AXIOMATIZING RELATIVISTIC DYNAMICS WITHOUT CONSERVATION POSTULATES

H. ANDRÉKA, J. X. MADARÁSZ, I. NÉMETI AND G. SZÉKELY

ABSTRACT. A part of relativistic dynamics (or mechanics) is axiomatized by simple and purely geometrical axioms formulated within first-order logic. A geometrical proof of the formula connecting relativistic and rest masses of bodies is presented, leading up to a geometric explanation of Einstein's famous $E = mc^2$. The connection of our geometrical axioms and the usual axioms on the conservation of mass, momentum and four-momentum is also investigated.

1. INTRODUCTION

The idea of elaborating the foundational analysis of the logical structure of spacetime theory and relativity theories (foundation of relativity) in a spirit analogous with the rather successful foundation of mathematics was initiated by several authors including David Hilbert [15], cf. also [14, 6th problem], Patrick Suppes [22], Alfred Tarski [13] and leading contemporary logician Harvey Friedman [11], [12].

There are several reasons for seeking an axiomatic foundation of a physical theory [23]. One is that the theory may be better understood by providing a basis of explicit postulates for the theory. Another reason is that if we have an axiom system we can ask ourselves which axioms are responsible for which theorems. For more on this kind of foundational thinking called reverse mathematics, see, e.g., Friedman [11] and Simpson [21]. Furthermore, if we have an axiom system for special or general relativity, we can ask what happens with the theory if we change one or more of the axioms. That could lead us to a new physically interesting theory. That is what happened with Euclid's axiom system for geometry when Bolyai and Lobachevsky altered the axiom of parallelism which lead to the discovery of hyperbolic geometry.

In the above spirit, in earlier works the Relativity and logic group of Rényi Mathematical Institute in Budapest built up relativity theories (both special and general) purely in the framework of first-order logic (FOL). This foundation of relativity is elaborated in strict parallel to the success story of the foundation of mathematics, cf. also [1].

Research supported by Hungarian National Foundation for Scientific Research grant No T43242 as well as by Bolyai Grant for Judit X. Madarász.

Why do we insist on staying within FOL as a framework? For good reasons, the foundation of mathematics has been carried through strictly within the framework of first-order logic. One of these reasons is that staying within FOL helps us to avoid tacit assumptions. Another reason is that FOL has a complete inference system while higher-order logic cannot have one by Gödel's incompleteness theorem, see, e.g., [25, p.505]. For more motivation for staying inside FOL as opposed to higher-order logic, see, e.g., [3], [4, Appendix 1: Why exactly FOL], [7], [10], [18], [26]. The same reasons motivate the effort of keeping the foundation of spacetime and relativity theory inside FOL.

In our earlier works we concentrated on the kinematics of relativity theories. The present paper is devoted to a part of relativistic dynamics or mechanics. In particular, we present an axiom system SpecRelDyn for relativistic inertial mass. It is an extension of our earlier axiom system **SpecRel** used for the kinematics of special relativity. Just as we did in SpecRel, we try to keep our axioms as few as possible and at the same time convincing, transparent and easy to comprehend even for someone not familiar with the basic concepts of physics. We also try to keep our axioms visualizable and purely geometrical. Based on SpecRelDyn, we present a purely geometrical proof for the theorem that relates the relativistic mass of a moving particle to its rest mass. The usual approach in standard relativity texts goes by assuming as new axioms the conservation of relativistic mass and conservation of momentum, cf. d'Inverno [9, p.43-36] and Rindler [19, pp.108-112]. These are very strong assumptions compared to ours, and by our above mentioned proof, these strong assumptions are not needed for introducing or explaining relativistic mass. We base our theory on more basic and more geometrical axioms. Being more basic and geometrical, these axioms are also more elementary and more self-evident.

In Section 2 we fix the first-order language for dynamics of special relativity theory. In Section 3 we recall the streamlined FOL axiom system **SpecRel** used for kinematics of special relativity theory from our previous works. In Section 4 we extend **SpecRel** to cover relativistic dynamics leading to Einstein's famous insight $E = mc^2$. In Section 5 we present a purely geometric axiom that is equivalent to conservation of mass and momentum. This axiom is also proved to be equivalent to the conservation of four-momentum. In Section 6 we sketch some possible future research directions.

2. A first-order logic frame for relativity theory

The motivation for our choice of vocabulary (basic concepts) is summarized as follows. We represent motion as changing spatial location in time. To do so, we will have reference-frames for coordinatizing events (sets of bodies) and, for simplicity, we will associate reference-frames with certain bodies which we will call *observers*. We visualize an observer as "sitting" in the origin of the space part of its reference-frame, or equivalently, "living" on the time-axis of the reference-frame. There will be another special kind of bodies which we will call *photons*. For coordinatizing events, we will use an arbitrary *ordered field* in place of the field of real numbers. Thus the elements of this field will be the *quantities* which we will use for marking time and space. In the axioms of dynamics we will use *relativistic masses* of bodies as a basic concept.

Allowing arbitrary ordered fields instead of the field of reals increases the flexibility of our theory and minimizes the amount of our mathematical presuppositions, see, e.g., Ax [7] for further motivation in this direction. Similar remarks apply to our flexibility oriented decisions below, e.g., the one to treat the dimension of spacetime as a variable.

Using observers in place of coordinate systems or reference frames is only a matter of didactic convenience and visualization. There are many reasons for using observers (or coordinate systems, or referenceframes) instead of a single observer-independent spacetime structure. One of them is that it helps us to weed unnecessary axioms from our theories; but we state and emphasize the logical equivalence between observer-oriented and observer-independent approaches to relativity theory elaborated in, e.g., [16, §4.5] and [5]. Motivated by the above, we now turn to fixing the first-order language of our axiom systems.

First we fix a natural number $\underline{d} \geq 2$ for the dimension of spacetime. Our language contains the following non-logical symbols:

- unary relation symbols **B** (for **bodies**), **IOb** (for inertial **observers**), **Ph** (for **photons**) and **Q** (for **quantities**),
- binary function symbols
 →,
 → and a binary relation symbol
 ≤
 (for the field operations and the ordering on Q),
- a 2+d-ary relation symbol W (for world-view relation), and
- a 3-ary relation symbol **M** (for **mass relation**).

