Article

On the Significance of the Stress–Energy Tensor in Finsler Spacetimes

Miguel Ángel Javaloyes [1], Miguel Sánchez [2,*] and Fidel F. Villaseñor [2]

[1] Departamento de Matemáticas, Campus de Espinardo, Universidad de Murcia, 30100 Murcia, Spain; majava@um.es
[2] Departamento de Geometría y Topología, Facultad de Ciencias & IMAG,
Campus de Fuentenueva s/n, Universidad de Granada, 18071 Granada, Spain; fidelfv@ugr.es
[*] Correspondence: sanchezm@ugr.es

Abstract: We revisit the physical arguments that led to the definition of the stress–energy tensor T in the Lorentz–Finsler setting (M, L) starting with classical relativity. Both the standard heuristic approach using fluids and the Lagrangian one are taken into account. In particular, we argue that the Finslerian breaking of Lorentz symmetry makes T an anisotropic 2-tensor (i.e., a tensor for each L-timelike direction), in contrast with the energy-momentum vectors defined on M. Such a tensor is compared with different ones obtained by using a Lagrangian approach. The notion of divergence is revised from a geometric viewpoint, and, then, the conservation laws of T for each observer field are revisited. We introduce a natural *anisotropic Lie bracket derivation*, which leads to a divergence obtained from the volume element and the non-linear connection associated with L alone. The computation of this divergence selects the Chern anisotropic connection, thus giving a geometric interpretation to previous choices in the literature.

Keywords: divergence in Finsler manifolds; stress–energy tensor; Finsler spacetime; Lorentz symmetry breaking; very special relativity

1. Introduction

This article has a double aim in Lorentz–Finsler geometry. The first one is to revisit the physical grounds of the stress–energy tensor T, Section 3. The possible extensions of the relativistic T are discussed from the viewpoint of both fluids mechanics and Lagrangian systems. The second one is to revise geometrically the notion of divergence, Section 4, yielding consequences about the conservation of T, Section 5. With this aim, we introduce new notions of the Lie bracket and the derivative associated with a nonlinear connection and applicable to anisotropic tensors fields, which appear naturally in Finsler geometry.

Finslerian modifications of General Relativity aim to find a tensor T collecting the possible anisotropies in the distribution of energy, momentum, and stress, which will serve as a source for the (now Lorentz–Finsler) geometry of the spacetime [1–5]. Some of these proposals may be waiting for experimental evidence, postponing then how the basic relativistic notions would be affected. However, such a discussion is relevant to understand the scope and implications of the introduced Finslerian elements. In a previous reference [6], the fundamentals of observers in the Finslerian setting were extensively studied, including its compatibility with the Ehlers–Pirani–Schild approach. Now we focus on the stress–energy tensor T.

The difficulty to study such a T is apparent. Recall that, using the principle of equivalence, General Relativity is reduced infinitesimally into the Special one, which provides a background for interpretations. However, in the Lorentz–Finsler case, the infinitesimal model is changed into a Lorentz norm (instead of scalar product), implying a breaking of Lorentz invariance. This is a substantial issue in its own right which has been studied

in the context of *Very Special Relativity* and others [7–11]. As an additional difficulty, the infinitesimal model changes with the point[1].

Two noticeable pre-requisites are the following: (a) only the value of the Lorentz–Finsler metric on causal directions is relevant [6,14] (this is briefly commented in the setup Section 2.3), and (b) there is a significant variety of possible extensions of the relativistic kinematic objects to the Finsler case, at least from the geometric viewpoint (see Appendix A). Taking into account these issues, the extension of the notion of the stress–energy tensor to the Finslerian setting is discussed in Section 3.

