

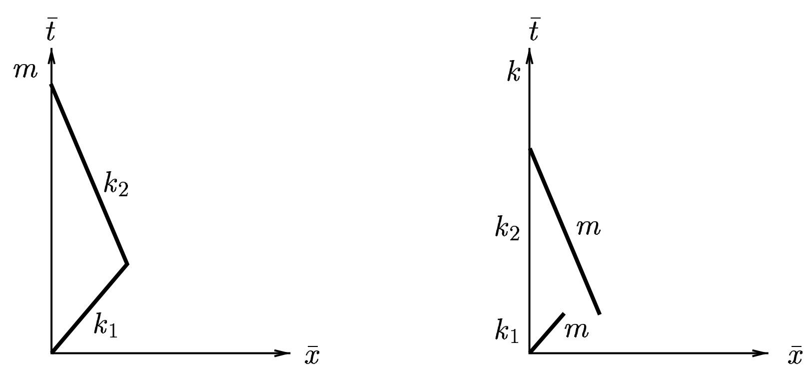


Figure 45: Twin paradox approximated by two inertial observers of different speeds.

Let us see what the above look like if we *refine* our approximation of the accelerated twin, i.e. if we approximate the accelerated twin by more and more inertial observers. From now on we assume that the life-line of k is symmetric in the sense

[152]This is one of our paradigmatic effects, the **Ax(symm)** version of "moving clocks slow down", cf. Thm. 2.8.7.

that his motion outwards is exactly of the same kind as his motion inward, i.e. k's life-line is symmetric w.r.t. the horizontal line containing q. Further, we assume that both m's and k's clocks show 0 at the turning point of k's life-line.[153] Also, for simplicity, we assume $\mathfrak{F} = \mathfrak{R}$.

Figure 46 shows m's and k's worldviews when k is approximated by three inertial observers, and when k is approximated by five inertial observers. We can see that as we approximate k by more and more inertial observers, the intervals on m's life-line that k will not observe become shorter and shorter, and eventually k will observe all events on m's life-line. Similarly, m's life-line will eventually become a continuous curve as k observes it (i.e. the displacements at the "pasting points" will eventually disappear).[154] (The word "eventually" here means "at the limit of this approximating process". We approximate so that the difference of speeds of the consecutive inertial observers approaches 0 and we choose the "pasting points" appropriately.)

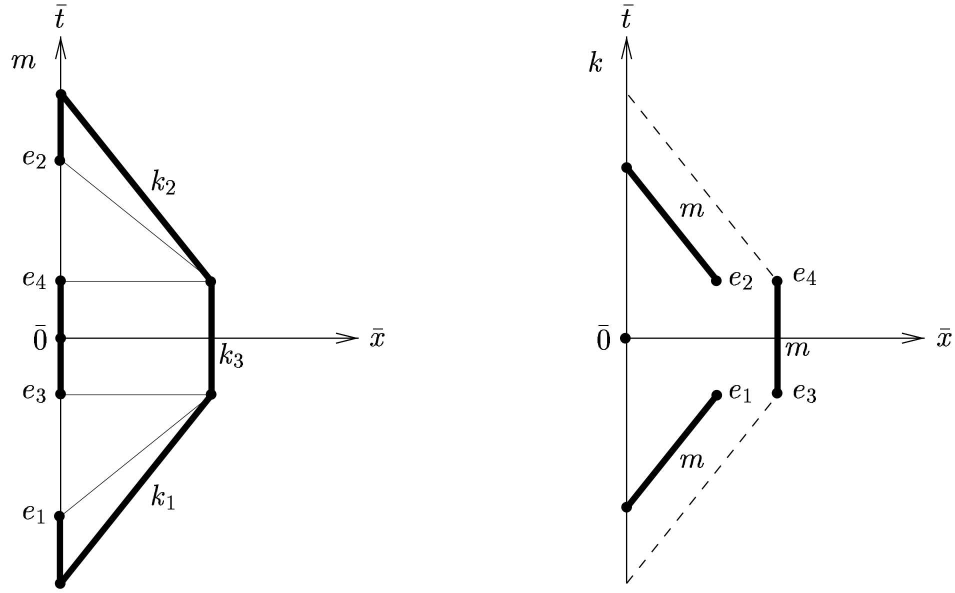
Figure 47 shows the "limit" of this approximating process. We concentrated on "smoothing out" the turning point on k's life-line, and we disregarded the initial and last segment of k's acceleration and deceleration (cf. the left-hand side of Figure 43).[155] Thus k goes outward with a constant speed v for a certain amount of time, then gradually (smoothly) he decelerates until he becomes momentarily at rest with respect to m, then he continues decelerating which means that he turns back and begins to gain speed[156] until he attains speed v again, and then he stops decelerating and approaches m with constant speed v until he reaches m. Cf. the left-hand side of Figure 47. This is how m observes k. Let us turn to how the accelerated twin k observes his inertial brother m. On Figure 47 we can see that k observes m first receding with a constant speed v, then m accelerates (increases (!) his speed), then m begins to decelerate till m becomes momentarily at rest w.r.t. k, and then m reverses this process. Thus the two life-lines are not alike: k's life-line, as observed by m, is "convex" in the sense that k's movement is uniform, it always decelerates.