We translate B(x), IOb(x), Ph(x) and Q(x) into natural language as "x is a body," "x is an observer," "x is a photon," and "x is a quantity." (A more careful wording would be "x is a possible body," "x is a possible observer," etc.) The bodies play the role of the "main characters" of our spacetime models and they are "observed" (coordinatized using the quantities) by the observers. This observation is coded by the world-view relation by translating $W(x, y, z_1, \ldots, z_d)$ as "observer x coordinatizes body y at spacetime location $\langle z_1, \ldots, z_d \rangle$," (that is, at space location $\langle z_2, \ldots, z_d \rangle$ at instant z_1). Finally we use the mass relation to speak about the relativistic masses of bodies according to observer x."

B(x), IOb(x), Ph(x), Q(x), $W(x, y, z_1, \ldots, z_d)$, M(x, y, z), x = y and $x \leq y$ are the atomic formulas of our first-order language, where x, y, z_1, \ldots, z_d can be arbitrary variables or terms built up from variables

by using the field-operations. The **formulas** of our first-order language are built up from these atomic formulas by using the logical connectives *not* (\neg) , *and* (\land) , *or* (\lor) , *implies* (\Longrightarrow) , *if-and-only-if* (\iff) and the quantifiers *exists* x $(\exists x)$ and *for all* x $(\forall x)$ for every variable x.

The **models** of this language are of the form

$$\langle U; B, IOb, Ph, Q, +, \cdot, \leq, W, M \rangle$$
,

where U is a non-empty set and B, IOb, Ph and Q are unary relations on U, etc. A unary relation on U is just a subset of U. Thus we use B, IOb etc. as sets as well, e.g., we write $k \in IOb$ in place of IOb(k).

We use the notation $\mathbb{Q}^n := \mathbb{Q} \times \ldots \times \mathbb{Q}$ (*n*-times) for the set of all *n*-tuples of elements of Q. If $p \in \mathbb{Q}^n$, then we assume that $p = \langle p_1, \ldots, p_n \rangle$, that is, $p_i \in \mathbb{Q}$ denotes the *i*-th component of the *n*-tuple *p*. We write W(m, b, p) in place of $W(m, b, p_1, \ldots, p_d)$, and we write $\forall p$ in place of $\forall p_1, \ldots, p_d$ etc.

We present each axiom at two levels. First we give an intuitive formulation, then we give a precise formalization using our logical notation (which can easily be translated into first-order formulas by inserting the definitions into the formalizations). We seek to formulate easily understandable axioms in FOL.

The first axiom expresses our very basic assumptions, such as: both photons and observers are bodies, etc.

AxFrame: $IOb \cup Ph \subseteq B$, $W \subseteq IOb \times B \times Q^d$, $M : IOb \times B \to Q$ is a function, M(k,b) > 0 for every observer k and body b, $B \cap Q = \emptyset$, + and \cdot are binary operations and \leq is a binary relation on Q.

To be able to add, multiply and compare measurements of observers, we put an algebraic structure on the set of quantities by the next axiom.

AxEOF: The quantity part $\langle Q; +, \cdot, \leq \rangle$ is a Euclidean¹ ordered field.

For the first-order logic definition of linearly ordered field, see, e.g. [8]. We use the usual field operations $0, 1, -, /, \sqrt{}$ definable within FOL. We also use the vector-space structure of Q^n , that is, if $p, q \in Q^n$ and $\lambda \in Q$, then $p+q, -p, \lambda \cdot p \in Q^n$; and $O := \langle 0, \ldots, 0 \rangle$ denotes the **origin**. The **Euclidean length** of $p \in Q^n$ is defined as $|p| := \sqrt{p_1^2 + \ldots + p_n^2}$, for any $n \ge 1$.

CONVENTION 2.1. We treat AxFrame and AxEOF as a part of our logical frame throughout this paper. Hence, without any further mentioning, they will be always assumed and will be part of every axiom system we propose herein.

¹That is, a linearly ordered field in which positive elements have square roots.

3. KINEMATICS

In this section we recall the streamlined axiom system SpecRel for kinematics of special relativity theory from our previous works. We note that SpecRel is extended in our works, e.g. in [17],[5], to deal with accelerated observers and general relativity.

 Q^d is called the **coordinate system** and its elements are referred to as **coordinate points**. We use the notations

$$p_{\sigma} := \langle p_2, \ldots, p_d \rangle$$
 and $p_{\tau} := p_1$

for the **space component** and for the **time component** of $p \in Q^d$, respectively.

The **event** $ev_k(p)$ is the set of bodies observed by observer k at coordinate point p is, that is,

$$ev_k(p) := \{ b \in \mathcal{B} : \mathcal{W}(k, b, p) \}$$

The **world-line** of body b according to observer k is defined as the set of coordinate points where b was observed by k, that is,

$$wl_k(b) := \left\{ p \in \mathbf{Q}^d : b \in ev_m(p) \right\}.$$

Now we formulate our first axiom on observers. (Historically this natural axiom goes back to Galileo Galilei or even to d'Oresme of around 1350, but probably it is much more ancient than that, see, e.g., [3, p.23, §5].)

AxSelf: Each observer k is motionless in the origin of the space part of his coordinate system, that is, his world-line is the time-axis:

$$\forall k \in \text{IOb} \quad wl_k(k) = \{ \langle \lambda, 0, \dots, 0 \rangle : \lambda \in \mathbf{Q} \}.$$

As a formula of first-order logic this axiom is:

 $\forall k \in \text{IOb} \ \forall p \in \mathbf{Q}^d \quad [\mathbf{W}(k,k,p) \iff p_2 = \ldots = p_d = 0].$

Now we formulate our axiom about the constancy of the speed of photons. For convenience, we choose 1 for this speed.

AxPh: The world-lines of photons are of slope 1, and moreover, for every observer, there is a photon through two coordinate points if their slope is 1:

$$\forall k \in \text{IOb } \forall p, q \in \mathbf{Q}^d \quad [|p_{\sigma} - q_{\sigma}| = |p_{\tau} - q_{\tau}| \iff \exists ph \in \text{Ph} \quad ph \in ev_k(p) \cap ev_k(q)].$$

This axiom is a well-known assumption of special relativity, see, e.g., [5], [9, §2.6]. In a more careful interpretation of our logical formalism, instead of "photons" and "bodies" we could speak about "possible world-lines of photons" and "possible world-lines of bodies," etc. We chose the present usage for brevity.