We start at the fluids approach. As a preliminary question, energy-momentum is discussed, in Section 3.1. We emphasize that, even though this is well-defined as a tangent vector in each tangent space $T_p M$, $p \in M$, different observers u, u' at p will use coordinates related by non-trivial linear transformations. Indeed, the latter will depend on both L and the chosen method to measure relative velocities. Moreover, when the stress–energy T is considered in Section 3.2, the arguments in Classical Mechanics and Relativity, which support its status as a tensor, hold only partially in the Lorentz–Finsler setting. Indeed, T acquires a nonlinear nature that is codified in an (observer-dependent) anisotropic tensor, rather than in a tensor on M.

The Lagrangian approach is discussed in Section 3.3. This approach has been developed recently by Hohmann, Pfeifer, and Voicu [15,16], who introduced an *energy-momentum scalar function*. Here, we discuss the analogies and differences of this function with the canonical relativistic stress–energy tensor $\delta S_{matter}/\delta g^{\mu\nu}$ and the 2-tensor T obtained from the fluids approach above. Relevant issues are the existence of different methods to obtain a 2-tensor starting at a scalar function, the recovery of this function from a matter Lagrangian, and the possibility to consider the Palatini Lagrangian as the background one (rather than Einstein–Hilbert-type Lagrangians used by the cited authors; recall that Palatini's becomes especially meaningful in the Finslerian case [17]). The important case of kinetic gases is considered explicitly (Example 2).

Once the definition of T has been discussed, we focus on its conservation, Section 5, revisiting first the divergence theorem, Section 4. This is crucial in the Finslerian setting because, as discussed before, the Lagrangian approach above does not guarantee a conservation law as the relativistic $\text{div}(G) = 0$.

Section 4 analyzes the divergence from a purely mathematical viewpoint. Now, L is regarded as pseudo-Finsler (the results will be useful not only in any indefinite signature but also in the classical positive definite case), and T will not be assumed to be symmetric a priori. Classically, the divergence of a vector field Z is defined with the derivation associated with the Lie bracket $[Z, X] = \mathfrak{L}_Z X$, applied to the volume element. In the Finslerian case, however, the Lie derivative and bracket do not make sense for arbitrary anisotropic vector fields. This difficulty was circumvented by Rund [18], who redefined $\text{div}(Z)$ in such a way that a type of divergence theorem held. However, the Lie viewpoint is restored here.

Section 4.1 Once a nonlinear connection HA (seen as a horizontal distribution on A) is prescribed, we can define a Lie bracket $l_Z^H X$ and, then, a Lie derivative $\mathfrak{L}_Z^H X$ (Definitions 1 and 2; Theorem 1 (C)). Noticeably, the former l_Z^H is expressible in terms of the infinitesimal flow of Z (Proposition 1).

Section 4.2 The divergence of Z is naturally defined by using this Lie bracket (Definition 3). For the computation of $\text{div}(Z)$, however, one can use an *anisotropic connection* ∇ (this can be seen as a Finsler connection dropping its vertical part, see Section 2) and a priori Chern's one is not especially privileged (Proposition 2).

Section 4.3. We give a general Finslerian version of the divergence theorem for any anisotropic vector field Z, emphasizing the role of the choice of an (admissible) vector field $V \colon M \to A$, which in the Lorentzian case can be interpreted as an observer field; this is expressed in terms of integration of forms in the spirit of Cartan's formula (Theorem 2, Remark 5). We also explain how the boundary term can be expressed in different ways

by using a normal either with respect to the pseudo-Riemannian metric g_V or to the fundamental tensor, which were the choices of Rund [18] and Minguzzi [19] resp.

Section 5 gives some applications to conservation laws.

Section 5.1. First, we discuss the definition of divergence for the case of T. Our definition for vector fields was not biased to the Chern anisotropic connection, but this will be used for $\mathrm{div}(T)$ (Definition 4). The reason is that $\mathrm{div}(T)$ should behave under contraction in a similar way as in the isotropic case (namely, as in Formula (11)), which privileges Chern's connection (Proposition 3).