[153] in order to get simpler drawings

[154] This is so because the extent of the paradigmatic effects increases with speed, they do not occur at speed 0, and the extent they occur with depends continuously of speed of movement. See Theorem 2.9.5 in §2.9.

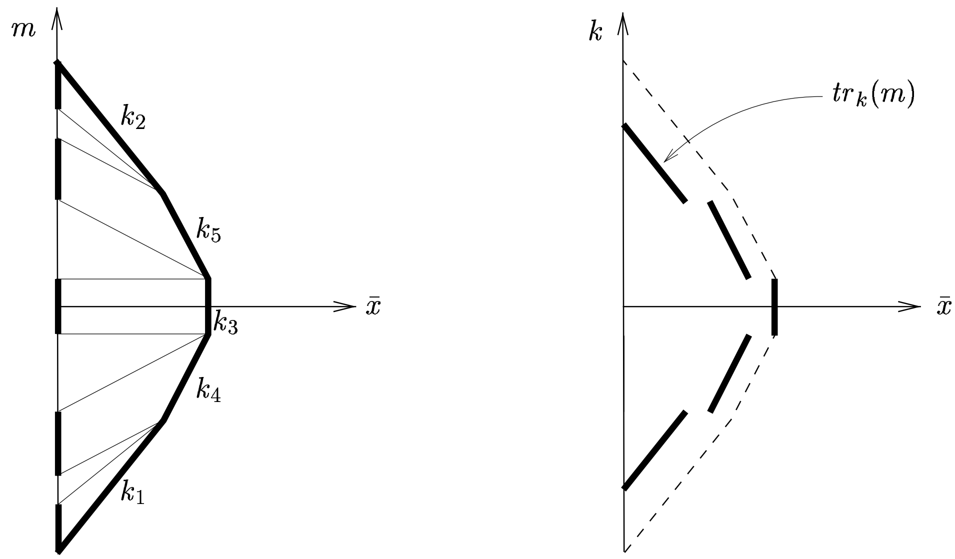
[155] For formulating and discussing the Twin Paradox, we do not need to assume that before event e_0 (or after e_3) the two twins k and m are at relative rest. Instead, we may assume that they simply meet at e_0 (moving with relative speed v). This way we can get rid of the initial (and final) acceleration without loosing anything essential. The acceleration "around" q, however, is essential, it cannot be "argued away" in the just used spirit.

[156] This is so because k's velocity changes gradually from $\vec{v}$ to $-\vec{v}$. So in terms of velocity, k's velocity is constantly decreasing. In terms of speed, this implies losing speed gradually from v to 0, and then gaining speed gradually from 0 to v again.



The traveling twin as approximated by three inertial observers

This is how m sees k This is how k sees m



The traveling twin as approximated by five inertial observers

Figure 46:

At the same time, m's life-line as k observes it is *both "convex and concave"*. This has to be so because of the following: $tr_m(k)$ and $tr_k(m)$ are both continuous (because in physics all movements are continuous), their initial and last segments are straight and parallel (because $v_m(k) = v_k(m)$ in the inertial parts of the journey), these segments are closer in m's life-line than in k's one (because for k less time has passed between departing and meeting, i.e. between e_0 and e_3, than for m), while the "width" of both life-lines are the same (because at the turning point k and m are at rest with respect to each other, so they see each other to be at the same spatial distance). See Figure 47.