AxEv : All observers coordinatize the same events:

$$\forall k, h \in \text{IOb} \ \forall p \in \mathbf{Q}^d \ \exists q \in \mathbf{Q}^d \ ev_k(p) = ev_h(q).$$

The **world-view transformation** between the world-views of observers k and h is the set of pairs of coordinate points $\langle p, q \rangle$ such that k and h observe the same event in p and q, respectively:

$$v_h^k := \{ \langle p, q \rangle \in \mathbf{Q}^d \times \mathbf{Q}^d : ev_k(p) = ev_h(q) \}.$$

As usual, ℓ is called a **line** iff there are $p, q \in \mathbb{Q}^d$ such that $q \neq O$ and $\ell = \{p + \lambda q : \lambda \in \mathbb{Q}\}.$

Remark 3.1. Assume $d \geq 3$ and AxSelf, AxPh and AxEv. Then

- (i) World-view transformations take lines to lines, see [5, Thm.11.11.(ii)].
- (ii) World-lines of observers are lines by (i) and AxSelf.
- (iii) No observer can travel faster than light, see [5, Thm.11.7].

By the next axiom we assume that observers use the same units of measurements.

AxSimDist : Any two observers agree as for the spatial distance between two events if these two events are simultaneous for both of them:

$$\forall k, h \in \text{IOb} \ \forall p, q, p', q' \in \mathbb{Q}^d \ \left[\left(ev_k(p) = ev_h(p') \land ev_k(q) = ev_h(q') \right) \\ \land p_\tau = q_\tau \land p'_\tau = q'_\tau \right) \implies |p_\sigma - q_\sigma| = |p'_\sigma - q'_\sigma| \right].$$

Let us introduce an axiom system for special relativistic kinematics:

$$\mathsf{SpecRel} := \{\mathsf{AxSelf}, \mathsf{AxPh}, \mathsf{AxEv}, \mathsf{AxSimDist}\}$$

Let $p, q \in \mathbb{Q}^d$. Then

$$\mu(p) := \begin{cases} \sqrt{p_{\tau}^2 - |p_{\sigma}|^2} & \text{if } p_{\tau}^2 - |p_{\sigma}|^2 \ge 0, \\ -\sqrt{|p_{\sigma}|^2 - p_{\tau}^2} & \text{otherwise} \end{cases}$$
(1)

is the (signed) **Minkowski length** of p and the **Minkowski distance** between p and q is defined as follows:

$$\mu(p,q) := \mu(p-q). \tag{2}$$

Function $f : \mathbb{Q}^d \to \mathbb{Q}^d$ is said to be a **Poincaré transformation** if it is a bijection and it preserves the Minkowski distance, that is, $\mu(f(p), f(q)) = \mu(p, q)$ for all $p, q \in \mathbb{Q}^d$. We note that every Poincaré transformation is a linear transformation composed by a translation. For proof of the following theorem see Thm.11.10 in [5].

Theorem 3.2. Assume $d \geq 3$ and SpecRel. Then w_h^k is a Poincaré transformation for every $k, h \in IOb$.

Thus from SpecRel if $d \ge 3$, we can deduce the most frequently quoted predictions of special relativity:

(i) "moving clocks slow down,"

- (ii) "moving meter-rods shrink" and
- (iii) "moving pairs of clocks get out of synchronism."

Moreover, SpecRel implies the exact amount of time-dilation, lengthcontraction and delay of clocks. So if $d \ge 3$, SpecRel captures the kinematics of special relativity well. For more detail, see, e.g., [3, 4, 5].

We often add axioms to SpecRel which do not change the spacetime structure, but are useful auxiliary or bookkeeping axioms. For example, AxThEx below states that each observer can make thought experiments in which he assumes the existence of "slowly moving" observers (see e.g. [5, p.622 and Thm.2.9(iii)]):

AxThEx: For each observer, in each spacetime location, in each direction, with any speed less than that of light it is possible to "send out" an observer whose time flows "forwards":

$$\forall k \in \text{IOb } \forall p, q \in \mathbb{Q}^d \ \exists h \in \text{IOb } \left[|(p-q)_{\sigma}| < (p-q)_{\tau} \implies p, q \in wl_k(h) \text{ and } w_h^k(q)_{\tau} < w_h^k(p)_{\tau} \right].$$

4. Dynamics

In this section we shall formulate our axioms on dynamics. The idea is that we use inelastic collisions for observing (or measuring) relativistic inertial mass. We could say that relativistic inertial mass is the quantity that shows the magnitude of the influence of the body on the state of motion of the body it collides with. The more a body changes the motion of bodies it collides with, the bigger its relativistic mass is.

To formulate our axioms on relativistic mass, first we define inelastic collisions. The sets $in_k(q)$ of incoming bodies and $out_k(q)$ of outgoing bodies of the collision at coordinate point q according to observer k are defined as bodies whose lifelines "end" and "start" at q respectively (see Fig.1):

$$in_k(q) := \{b \in \mathcal{B} : q \in wl_k(b) \land \forall p \in wl_k(b) \ [p_\tau < q_\tau \lor p = q]\},\$$
$$put_k(q) := \{b \in \mathcal{B} : q \in wl_k(b) \land \forall p \in wl_k(b) \ [p_\tau > q_\tau \lor p = q]\}.$$

Bodies b and c collide inelastically originating body d according to observer k, in symbols inecoll_k(b, c : d), iff $b \neq c$ and there is a coordinate point q such that $in_k(q) = \{b, c\}$ and $out_k(q) = \{d\}$, see the right-hand side of Fig.1.

Recall that by AxFrame, M : IOb×B \rightarrow Q is a function and M(k, b) > 0 for every observer k and body b. If k is an observer and b is a body then we call $m_k(b) := M(k, b)$ the relativistic mass of body b according to observer k, or equivalently, "... in the world-view of k".

The **spacetime location** $loc_k(b,t)$ of body b at time instance $t \in \mathbf{Q}$ according to observer k is defined to be the coordinate point p for which $p \in wl_k(b)$ and $p_{\tau} = t$ if there is such a unique p, and it is undefined otherwise, see Fig.2.