Section 5.2. As an interlude about the appearance of Chern's ∇, a comparison with the possible use of Berwald's and previous approaches in the literature is done.

Section 5.3. A conservation law for the flow of $T_V(X_V)$ is obtained (Corollary 2), stressing three hypotheses on the vanishing for V of elements related to the stress–energy T ($\mathrm{div}(T) = 0$), the anisotropic vector X ($\mathsf{l}_X^{\mathrm{H}} g = 0$, generalizing the isotropic case) and a derivative of V. The latter hypothesis is genuinely Finslerian, and it means that some terms related to the nonlinear covariant derivative $\mathrm{D}V$ must vanish globally (V can always be chosen such that they vanish at some point). It is worth pointing out that our general formula for the integral of the divergence (36) recovers the classical interpretation of the divergence as an infinitesimal growth of the flow (now observer-dependent). So, $\mathrm{div}(T) = 0$ is equivalent to the conservation of energy-momentum in the instantaneous rest-space of each observer—see Remark 10.

We finish by applying this general result to two examples. First, we apply it to Lorentz norms, showing that the conservation laws of Special Relativity still hold even though, now, the conserved quantity may be different for different observers. As a second example, we give natural conditions so that the flow of $T_V(X_V)$ (whenever it exists as a Lebesgue integral, eventually equal to $\pm\infty$) is equal in two Cauchy hypersurfaces of a globally hyperbolic Finsler spacetime. Indeed, we refine a previous result by Minguzzi [19], who assumed that L was defined on the whole $\mathrm{T}M$ and $T_V(X_V)$ was compactly supported. We show that a combination of Rund's and Minguzzi's methods to compute the boundary terms allows one to obtain appropriate decay rates (namely, the properly Finslerian hypothesis (49)), which ensure the conservation.

2. Preliminaries and Setup

First, let us set up some notation. In all the present text, M is a connected smooth (C^∞) manifold of dimension $n \geq 2$. As in previous references [17,20], any coordinate chart $(U, (x^1, \ldots, x^n))$ of M naturally induces a chart $(\mathrm{T}U, (x^1, \ldots, x^n, y^1, \ldots, y^n))$ of $\mathrm{T}M$ defined by the fact that

$$v = y^i(v) \left. \frac{\partial}{\partial x^i} \right|_{\pi(v)}$$

for $v \in \mathrm{T}U$, where $\pi \colon \mathrm{T}M \to M$ is the canonical projection. We abbreviate

$$\frac{\partial}{\partial x^i} =: \partial_i, \qquad \frac{\partial}{\partial y^i} =: \dot{\partial}_i;$$

these are vector fields on $\mathrm{T}U$. At any rate, we will express our results in coordinate-free and geometric terms.

2.1. Anisotropic Tensors

We shall employ the framework of anisotropic tensors, following [20–22], as it is simpler than previous ones. An open subset $A \subseteq \mathrm{T}M$ with $\pi(A) = M$ is fixed; the elements $v \in A$ are called *observers*. We will denote by $\mathcal{T}_s^r(M_A)$ the space of (smooth) r-contravariant s-covariant A-anisotropic tensor fields ($r, s \in N \cup \{0\}$) and by $\mathcal{T}(M_A) := \bigoplus_{r,s} \mathcal{T}_s^r(M_A)$ the full anisotropic tensor algebra. $\mathcal{F}(A) = \mathcal{T}_0^0(M_A)$ will be the space of functions on A. This time we will also put $\mathfrak{X}(M_A) := \mathcal{T}_0^1(M_A)$ for the space of anisotropic vector fields and $\Omega_s(M_A)$ for the space of anisotropic s-forms (alternating anisotropic tensors, so that