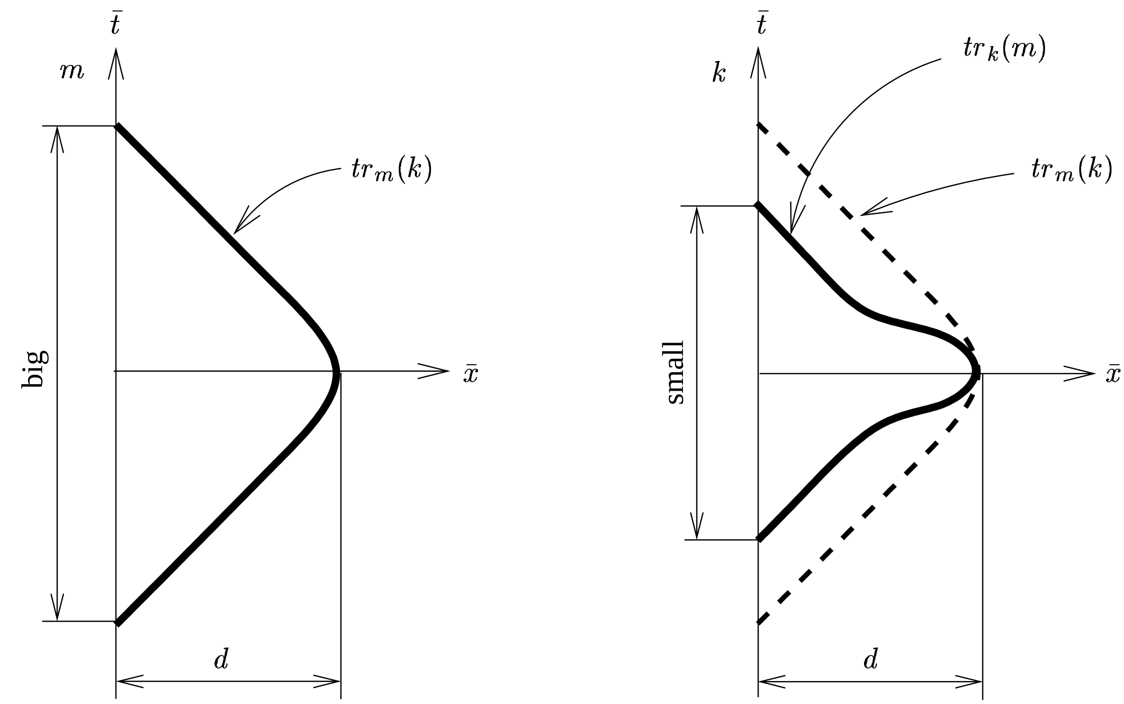


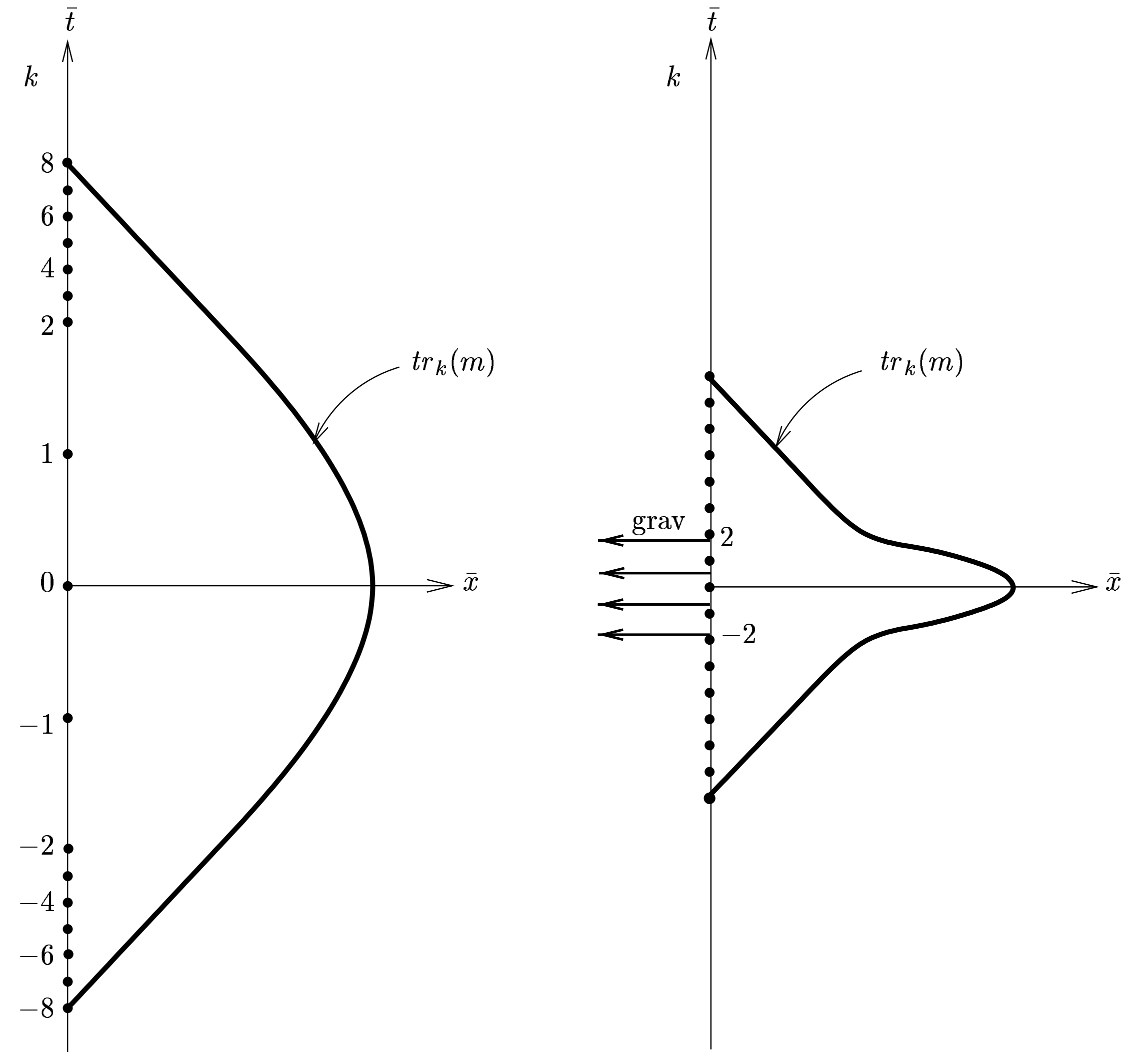
Figure 47: The inertial brother's life-line is different from that of the accelerated one.

We will return to the twin paradox in the chapter on accelerated observers, where we will begin to study gravity, too. Jumping ahead for a short while, let us see how k will "explain" m's strange movement (life-line) by using his knowledge about gravity. This explanation serves also to explain how the "laws of physics" can be the same for m and for k despite of the fact that they observe their brothers as behaving rather differently. The reader does not have to understand the explanation which comes below, since it uses (i) Einstein's equivalence principle (between acceleration and gravity) and some of the effects of gravity which we will prove in the chapter on accelerated observers, namely that (ii) gravity causes clocks run slow relative to clocks far away from the "source of" gravity, and (iii) in some sense gravity does not affect processes which happen sufficiently far away from the source of gravity. Therefore we advise the reader to read the explanation below as a "fairy tale" (which, in turn, will become easily understandable after studying the basic parts of the theory of accelerated observers).

The accelerated brother k thinks that he is at rest and m is moving away from him with speed v. When m is at a distance already, a gravitational field appears in k's worldview where k stands. To remain motionless despite of this strange gravitational field (which appeared "out of nowhere" so to speak), k starts up the engine of his spaceship to balance the effect of gravity. (As a constrast, m thinks that k started his engines in order to decelerate.) This gravity slows down k's clock, and this explains why, for k, m appears to accelerate first when gravity appears. See Figure 48. From this time on, since this gravity "pulls" m towards k, m begins to decelerate till it comes to a momentary rest w.r.t. k, then turns back and begins to "fall back" towards k with growing speed. When gravity disappears (then k stops the engine in order to stay motionless), m first slows down[157], and then reaches speed v and continues to approach k with constant speed v. (In passing we note that the reader might have the impression that for k sometimes m moves faster than light. However, this is not the case, because as a side-effect of *gravity* in k's coordinate system, at places far away from k the speed of light becomes larger than usual. This will be seen in the chapter on accelerated observers.)