FIGURE 1. Illustration of relations $in_k(q)$, $out_k(q)$ and $inecoll_k(b, c: d)$

The **center of mass** $cen_k(b, c, t)$ of bodies b and c at time instance t according to observer k is defined to be the coordinate point q such that $q_{\tau} = t$ and q_{σ} is the point on the line-segment between $loc_k(b, t)$ and $loc_k(c, t)$ whose distances from these two end-points have the same proportion as that of the relativistic masses of b and c; and it is closer to the "more massive" body, i.e.:

$$m_k(b) \cdot \left(loc_k(b,t) - cen_k(b,c,t) \right) = m_k(c) \cdot \left(cen_k(b,c,t) - loc_k(c,t) \right)$$

if $loc_k(b, t)$ and $loc_k(c, t)$ are defined, and $cen_k(b, c, t)$ is undefined otherwise, see Fig.2. We note that an explicit definition for $cen_k(b, c, t)$ is the following:

$$cen_k(b, c, t) = \frac{m_k(b)}{m_k(b) + m_k(c)} \cdot loc_k(b, t) + \frac{m_k(c)}{m_k(b) + m_k(c)} \cdot loc_k(c, t),$$

(if $loc_k(b, t)$ and $loc_k(c, t)$ are defined and $cen_k(b, c, t)$ is undefined otherwise). The **center-line of mass** of bodies b and c according to observer k is defined as

$$\underline{cen_k(b,c)} := \{cen_k(b,c,t) : t \in \mathbf{Q} \text{ and } cen_k(b,c,t) \text{ is defined} \}.$$

Intuitively, the center-line of mass is the world-line of the center of mass. The segment determined by $p, q \in \mathbf{Q}^d$ is defined as:

$$[p,q] := \{\lambda \cdot p + (1-\lambda) \cdot q : \lambda \in \mathbf{Q}, \ 0 \le \lambda \le 1\}.$$

We call $H \subseteq \mathbb{Q}^d$ line segment iff H is connected (i.e., $[p,q] \subseteq H$ for all $p, q \in H$), H has at least two elements, and H is contained in a line.

Bodies whose world-lines are line segments are called **inertial bod-ies**, and their set is defined as:

$$\mathbf{Ib} := \{ b \in \mathbf{B} : \forall k \in \mathbf{IOb} \quad wl_k(b) \text{ is a line segment} \}$$

We note that $cen_k(b, c)$ is a line segment or a point or the empty set and $wl_k(b) \cap wl_k(c) \subseteq cen_k(b, c)$ for every $k \in IOb$ and $b, c \in Ib$.



FIGURE 2. Illustration of $cen_k(b, c, t)$, $cen_k(b, c)$ and of axiom AxCenter

We are ready now to formalize that the relativistic mass is a quantity that shows the magnitude of the influence of the body on the state of motion of the body it collides with.

AxCenter: If inertial bodies b and c collide inelastically originating single inertial body d, then the world-line of d is the continuation of the center-line of mass of b and c (see Fig.2):

$$\forall k \in \text{IOb } \forall b, c, d \in \text{Ib} \quad [inecoll_k(b, c : d) \implies cen_k(b, c) \cup wl_k(d) \subseteq \ell \text{ for some line } \ell].$$

The main axiom of SpecRelDyn is AxCenter which, in some sense, can be taken as the definition of relativistic mass. The remaining axioms of our axiom system will be simplifying or book-keeping axioms to make life simpler.

AxCenter is an axiom in Newtonian Dynamics, too, where the mass $m_k(b)$ of a body b is observer-independent in the sense that it does not depend on the observer k. However, in special relativity, AxCenter implies that the mass of a body necessarily depends on the observer. The reason for this fact is that the simultaneities of the different observers in special relativity differ from each other, and this implies that the proportions involved in AxCenter change, too. See Prop.4.1 and Fig.3 below.

Proposition 4.1. Assume SpecRel and AxCenter. Let $k, h \in IOb$, $b, c, d \in Ib$ be such that $inecoll_k(b, c : d)$, $inecoll_h(b, c : d)$ and h is not at rest w.r.t. k. Then

$$\frac{m_k(b)}{m_k(c)} \neq \frac{m_h(b)}{m_h(c)}.$$



FIGURE 3. Illustration for Prop.4.1. The proportion of the bold and dotted segments on the horizontal line is different from that on the slanted one.

We omit the proof of Prop.4.1, but Fig.3 is an illustration for it.

The velocity $\vec{v}_k(b)$ and speed $v_k(b)$ of body b according to observer k are defined as:

$$\vec{v}_k(b) := \frac{p_\sigma - q_\sigma}{p_\tau - q_\tau}, \text{ for } p, q \in wl_k(b) \text{ with } p_\tau \neq q_\tau, \text{ and } \vec{v}_k(b) := |\vec{v}_k(b)|$$

if $wl_k(b)$ is a subset of a line and contains coordinate-points p and q with $p_{\tau} \neq q_{\tau}$, and they are undefined otherwise.

The **rest mass** $m_0(b)$ of body b is defined to be $\lambda \in \mathbb{Q}$ if there is an observer according to which b is at rest and the relativistic mass of b is λ , and for every observer according to which b is at rest the relativistic mass of b is λ , that is, $m_0(b) = \lambda$ iff

$$\exists k \in \text{IOb} (v_k(b) = 0 \land m_k(b) = \lambda) \land \forall k \in \text{IOb} (v_k(b) = 0 \Longrightarrow m_k(b) = \lambda)$$

By Rmk.3.1, assuming $d \geq 3$, AxSelf, AxEv and AxPh, if the rest mass of body b is defined then b is slower than light, that is, $v_k(b)$ is defined and $v_k(b) < 1$ for every observer k. In particular, photons do not have rest masses, but see Remark 4.4(2) later.

CONVENTION 4.2. We use the equation sign "=" in the sense of existential equality (of partial algebra theory [2]), that is, $\alpha = \beta$ abbreviates that both α and β are defined and they are equal. See [16, Conv.2.3.10, p.31] and [4, Conv.2.3.10, p.61].

We have seen that AxCenter implies that the relativistic mass $m_k(b)$ has to depend on both b and k. The next axiom states that the relativistic mass of a body depends at most on its rest mass and its velocity.