$\Omega_1(M_A) := \mathcal{T}_1^0(M_A))$. The space $\mathcal{T}(M)$ of classical tensor fields will be seen as a subspace of $\mathcal{T}(M_A)$, formed by the *isotropic* elements, namely, those which depend only on the point $p \in M$ and not on the observer at it. In particular, $\mathfrak{X}(M) \subseteq \mathfrak{X}(M_A)$. There is a distinguished element of $\mathfrak{X}(M_A)$: the *canonical (or Liouville) anisotropic vector field*,

$$C = y^i \, \partial_i, \qquad C_v := v.$$

For an open set $U \subseteq M$; we will put $\mathfrak{X}^A(U)$ for the set of (local) *observer fields*, that is, those $V \in \mathfrak{X}(U)$ such that $V_p \in A \cap \mathrm{T}_p M$ for all $p \in U$. Given one of these and $T \in \mathcal{T}_s^r(M_A)$, their composition, denoted by $T_V \in \mathcal{T}_s^r(U)$, makes sense. Finally, for $X \in \mathfrak{X}(M_A)$, there is also a canonical derivation $\dot{\partial}_X \colon \mathcal{T}_s^r(M_A) \to \mathcal{T}_s^r(M_A)$: the *vertical derivative along X*,

$$\left(\dot{\partial}_X T\right)_v := \lim_{t \to 0} \frac{T_{v+tX_v} - T_v}{t}, \qquad \left(\dot{\partial}_X T\right)_{j_1,\ldots,j_s}^{i_1,\ldots,i_r} = X^{j_{s+1}} \dot{\partial}_{j_{s+1}} T_{j_1,\ldots,j_s}^{i_1,\ldots,i_r}.$$

2.2. Nonlinear and Anisotropic Connections

In this article, a *nonlinear connection on $A \to M$* is defined as a (horizontal) subbundle $\mathrm{H}A \subseteq \mathrm{T}A$ such that $\mathrm{T}A = \mathrm{H}A \oplus \mathrm{V}A$, where $\mathrm{V}A := \mathrm{Ker}(\mathrm{d}\pi)|_A$ is the vertical subbundle. For other options and the rudiments, see [20]. Nonlinear connections are characterized by their *nonlinear coefficients* N_j^i,

$$\mathrm{H}_v A = \mathrm{Span}\{\delta_i|_v\}, \qquad \delta_i := \frac{\delta}{\delta x^i} := \frac{\partial}{\partial x^i} - N_i^j \frac{\partial}{\partial y^j}, \tag{1}$$

and also by their *nonlinear covariant derivative* $\mathrm{D}_X \colon \mathfrak{X}^A(U) \to \mathfrak{X}(U)$,

$$\mathrm{D}_X V := X^j \left(\frac{\partial V^i}{\partial x^j} + N_j^i(V) \right) \partial_i, \tag{2}$$

for $X \in \mathfrak{X}(U)$. They also provide (at least locally) a *nonlinear parallel transport* of observers $v \in A \cap \mathrm{T}_{\gamma(0)} M$ along curves $\gamma \colon [0, t] \to M$. Namely, a map $P_t \colon A_{\gamma(0)} \to A_{\gamma(t)}$ defined as $P_t(v) = V(t)$, being V the only vector field along γ such that $V(0) = v$ and $D_{\dot{\gamma}} V = 0$ (see (Definition 12 in [20]) and the comment below).

An *A-anisotropic connection* is an operator $\nabla \colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M_A)$ satisfying the usual Koszul derivation properties—see [17,21,22]. In a chart domain U, they are characterized by their Christoffel symbols $\Gamma_{jk}^i \colon A \cap TU \to \mathbb{R}$,

$$\nabla_{\partial_j} \partial_k =: \Gamma_{jk}^i \partial_i.$$

They can be seen as vertically trivial linear connections on the vector bundle $\mathrm{V}A \to A$ (Theorem 3 in [20]). On the other hand, every anisotropic connection has an *underlying nonlinear connection*, the only one with nonlinear coefficients