In the above we used the expression "k observes" in place of the expression "k sees", because we wanted to emphasize that we meant everything according to k's coordinate system, and not according to how k actually "sees" via photons. Let us briefly turn to the visual effects, i.e. let us see how m and k visually see the journey via photons. See Figure 49. Again, we will find that the two brothers see the journey differently. The inertial brother will see k such that k travels outward (with slowed down clocks) for a long time and then he approaches (with fast running clocks) for a very short time. On the other hand, k will see that his inertial brother m travels outward (with slow clocks) for about the first half of the time needed for the whole experiment (i.e. until event $w_m(q)$ which is when m thinks that k turns around), and from that time on m approaches (with fast clocks). Thus for k, m's outward and inward parts of the journey (as k sees via photons) lasted approximately for

[157] because of the already mentioned effect of gravity on k's clocks



k knows that his clock slows down
in the interval of "acceleration"
(i.e. in the time interval around $t = 0$)

the same drawing (as on the left)
but taking the readings of
k's clock seriously

Figure 48: When k starts up his engine, k's clock slows down, and thus m's movement seems to speed up (m seems to accelerate).

the same time, while m will see (via photons) that k's journey outward lasted much longer than k's journey backward. If k decelerates only for a short time around its turning point, then this difference of ratio of outward and inward trips as k and m see them via photons will remain.

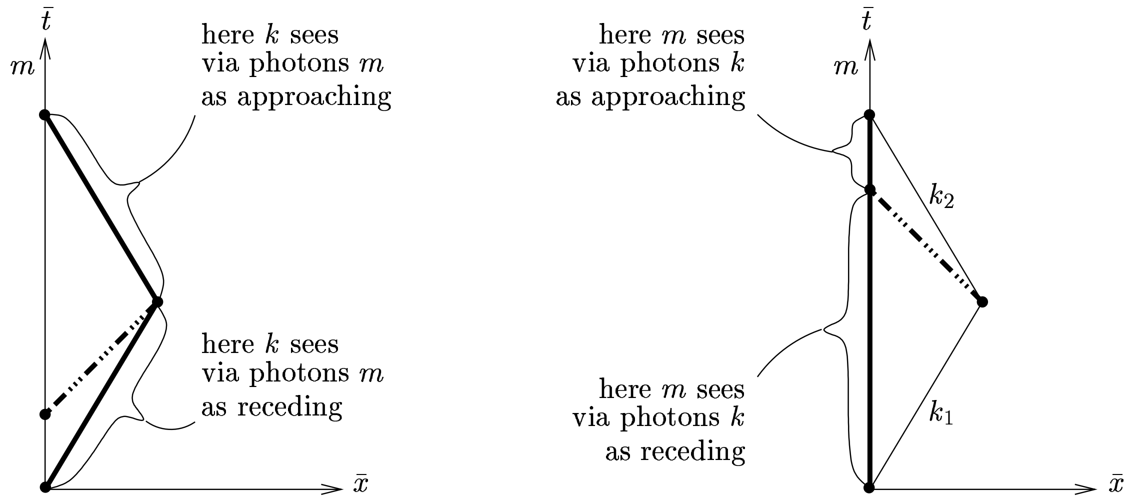


Figure 49: The two brothers' visual observations of each other's journey are also different.

2.8.5 Our central axiom system Specrel.

We conclude this part with introducing one of our central axiom systems for special relativity. Theorem 2.9.5 in §2.9 (p.155) states that the world-view transformations in models of **Basax** + **Ax(symm)** are all so called Poincaré transformations, i.e. such world-view transformations which occur in the standard models of special relativity. Thus, in models of **Basax**+**Ax(symm)** all the usual formulae for coordinate-transformations, used in the physics books, are valid. Therefore, models of **Basax** + **Ax(symm)** are very close to the standard, Minkowskian models. In proofs we often will need the auxiliary axioms **Ax(***Triv***)** and **Ax(**$\|$**)**. Though **Ax(**$\|$**)** follows from **Basax** + **Ax(symm)**, for later weaker versions of **Basax** this will not be so, e.g. **Bax** + **Ax(symm)** $\not\models$ **Ax(**$\|$**)**, here **Bax** is an axiom system to be introduced in §3.4.2. Therefore we define one of our stronger kind[158] of special relativity theories as follows.

Definition 2.8.20

$$\mathbf{Ax(symm)}^{\dagger} \stackrel{\text{def}}{=} \mathbf{Ax(symm)} + \mathbf{Ax}(\mathit{Triv}) + \mathbf{Ax}(\|)$$

$$\mathbf{Specrel} \stackrel{\text{def}}{=} \mathbf{Basax} + \mathbf{Ax(symm)}^{\dagger}.$$

$\lhd$

Specrel is a first-order-logic theory of special relativity that is basically equivalent with the standard version of special relativity theory. The only omissions (missing from **Specrel**) are some auxiliary axioms that we almost never use, see the definitions of **BaCo** and[159] Minkowski model in §3.8, pages p.298, p.331. For more on this see §3.8.