AxSpeed: The relativistic masses of two inertial bodies are the same if both of their rest masses and speeds are equal:

 $\forall k \in \text{IOb } \forall b, c \in \text{Ib}$

$$\left[\left(m_0(b) = m_0(c) \land v_k(b) = v_k(c) \right) \implies m_k(b) = m_k(c) \right].$$

Our last axiom on dynamics states that each observer can make experiments in which he makes inertial bodies of arbitrary rest masses and velocities inelastically collide:

Ax∀inecoll : For every observer, every kind of possible inelastic collision is realized by inertial bodies having rest mass:

 $\forall k \in \text{IOb } \forall v_1, v_2 \in \mathbf{Q}^{d-1} \ \forall m_1, m_2 \in \mathbf{Q} \quad (|v_1| < 1 \land |v_2| < 1 \\ \land m_1 > 0 \land m_2 > 0 \implies \exists b, c, d \in \text{Ib } [inecoll_k(b, c : d) \\ \land \ \vec{v}_k(b) = v_1 \land \ \vec{v}_k(c) = v_2 \land \ m_0(b) = m_1 \land \ m_0(c) = m_2]).$

Let us extend SpecRel with the axioms of dynamics above.

 $SpecRelDyn := {AxCenter, AxSpeed, Ax\forall inecoll, AxThEx} \cup SpecRel$

We note that **SpecRelDyn** is provably consistent. Moreover it has non-trivial models, see Prop.5.6.

The following theorem gives the connection between the rest mass and the relativistic mass of an inertial body. Its conclusion is a well known result of special relativity. We will see that our theorem is stronger than the corresponding result in the literature since it contains fewer assumptions.

Theorem 4.3. Assume $d \ge 3$ and SpecRelDyn. Let k be an observer and b be an inertial body having rest mass. Then

$$m_0(b) = \sqrt{1 - v_k(b)^2} \cdot m_k(b).$$

Proof. Let k be an observer and let a be an inertial body having rest mass. Let $v := v_k(a)$, $m_0 := m_0(a)$ and $m(v) := m_k(a)$. We would like to prove that $m_0 = \sqrt{1 - v^2} \cdot m(v)$. It holds if v = 0 by the definition of rest mass. Now assume that $v \neq 0$. We are in the world-view of observer k. Let inertial bodies b and c collide inelastically originating inertial body d such that the rest masses of b and c are m_0 the speed of b is v and the speed of c is 0. See Fig.4. Such b, c and d exist by $Ax \forall inecoll$. There are distinct points B and C on the world-lines of b and c, respectively, such that $B_{\tau} = C_{\tau}$. Let such B and C be fixed and let $t := B_{\tau} = C_{\tau}$. Let D be the center of mass of b and c at t. The relativistic masses of b and c according to k are m(v)and m_0 , respectively, by AxSpeed and the definition of rest mass. Let |pq| := |p-q|. By definition of center of mass, $m(v) \cdot |BD| = m_0 \cdot |DC|$. Thus

$$m_0 = \frac{|BD|}{|CD|} \cdot m(v). \tag{3}$$



FIGURE 4. Illustration for the proof of Thm.4.3

Let A be the point where the world-lines of b, c and d meet. By AxCenter, $cen_k(b,c) \cup wl_k(d) \subseteq AD$. Let k' be an observer such that $v_{k'}(b) = 0$. Such a k' exists since b has rest mass. We can assume that the clocks of k and k' show 0 at A, that is, $A_{\tau} = w_{k'}^k(A)_{\tau} = 0$, and the clock of k shows -1 at C, that is, $C_{\tau} = -1$. By applying the "time-dilation theorem" of SpecRel (see [5, Thm.11.6.(2)]) we get that the clock of k' shows $-\sqrt{1-v^2}$ or $\sqrt{1-v^2}$ at B. We can assume that the clock of k' shows $-\sqrt{1-v^2}$ at B.

By AxThEx there is an observer h for which b and c have opposite velocities and $inecoll_h(b, c : d)$. Let such an h be fixed.

The world-view transformation w_h^k between the world-views of k and h is an affine transformation, that is, a linear transformation composed by a translation by [5, Thm.11.10.]. Thus w_h^k takes lines to lines.

Let us turn our attention to the world-view of h. See the right-hand side of Fig.4. Let A', B', C' and D' be the w_h^k images of A, B, C and D, respectively. Since w_h^k is an affine transformation,

$$\frac{|BD|}{|CD|} = \frac{|B'D'|}{|C'D'|}.$$
(4)

We will prove that

$$\frac{|B'D'|}{|C'D'|} = \frac{|A'B'|}{|A'C'|}.$$
(5)

Let ℓ be the line parallel to the time-axis \bar{t} and passing through A'. Since the rest masses and the speeds of b and c coincide, their relativistic masses coincide by AxSpeed. Therefore $cen_h(b,c) \subseteq \ell$. By AxCenter, $wl_h(d) \subseteq \ell$. The world-view transformation takes lines to lines and world-lines to world-lines. Thus w_h^k takes $wl_k(d) \subseteq AD$ to $wl_h(d) \subseteq \ell$. Therefore D' is the intersection of ℓ and B'C'.

Let $E' \in A'D'$ be such that E'C' is parallel to A'B'. The triangles B'D'A' and C'D'E' are similar. Thus

$$\frac{|B'D'|}{|C'D'|} = \frac{|A'B'|}{|E'C'|}.$$
(6)

Since b and c have opposite speeds and A'B' is parallel with C'E', angles E'A'C' and A'E'C' are congruent. Thus |E'C'| = |A'C'|. By this and (6), we conclude that (5) above holds.

The clocks of k' and k show 0 at A', the clock of k' shows $-\sqrt{1-v^2}$ at B' and the clock of k shows -1 at C'. The speeds of k and k' coincide for h. Thus the clocks of k and k' slow down with the same rate for h by [5, Thm.11.6.(2)]. Therefore

$$\frac{|A'B'|}{|A'C'|} = \sqrt{1 - v^2}.$$
(7)

By (3), (4), (5) and (7), we get that

$$m_0 = \sqrt{1 - v^2} \cdot m(v);$$

and that is what we wanted to prove.