$$N_j^i := \Gamma_{jk}^i y^k.$$

As a consequence, they define the *covariant derivative* $\nabla \colon \mathcal{T}_s^r(M_A) \to \mathcal{T}_{s+1}^r(M_A)$ for any anisotropic tensor:

$$\nabla_{j_{s+1}} T_{j_1,\ldots,j_s}^{i_1,\ldots,i_r} = \delta_{j_{s+1}} T_{j_1,\ldots,j_s}^{i_1,\ldots,i_r} + \sum_{\mu=1}^r \Gamma_{j_{s+1}k}^{i_\mu} T_{j_1,\ldots,j_s}^{i_1,\ldots,k,\ldots,i_r} - \sum_{\nu=1}^s \Gamma_{j_{s+1}j_\nu}^k T_{j_1,\ldots,k,\ldots,j_s}^{i_1,\ldots,i_r}.$$

2.3. Lorentz–Finsler Metrics

From now on, we will always assume that A is conic ($\lambda v \in A$ for $v \in A$ and $\lambda \in (0, \infty)$). We shall follow the definitions and conventions in [20,23]. In particular, a *Finsler spacetime (M, L)* is a (connected) manifold M endowed with a *(properly) Lorentz–Finsler*

metric $L \colon \overline{A} \subseteq \mathrm{T}M \setminus \mathbf{0} \to [0, \infty)$. L is required to be smooth, positive homogeneous, and, when restricted to each $A_p := \mathrm{T}_p M \cap A$ ($p \in M$), its vertical Hessian g is non-degenerate with signature $(+, -, \ldots, -)$; A_p must be connected and salient, and its boundary in $\mathrm{T}M \setminus \mathbf{0}$, which must be equal to $L^{-1}(0)$, is a (strong) *cone structure* $\mathcal{C}$. In particular, at each point p, L is a *Lorentz norm*. By positive homogeneity, L is determined by its *indicatrix* $L^{-1}(1)$.

Notice that the cone $\mathcal{C}$ yields a natural notion of timelike, lightlike, and spacelike tangent vectors, but L is not defined on the latter. Indeed, we are not interested in the value of L on spacelike vectors by physical reasons, which are analyzed in [6]. Roughly, only particles (massive, massless) can be measured, and, so, experimental evidences only can affect Σ and $\mathcal{C}$. Even though this also happens in classical relativity, the value of the Lorentz metric on the (future-directed) timelike vectors is enough to extend it to all the directions. Indeed, the anisotropies in Finsler spacetimes should be regarded as originated by the distribution of matter and energy in the causal directions rather than by (unobservable) spacelike anisotropies.

Even though it is the Lorentz–Finsler case that has a physical interpretation, in all other aspects, the theory carries on if L is just *pseudo-Finsler*, namely, positively 2-homogeneous with non-degenerate g on A. In fact, this is the context in which we will develop Sections 4 and 5, as they are of a more mathematical character.

The *Cartan tensor* of L is

$$C := \frac{1}{2} \dot{\partial} g, \qquad C_{ijk} = \frac{1}{2} \frac{\partial g_{ij}}{\partial y^k}.$$

It is actually symmetric, so one can define the *mean Cartan tensor* as

$$C^{\mathbf{m}}(X) := \mathrm{trace}_g \{ C(X, -, -) \}, \qquad (C^{\mathbf{m}})_j = g^{ik} C_{ijk} =: C_j, \tag{3}$$

for $X \in \mathfrak{X}(M_A)$. L has also a canonically associated connection: the *metric nonlinear connection*, $\mathrm{H}A$, of nonlinear coefficients

$$N^i_j := \gamma^i_{jk} y^k - C^i_{jk} \gamma^k_{ab} y^a y^b, \qquad \gamma^i_{jk} := \frac{1}{2} g^{ic} \left(\frac{\partial g_{cj}}{\partial x^k} + \frac{\partial g_{ck}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^c} \right). \tag{4}$$