[158]As indicated in the introduction, in the present work we will have stronger axiomatic versions as well as weaker axiomatic versions of (special) relativity. At each point, we will choose between the stronger or weaker versions depending on our purposes at that point, cf. e.g. items II, III, V in §1.1 herein.

[159]**BaCo** is a complete axiomatization of (what we consider as) usual special relativity.

2.9 Connections with standard Lorentz and Poincaré transformations

In order to compare our results with the literature, we recall some standard concepts from the literature.

Definition 2.9.1 Assume $\mathfrak{F} = \langle \mathbf{F}, \leq \rangle$ is an ordered field and $n \geq 2$.

1. $Linb = Linb(n, \mathfrak{F})$ denotes the set of bijective linear transformations of ${}^n\mathbf{F}$.

2. Let $p \in {}^nF$. Then $\tau_p : {}^nF \longrightarrow {}^nF$ denotes the *translation* by vector p, defined as follows:
$$\tau_p \stackrel{\text{def}}{=} \langle\, q + p \;:\; q \in {}^nF \,\rangle \,.$$
$Tran = Tran(n, \mathfrak{F})$ denotes the set of translations of ${}^n\mathbf{F}$, i.e.
$Tran \stackrel{\text{def}}{=} \{\, \tau_p \;:\; p \in {}^nF \,\}$.

3. A function $\mathsf{f} : {}^nF \longrightarrow {}^nF$ is called an *affine transformation* of ${}^n\mathbf{F}$ iff it is a composition of an bijective linear transformation and a translation, i.e.
$\mathsf{f} = \mathsf{g} \circ \tau_p$, for some $\mathsf{g} \in Linb(n, \mathfrak{F})$ and $p \in {}^nF$.

   $Aftr = Aftr(n, \mathfrak{F})$ denotes the set of affine transformations of ${}^n\mathbf{F}$.

4. Let $p, q \in {}^nF$. Then the square of their *Minkowski-distance* $g^2_\mu(p, q)$ is defined as follows:
$$g^2_\mu(p,q) \stackrel{\text{def}}{=} \left| (q_0 - p_0)^2 - \left(\sum\nolimits_{0<i\in n} (q_i - p_i)^2 \right) \right| .$$
We note that $g^2_\mu : {}^nF \times {}^nF \longrightarrow F$.

5. By a *Lorentz transformation* of ${}^n\mathbf{F}$ we understand $\mathsf{f} \in Linb$ such that f preserves the square of Minkowski-distance, that is,
$$(\star) \qquad (\forall p, q \in {}^nF)\ g^2_\mu(p,q) = g^2_\mu(\mathsf{f}(p), \mathsf{f}(q)) \,.$$

   $Lor = Lor(n, \mathfrak{F})$ denotes the set of Lorentz transformations of ${}^n\mathbf{F}$.

6. By a *standard Lorentz transformation*[160] we understand a Lorentz transformation f such that
$$\mathsf{f}[\bar{t}\,], \mathsf{f}[\bar{x}\,] \subseteq \mathsf{Plane}(\bar{t}, \bar{x}) \quad \text{and} \quad (\forall 1 < i \in n)\ \mathsf{f}(1_i) = 1_i \,.$$

[160] Or equivalently a Lorentz transformation in standard configuration.

$SLor = SLor(n, \mathfrak{F})$ denotes the set of standard Lorentz transformations of ${}^n\mathbf{F}$.

7. By a <u>*Poincaré transformation*</u> of ${}^n\mathbf{F}$ we understand $\mathsf{f} \in Aftr$ such that f preserves the square of Minkowski-distance, that is, $(\star)$ in item 5 holds for f.

   $Poi = Poi(n, \mathfrak{F})$ denotes the set of Poincaré transformations.[161]

8. An bijective linear transformation f of ${}^n\mathfrak{F}$ is called an <u>*expansion*</u>[162] iff

$$(\exists 0 < \lambda \in F) \quad \mathsf{f} = \langle\, \lambda \cdot p \;:\; p \in {}^nF \,\rangle \,.$$

   $Exp = Exp(n, \mathfrak{F})$ denotes the set of expansions.