Remark 4.4. (1) The conclusion of Thm.4.3 fails if we omit any one of the axioms AxCenter, AxSpeed, Ax \forall inecoll, AxThEx from SpecRelDyn. However, it remains true if we omit AxSimDist and weaken Ax \forall inecoll and AxThEx to the following two axioms, respectively:

Ax∃inecoll: According to every observer, for every inertial body a having rest mass there are inertial bodies b and c colliding inelastically originating an inertial body such that a, b and c have the same rest masses, a and b have the same speeds and the speed of c is 0:

$$\forall k \in \text{IOb} \ \forall a \in \text{Ib} \ \exists b, c, d \in \text{Ib} \ (m_0(a) = m_0(a) \implies [m_0(a) = m_0(b) = m_0(c) \ \land v_k(b) = v_k(a) \ \land \ v_k(c) = 0 \ \land \text{inecoll}_k(b, c : d)]).$$

AxMedian : For every two inertial bodies having rest masses and colliding inelastically, there is an observer for which these two inertial bodies have opposite velocities and collide inelastically:

$$\forall k \in \text{IOb } \forall b, c, d \in \text{Ib} \quad \left[\text{inecoll}_k(b, c : d) \implies \\ \exists h \in \text{IOb} \quad \left(\vec{v}_h(b) = -\vec{v}_h(c) \land \text{inecoll}_h(b, c : d) \right) \right]$$

(2) According to our definition, photons do not have rest masses because no observer sees them at rest, by AxPh. However, they do have nonzero relativistic masses, by AxFrame. In the light of Thm.14.3 then it is natural to extend the rest mass concept for photons as $m_0(ph) = 0$ for all $ph \in$ Ph. This is often done in the physics literature. In the light of Einstein's $E = mc^2$, one could say that "photons are pure energy," because they have nonzero relativistic masses, but zero rest masses.



FIGURE 5. Illustration of axioms Ax∃inecoll and AxMedian

On Einstein's $E = mc^2$: The conclusion $m_0(b) = \sqrt{1 - v_k(b)^2} \cdot m_k(b)$ of our Thm.4.3 above is used in the relativity textbook Rindler [19, pp.111-114] to explain the discovery and meaning of Einstein's famous insight $E = mc^2$. We could repeat literally this part of the text of [19] to arrive at $E = mc^2$ in the framework of our theory SpecRelDyn based on the axiom AxCenter. We postpone this to section 5, because there we will have developed more "ammunition," hence the didactics can be made more inspiring.

5. Conservation of relativistic mass and linear-momentum

We can view AxCenter as stating that the center of mass of an isolated system consisting of two bodies moves along a straight line regardless whether the two bodies collide or not. It is natural to generalize AxCenter to more than two bodies (but permitting only two-by-two inelastic collisions). Let AxCenter_n denote, temporarily, a version of AxCenter which concerns an isolated system consisting of *n* bodies. Thus AxCenter is just AxCenter₂ in this series of stronger and stronger axioms. We will see that it does not imply AxCenter₃ (cf. Prop.5.4), thus AxCenter₃ is strictly stronger than AxCenter. However, it can be shown (see [6]) that the rest of the axioms in this series are all equivalent to AxCenter₃. This motivates our introducing SpecRelDyn⁺ by replacing AxCenter in SpecRelDyn with the stronger AxCenter₃. The theory $SpecRelDyn^+$ is still very geometric and observation-oriented in spirit.

We are going to introduce $AxCenter_3$, we will denote it as $AxCenter^+$. The center-line of mass $cen_k(a, b, c)$ of three bodies a, b and c according to observer k is defined in a completely analogous way as for two bodies, as follows. The **center of mass** $cen_k(a, b, c, t)$ of bodies a, b and caccording to observer k at time instance t is defined as:

$$m_k(a) \cdot (cen_k(a, b, c, t) - loc_k(a, t)) + m_k(b) \cdot (cen_k(a, b, c, t) - loc_k(b, t)) + m_k(c) \cdot (cen_k(a, b, c, t) - loc_k(c, t)) = 0$$

if $loc_k(a, t)$, $loc_k(b, t)$ and $loc_k(c, t)$ are defined and it is undefined otherwise. We note that an explicit definition for $cen_k(a, b, c, t)$ is the following:

$$cen_k(a, b, c, t) = \frac{m_k(a)}{m_k(a) + m_k(b) + m_k(c)} \cdot loc_k(a, t) + \frac{m_k(b)}{m_k(a) + m_k(b) + m_k(c)} \cdot loc_k(b, t) + \frac{m_k(c)}{m_k(a) + m_k(b) + m_k(c)} \cdot loc_k(c, t)$$

The **center-line of mass** of bodies a, b and c according to observer k is defined as

$$cen_k(a,b,c) := \{cen_k(a,b,c,t) : t \in \mathbb{Q} \text{ and } cen_k(a,b,c,t) \text{ is defined} \}.$$

AxCenter⁺: If a is an inertial body and inertial bodies b and c collide inelastically originating inertial body d, then the centerline of a and d is the continuation of the center-line of a, b and c, i.e. there is a line that contains both the center-line of a, b and c and the center-line of a and d (see Fig.6):

$$\forall k \in \text{IOb } \forall a, b, c, d \in \text{Ib} \quad [inecoll_k(b, c : d) \implies cen_k(a, b, c) \cup cen_k(a, d) \subseteq \ell \text{ for some line } \ell].$$

Let us replace AxCenter with AxCenter⁺ in SpecRelDyn:

$SpecRelDvn^+ :=$	{AxCenter ⁺	. AxSpeed. Ax∀ined	coll. AxThEx	$\cup SpecRel$
			-)	

We note that $SpecRelDyn^+$ is consistent. Moreover it has non-trivial models, see Prop.5.6.

CONVENTION 5.1. Throughout the paper, there appear "highlighted" statements like AxCenter⁺ above which associate a name like AxCenter⁺ to a formula of our first-order logic language (for SpecRelDyn). It is important to note that these formulas are not automatically elevated to the rank of an axiom. Instead, they serve as potential axioms or even as potential statements to appear in theorems, hence they are nothing but distinguished formulas of our language.



FIGURE 6. Illustration of AxCenter⁺

AxCenter determines the velocity of the body emerging from an inelastic collision, and we will see that $AxCenter^+$ determines also the relativistic mass of the body emerging from the collision.

ConsMass: Conservation of relativistic mass:

$$\forall k \in \text{IOb} \ \forall b, c, d \in \text{Ib} \quad [inecoll_k(b, c : d) \implies$$

 $m_k(b) + m_k(c) = m_k(d)].$

The **linear-momentum** of body *b* according to observer *k* is defined to be $m_k(b) \cdot \vec{v}_k(b)$ if $\vec{v}_k(b)$ is defined, and it is undefined otherwise.