This is the underlying nonlinear connection of several anisotropic connections. One is the (Levi–Civita)–Chern ∇, the only symmetric anisotropic connection that parallelizes g. It is the horizontal part of Chern–Rund's and Cartan's classical connections, and it has Christoffel symbols

$$\Gamma^i_{jk} := \frac{1}{2} g^{il} \left(\frac{\delta g_{lj}}{\delta x^k} + \frac{\delta g_{lk}}{\delta x^j} - \frac{\delta g_{jk}}{\delta x^l} \right), \tag{5}$$

where the δ_i are those associated with (4). Another one is the Berwald $\widehat{\nabla}$. This is the horizontal part of Berwald's and Hashiguchi's classical connections, and it has Christoffel symbols

$$\widehat{\Gamma}^i_{jk} := \frac{1}{2} g^{il} \left(\frac{\delta g_{lj}}{\delta x^k} + \frac{\delta g_{lk}}{\delta x^j} - \frac{\delta g_{jk}}{\delta x^l} \right) + \mathrm{Lan}^i_{jk}. \tag{6}$$

Here, Lan^i_{jk} are the components of a tensor metrically equivalent to the *Landsberg tensor of L*, which, among many other methods, can be defined as

$$\mathrm{Lan}_{ijk} := \frac{1}{2} g_{lm} \dot{\partial}_i \dot{\partial}_j N^l_k y^m$$

for the N^l_k of (4) (see (37 in [21])). The Landsberg tensor is actually symmetric too, so one can define the *mean Landsberg tensor of L* as

$$\mathrm{Lan}^{\mathbf{m}}(X) := \mathrm{trace}_g \{ \mathrm{Lan}(X, -, -) \}, \qquad (\mathrm{Lan}^{\mathbf{m}})_j = g^{ik} \mathrm{Lan}_{ijk} =: \mathrm{Lan}_j. \tag{7}$$

3. Basic Interpretations on the Stress–Energy Tensor *T*

Let us start with a discussion at each event $p \in M$ of a Finsler spacetime (M, L). We can consider $T_p M$ endowed with the Lorentz norm $L|_{T_p M}$. In most of this section, the discussion relies essentially on the particular case when M is a real affine n-space with associated vector space V (which plays the role of $T_p M$ in the general case) and L is a Lorentz–Finsler norm on V with indicatrix Σ and cone $\mathcal{C}$ included in V. Given $u, u' \in \Sigma$, consider the corresponding fundamental tensors g_u and $g_{u'}$ and take orthonormal bases B_u, $B_{u'}$, obtained extending u, u'. In a natural way, these bases live in $T_u V, T_{u'} V$, and they can be identified with bases in V itself. Assuming this, the change in coordinates between B_u, $B_{u'}$ is linear but not a Lorentz transformation, in general.

Extending the interpretations in relativity, $p \in M$ is an *event*; the affine simplification includes the case of Very Special Relativity [7,8,10]; $u \in \Sigma$ can be regarded as an *observer*; the tangent space to the indicatrix $T_u \Sigma$ (i.e., the subspace g_u-orthogonal to u in $T_u V \equiv V$) becomes the *rest-space* of the observer u; and B_u is an *inertial reference frame* for this observer. The *Lorentz invariance breaking* corresponds to the fact that the bases B_u and $B_{u'}$ are orthonormal for the different metrics $g_u, g_{u'}$, and, thus, the linear transformation between the coordinates of B_u and $B_{u'}$ (when regarded as elements of the same vector space $T_u V \equiv V \equiv T_{u'} V$) is not a Lorentz one. If the affine simplification is dropped, such elements (observers and rest-spaces) must be regarded as instantaneous at $p \in M$.