$\lhd$

CONVENTION 2.9.2 For better readability, the elements of Exp and Lor will often be denoted by exp and lor, respectively. Similarly for $Linb$, $SLor$, Poi, $Rhomb$ etc.

$\lhd$

For completeness, we note that our distinguished sets of transformations are contained in each other in the following way:

$$\begin{array}{ccccccc} SLor & \subset & Lor & \subset & Linb & \supset & Exp \\ & & \cap & & \cap & & \\ Tran & \subset & Poi & \subset & Aftr\,, & & \end{array}$$

where $\begin{array}{c} A \\ \cap \\ B \end{array}$, $B \supset A$, etc. all denote that A is a proper subset of B.

Thm. 2.9.4 below is a kind of <u>*characterization*</u> of the world-view transformations f_{mk} in $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ })$. Intuitively, it says (assuming $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ })$) that a world-view transformation f_{mk} is always a composition of a Poincaré transformation, an expansion, and a map $\widetilde{\varphi}$ induced by an automorphism φ of the ordered field $\mathfrak{F}$ (cf. Notation 2.9.3 below for $\widetilde{\varphi}$). Moreover all such compositions are world-view transformations (of some **Basax** model), if we assume that $\mathfrak{F}$ is Euclidean. To formulate this theorem we need Notation 2.9.3 below.

[161] An equivalent definition says that a Poincaré transformation is a composition of a Lorentz transformation and a translation, i.e. is of the form $lor \circ \tau_p$, for some $lor \in Lor$ and $p \in {}^nF$.

[162] We note that the official name for an expansion is a <u>*transformation of similitude*</u>. (Coxeter [62] uses the word <u>*dilation*</u> while Burke [52] calls it an expansion. Sometimes it is also called <u>*homothetic transformation*</u>). For reasons of convenience we restricted the notion of an expansion for multiplying with positive λ's only.

Notation 2.9.3

- $Aut(\mathfrak{F})$ denotes the set of automorphisms of the ordered field $\mathfrak{F}$. For any algebraic structure or model $\mathfrak{A}$, $Aut(\mathfrak{A})$ is defined similarly (i.e. is the set of automorphisms of $\mathfrak{A}$).

- $\widetilde{\varphi}$ denotes the function induced by other function φ the following way. Assume $\varphi : F \longrightarrow F$. Then the induced function $\widetilde{\varphi} : {}^nF \longrightarrow {}^nF$ is defined the natural way, i.e.

$$\widetilde{\varphi}(p) \stackrel{\text{def}}{=} \langle \varphi(p_0), \varphi(p_1), \ldots, \varphi(p_{n-1}) \rangle, \quad \text{for every } p \in {}^nF.$$

$\lhd$

THEOREM 2.9.4 (Characterization of the world-view transformations in models of Basax) *Assume* $\mathbf{Basax} + \mathbf{Ax}(\sqrt{\ })$. *Let* $m, k \in Obs$. *Then:*

(i) $\mathsf{f}_{mk} = poi \circ exp \circ \widetilde{\varphi}$, *for some* $poi \in Poi$, $exp \in Exp$ *and* $\varphi \in Aut(\mathfrak{F})$.

(ii) *Assume in addition that* $\mathsf{f}_{mk}(\bar{0}) = \bar{0}$. *Then*
$\mathsf{f}_{mk} = lor \circ exp \circ \widetilde{\varphi}$, *for some* $lor \in Lor$, $exp \in Exp$ *and* $\varphi \in Aut(\mathfrak{F})$.

(iii) *Let* $\mathfrak{F}$ *be a fixed Euclidean ordered field. Assume* f *is a composition of a Poincaré transformation, an expansion, and a map* $\widetilde{\varphi}$, *for some* $\varphi \in Aut(\mathfrak{F})$. *Then there is a* **Basax** *model* $\mathfrak{N}$ *with ordered field reduct* $\mathfrak{F}$ *such that* $\mathsf{f} = \mathsf{f}_{m'k'}$, *for some* $m', k' \in Obs^{\mathfrak{N}}$.

The **proof** will be given in §3.7. ∎

Let us recall that the symmetry axiom **Ax(symm)** was introduced in §2.8 on p.127.

The following theorem says that, under assuming **Basax** + **Ax(symm)** + $\mathbf{Ax}(\sqrt{\ })$, a world-view transformation f_{mk} is a Poincaré transformation. Moreover all Poincaré transformations over a Euclidean $\mathfrak{F}$ are world-view transformations in some **Basax+Ax(symm)** model. That is, **Ax(symm)** implies that expansions and automorphisms are not needed in the above characterization of world-view transformations.