ConsMoment: Conservation of linear-momentum:

$$\forall k \in \text{IOb } \forall b, c, d \in \text{Ib} \quad [inecoll_k(b, c : d) \implies \\ m_k(b) \cdot \vec{v}_k(b) + m_k(c) \cdot \vec{v}_k(c) = m_k(d) \cdot \vec{v}_k(d)].$$

The following theorem states that $AxCenter^+$ is equivalent to the conjunction of any two of the formulas AxCenter, ConsMass, ConsMoment, but it is strictly stronger than any one of them (Prop.5.4). This means, in some sense, that ConsMass represents the "difference" between AxCenter and $AxCenter^+$, and the same holds for ConsMoment.

Theorem 5.2. Assume AxSelf. Items (i)–(iv) below are equivalent

- (i) AxCenter⁺.
- (ii) ConsMass \land ConsMoment.
- (iii) ConsMass \land AxCenter.
- (iv) ConsMoment \land AxCenter.

The proof of Thm.5.2 is in [6].

Corollary 5.3. Assume SpecRelDyn⁺. Let $k \in IOb, b, c, d \in Ib$ and assume *inecoll*_k(b, c : d). Then

$$m_k(d) = m_k(b) + m_k(c), \quad \text{but}$$

$$m_0(d) > m_0(b) + m_0(c), \quad \text{whenever } m_0(b), m_0(c) \text{ exist and } \vec{v}_k(b) \neq \vec{v}_k(c)$$

The proof is in [6], but for the idea of the proof see below.

Returning to $E = mc^2$: Cor.5.3 above can be used for arriving at Einstein's insight $E = mc^2$ analogously to how it is done in the relativity textbooks Rindler [19] and d'Inverno [9]. Namely, we have seen above, in Cor.5.3, that under appropriate arrangement, rest mass can be created. Created from what? Well, from kinetic energy (energy of motion). This points in the direction of Einstein's connecting mass with energy. In more detail, let us start with two bodies b_1, b_2 of rest mass m_0 . Let us accelerate the two bodies towards each other and let them collide inelastically, so that they stick together forming the new body " $b_1 + b_2$ " (deliberately sloppy notation). Assume $b_1 + b_2$ is at rest relative to the observer conducting the experiment. Then the rest mass $m_0(b_1+b_2)$ is the sum of relativistic masses $m_k(b_1)$ and $m_k(b_2)$ by Cor.5.3. Assuming that at collision the speed of both b_1 and b_2 was v, we have $m_0(b_1 + b_2) = m_0(b_1)/\sqrt{1 - v^2} + m_0(b_2)/\sqrt{1 - v^2}$, by Thm.4.3, which is definitely bigger than $m_0(b_1) + m_0(b_2)$ if $v \neq 0$. So, rest mass was created from the kinetic energy supplied to our test bodies b_1, b_2 when we accelerated them towards each other. So far, we have a qualitative argument (based on our SpecRelDyn⁺) in the direction that energy (in our example kinetic) can be "transformed" to "create" mass. A quantitatively (and physically) more detailed analysis of $E = mc^2$ in terms of Thm.4.3 is given in [19, pp.111-114] to where we refer the reader for more detail and for the "second part" of the argument. The "first part" was provided by Thm.4.3 and Cor.5.3.

Let φ be a formula and Σ be a set of formulas. $\Sigma \models \varphi$ denotes that φ is true in all models of Σ (i.e. φ is a logical consequence of Σ). $\Sigma \not\models \varphi$ denotes that there is a model of Σ in which φ is not true.

Proposition 5.4.

The proof of Prop.5.4 is in [6].

In the literature, the conservation of relativistic mass and that of linear-momentum are used to derive the conclusion of Thm.4.3. By Prop.5.4 above, our axiom system SpecRelDyn implies neither ConsMass nor ConsMoment. By Thm.5.2, ConsMass and ConsMoment together imply the key axiom AxCenter of SpecRelDyn. So Thm.4.3 is stronger

than the corresponding result in the literature since it requires fewer assumptions.

Thm.5.2 also states that the conservation axioms can be replaced by the natural, purely geometrical symmetry postulate $AxCenter^+$ without loss of predictive or expressive power. Since the conservation axioms **ConsMass** and **ConsMoment** are not "purely geometrical" and they are less observation-oriented than $AxCenter^+$, we feel that it may be more convincing to use AxCenter or $AxCenter^+$ in an axiom system when we introduce the basics of relativistic dynamics.

Let $k \in IOb$ and $b \in Ib$. The **four-momentum** $P_k(b)$ of inertial body b according to observer k is defined to be the element of Q^d whose time component and space component are the relativistic mass and linear-momentum of body b according to observer k, respectively, see Fig.7. That is,

$$P_k(b)_{\tau} = m_k(b)$$
 and $P_k(b)_{\sigma} = m_k(b) \cdot \vec{v}_k(b)$.

It is not difficult to see, using Thm.4.3, that $P_k(b)$ is parallel to the world-line of b and its Minkowski-length is $m_0(b)$.



FIGURE 7. Illustration of four-momentum $P_k(b)$

ConsFourMoment : Conservation of four-momentum:

$$\forall k \in \text{IOb } \forall b, c, d \in \text{Ib} \quad [inecoll_k(b, c : d) \implies P_k(b) + P_k(c) = P_k(d)]$$

The following is an immediate corollary of Thm.5.2.

Corollary 5.5. AxSelf \vdash (AxCenter⁺ \iff ConsFourMoment).

Let us return to discussing the merits of using AxCenter⁺ in place of the more conventional preservation principles. In the context of Cor.5.5, ConsFourMoment has the advantage that it is computationally direct and simple, while AxCenter⁺ has the advantage that it is more observational, more geometrical, and more basic in some intuitive sense. Let us finally state a theorem about the existence of nontrivial models of our axiom systems.

Proposition 5.6. SpecRelDyn⁺ \cup {IOb $\neq \emptyset$ } is consistent.

The proof of Prop. 5.6 is in [6].

A related work with somewhat different aims is [20].

6. Concluding Remarks

We have introduced a purely geometrical axiom system of special relativistic dynamics which is strong enough to prove the formula connecting relativistic and rest masses of bodies. We have also studied the connection of our key axioms AxCenter and AxCenter⁺ and the usual axioms about the conservation of mass, momentum and fourmomentum. Connections with Einstein's $E = mc^2$ were also discussed. The contents of the present paper represent only the first steps towards a logical conceptual analysis of relativistic dynamics or mechanics. A glimpse to Chap.6 (pp.108-130) "Relativistic particle mechanics" of the textbook Rindler [19] suggests the topics to be covered in future work in this line. In a direction orthogonal to this, looking at the logical issues in [4] and [5] suggests questions and investigations to be carried out into the logical analysis of relativistic dynamics.