THEOREM 2.9.5 (Characterization of the world-view transformations in models of Basax + Ax(symm)) *Assume* **Basax + Ax(symm) + Ax($\sqrt{\ }$)**. *Let $m, k \in Obs$. Then (i)–(iii) below hold.*

(i) $\mathsf{f}_{mk} \in Poi$.

(ii) *Assume in addition that* $\mathsf{f}_{mk}(\bar{0}) = \bar{0}$. *Then* $\mathsf{f}_{mk} \in Lor$.

(iii) *Let $\mathfrak{F}$ be a fixed Euclidean ordered field. Let $\mathsf{f} \in Poi(n, \mathfrak{F})$. Then there is a* **Basax + Ax(symm)** *model $\mathfrak{N}$ whose ordered field reduct is $\mathfrak{F}$ such that* $\mathsf{f} = \mathsf{f}_{m'k'}$, *for some $m', k' \in Obs^{\mathfrak{N}}$.*

The **proof** will be given in §3.8. Here we show the idea of proof of **(ii)** in the case of $n = 2$. Assume that $\mathfrak{M} \models$ **Basax + Ax(symm) + Ax($\sqrt{\ }$)**, $m, k \in Obs$ and $\mathsf{f}_{mk}(\bar{0}) = \bar{0}$. Then each of m and k thinks that the other's clock is slow (by Thm.2.8.7), and moreover the rate of slowing down is the same for both of them (see Thm.2.8.9). Figure 50 shows how this is possible. By using this figure, it is not difficult to show that the unique place where $e = \mathsf{f}_{mk}(1_t)$ can be is such that the Minkowski-distance between $\bar{0}$ and e is 1.

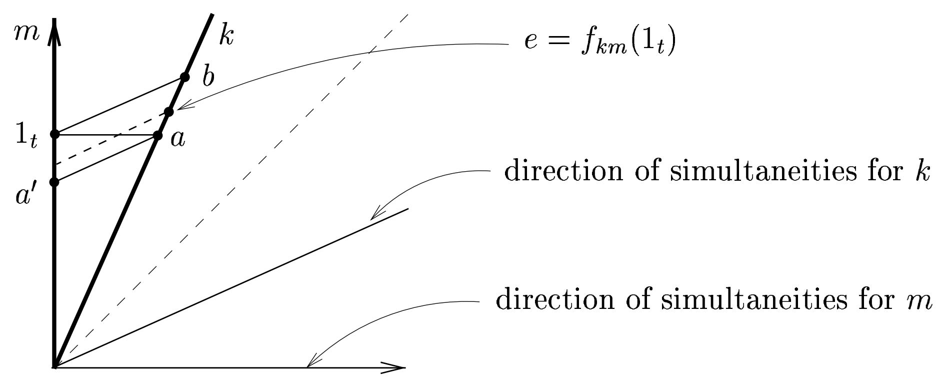


Figure 50: Both m and k think that the other's clock slows down iff $f_{mk}(1_t)$ is in between a and b. The rates of slowing down will be equal at a unique point. This unique point is closer to a than to b, and a geometrical construction for it is given in § 3. The Minkowski-distance between $\bar{0}$ and e is 1.

∎

Recall that the set $Rhomb(n, \mathfrak{F})$ of rhombus transformations was defined in Def.2.3.18 (p.72).

The following theorem says that, under some mild assumptions, rhombus transformations are compositions of standard Lorentz transformations and expansions. Moreover all such compositions are rhombus transformations.

THEOREM 2.9.6

(i) *Assume $\mathfrak{F}$ is Euclidean, i.e. that $\mathfrak{F} \models \mathbf{Ax}(\sqrt{\ })$. Then*

$$Rhomb(n, \mathfrak{F}) = \{\, slor \circ exp \ : \ slor \in SLor \quad and \quad exp \in Exp \,\}.$$

(ii) *Assume $n > 2$. Then*

$$Rhomb(n, \mathfrak{F}) = \{\, slor \circ exp \ : \ slor \in SLor \quad and \quad exp \in Exp \,\}.$$

(iii)
$$Rhomb(n, \mathfrak{F}) \supseteq \{\, slor \circ exp \ : \ slor \in SLor \quad and \quad exp \in Exp \,\}.$$

The **proof** will be filled in later. ∎