In this paper we began axiomatizing dynamics in special relativity. This axiomatization of dynamics is extended to the theory of accelerated observers AccRel in [24]. (For the FOL theory AccRel we refer to [17].) In a similar spirit, these ideas can be naturally extended to the FOL theory GenRel of general relativity (see e.g. [5]).

AxPh reveals that (in our present axiom systems) we think of photons as "possible bodies", and the real meaning of AxPh is that "it is possible for a photon to move from p to q iff ...". The situation is similar with axioms AxThEx, Ax \forall inecoll. So, a notion of possibility plays a role here. In the present paper we work in an extensional framework, as is customary in geometry and in spacetime theory. It would be more natural to treat this "possibility phenomenon" in a modal logic framework, and this is more emphatically so in dynamics. It would be most interesting to explore the use of a modal logic framework in our logical analysis of relativity theory.

Acknowledgements. Thanks go to Zalán Gyenis, Leon Horsten, Thomas Mueller, Adrian Sfarti and Renata Tordai for helpful and fruitful discussions, suggestions and remarks.

References

[1] Logic and philosophy of mathematics. Homepage of a seminar in Department of Logic, Institute of Philosophy, Eötvös University Budapest. http://phil.elte.hu/cgi-bin/mailman/listinfo/lpom.

- [2] H. Andréka, P. Burmeister, and I. Németi. Quasivarieties of partial algebras a unifying approach towards a two-valued model theory for partial algebras. *Studia Sci. Math. Hungar.*, 16:325–372, 1981.
- [3] H. Andréka, J.X. Madarász, and Németi I. Logical axiomatizations of spacetime; samples from the literature. In A. Prékopa and E. Molnár, editors, *Non-Euclidean Geometries: János Bolyai Memorial Volume*, pages 155–185. Springer, 2006. http://www.math-inst.hu/pub/algebraic-logic/lstsamples.ps.
- [4] H. Andréka, J.X. Madarász, and I. Németi. On the logical structure of relativity theories. Research report, Alfréd Rényi Institute of Mathematics, Hungar. Acad. Sci., Budapest, 2001. With contributions from A. Andai, G. Sági, I. Sain and Cs. Tőke. http://www.math-inst.hu/pub/algebraic-logic/Contents.html. 1200 pp.
- H. Andréka, J.X. Madarász, and I. Németi. Logic of space-time and relativity. In *Handbook of Spatial Logics*, pages 607–711. Springer, 2007. http://ftp.mathinst.hu/pub/algebraic-logic/Logicofspacetime.pdf.
- [6] H. Andréka, J.X. Madarász, I. Németi, and G. Székely. E=mc² derived from geometrical axioms. Research report, Alfréd Rényi Institute of Mathematics, Hungar. Acad. Sci., Budapest, 2007.
- [7] J. Ax. The elementary foundations of spacetime. Found. Phys., 8(7-8):507-546, 1978.
- [8] C.C. Chang and H.J. Keisler. *Model Theory*. North–Holland, Amsterdam, 1973, 1990.
- [9] R. d'Inverno. Introducing Einstein's Relativity. Clarendon, Oxford, 1992.
- [10] J. Ferreirós. The road to modern logic an interpretation. B. Symb. Log., 7:441–484, 2001.
- [11] H. Friedman. On foundational thinking 1. Posting in FOM (Foundations of Mathematics) Archives www.cs.nyu.edu (January 20, 2004).
- [12] H. Friedman. On foundations of special relativistic kinematics 1. Posting No 206 in FOM (Foundations of Mathematics) Archives www.cs.nyu.edu (January 21, 2004).
- [13] L. Henkin, P. Suppes, and A. Tarski, editors. The axiomatic method with special reference to geometry and physics, Studies in Logic and the Foundations of Mathematics. North–Holland, Amsterdam, 1959.
- [14] D. Hilbert. Mathematische Behandlung der Axiome der Physik. Akad. Wiss. Göttingen, pages 272–273, 1900. http://www.mathematik.unibielefeld.de/ kersten/hilbert/prob6.html.
- [15] D. Hilbert. Über den Satz von der Gleichheit der Basiswinkel im gleichschenkligen Dreieck. Proc. London Math. Soc., 35:50–68, 1902/1903.
- [16] J.X. Madarász. Logic and relativity (in the light of definability theory). PhD thesis, Eötvös Loránd Univ., Budapest, 2002. http://www.mathinst.hu/pub/algebraic-logic/Contents.html.
- [17] J.X. Madarász, I. Németi, and G. Székely. Twin paradox and the logical foundation of relativity theory. *Foundations of Physics*, 36(5):681–714, 2006. arXiv:gr-qc/0504118.
- [18] V. Pambuccian. Axiomatizations of hyperbolic and absolute geometries. In A. Prékopa and E. Molnár, editors, Non-Euclidean Geometries: János Bolyai Memorial Volume, pages 119–153. Springer, 2006.
- [19] W Rindler. Relativity: Special, General, and Cosmological. Oxford University Press, 2001.
- [20] B. Rothenstein and S. Popescu. Relativistic dynamics without conservation laws. arXiv:physics/0605203v1, May 2006.

- [21] S. G. Simpson, editor. *Reverse Mathematics 2001*. Lecture Notes in Logic. Association for Symbolic Logic, 2005.
- [22] P. Suppes. Axioms for relativistic kinematics with or without parity. In Henkin et al. [13].
- [23] P. Suppes. The desirability of formalization in science. J. philos., 65:651–664, 1968.
- [24] G. Székely. First-order Logic Investigation of Special Relativity extended to Accelerated Observers. PhD thesis, Eötvös Loránd Univ., Budapest, 2008. in preparation.
- [25] J. Väänänen. Second-order logic and foundations of mathematics. B. Symb. Log., 7:504–520, 2001.
- [26] J. Woleński. First-order logic: (philosophical) pro and contra. In Vincent F. Hendrics et al., editor, *First-Order Logic Revisited*, pages 369–398. Logos Verlag, Berlin, 2004.

Alfréd Rényi Institute of Mathematics of the Hungarian Academy of Sciences Budapest P.O.Box 127, H-1364 Hungary andreka@renyi.hu, madarasz@renyi.hu, nemeti@renyi.hu, turms@renyi.hu.