It is worth emphasizing that, according to the viewpoint introduced in [14] and discussed extensively in [6], the space-like directions are not physically relevant for the Lorentz–Finsler metric. However, each (instantaneous) observer does have a restspace with a Euclidean scalar product. In the case of classical relativity, Lorentz-invariance permits natural identifications between these rest-spaces, and they become consistent with the value of the scalar product on space-like directions. Certainly, a Lorentz norm L could be extended outside these directions (maintaining the Lorentz signature for its fundamental tensor), but this can be done in many different ways, and no relation with the scalar products $g_u, u \in \Sigma$ would hold.

The dropping of natural identifications associated with the Lorentz invariance implies that many notions that are unambiguously defined in classical relativity admit many different alternatives now. In the Appendix, we analyze some of them for the *relative velocity* between observers as well as other kinematical concepts. This is taken into account in the following discussion about how the Finslerian setting affects the notion of the *energy–momentum–stress* tensor.

3.1. Particles and Dusts: Anisotropic Picture of Isotropic Elements

In principle, there is no reason to modify the classical relativistic interpretation of $p = mu$ as the *(energy-) momentum vector* of a particle of *(rest) mass* $m > 0$ moving in the observer's direction $u \in \Sigma$. Moreover, if the particle moves in such a way that m is constant, it will be represented by a unit time-like curve $\gamma(\tau)$ such that $p(\tau) = m\gamma'(\tau)$ will be its instantaneous momentum at each *proper time* τ. The (covariant) derivative $p' = m\gamma''$ would be the force F acting on the particle, which is necessarily $g_{\gamma'}$-orthogonal to γ' (i.e., the force lies in the instantaneous rest-space of the particle). Then, the relativistic conservation of the momentum in the absence of external forces would retain its natural meaning, namely, if the particle represented by (m, γ) splits into two (m_1, γ_1) and (m_2, γ_2) at some τ_0 then $m\gamma'(\tau_0) = m_1 \gamma'_1(\tau_0) + m_2 \gamma'_2(\tau_0)$.

The Appendix suggests that the way how an observer u may measure the energy-momentum and conservation may be non-trivial. In particular, if one assumes that an observer u measures $m\gamma' \in T_p M$ by using a g_u-orthonormal basis B_u in general, $g_u(m\gamma', m\gamma') \neq m^2 (= L(m\gamma'))$. Moreover, as we have already commented, the coordinates for other observer u' will not transform by means of Lorentz transformation. However, as the transformation of their coordinates is still linear, and both of them will write consistently $m\gamma'(\tau_0) = m_1 \gamma'_1(\tau_0) + m_2 \gamma'_2(\tau_0)$ in their coordinates.

Particles are also the basis to model dusts, which constitute the simplest class of relativistic fluids. A dust is represented by a number-flux vector field $N = nU$, where U represents the intrinsic velocity of the particle in the dust, i.e., a comoving observer, and n is the density of the dust for each momentaneously comoving reference frame. Comparing with the case of energy momentum, N is also an intrinsic object that lives at the tangent space of each point, and U gives the privileged observer who measures n. However, the measures of n by different observers involve different measures of the volume. As explained in the Appendix, the length contraction may be fairly unrelated to the relative velocities of the observers. This implies a more complicated transformation of the coordinates by different observers. Anyway, the transformations between these coordinates would remain linear, and, so, they could still agree in the fact that they are measuring the same intrinsic vector field.

Summing up, in the case of both particles and dusts, one assumes that the physical property lives in V (or, more properly, in each tangent space $T_p M$ of the affine space), and there is a privileged (comoving) observer u. The transformation of coordinates for another observer u' may be complicated, but, at the end, it is a linear transformation that can be determined by specifying the geometric quantities that are being measured as well as the geometry of Σ. Thus, by using the coordinates measured by each observer one could construct and anisotropic vector field at each $p \in M$, which will fulfill some constraints, as the measurement by one of the observers (in particular, the privileged one) would determine the measurements by all the others.

3.2. Emergence of an Anisotropic Stress–Energy Tensor

The situation, however, is subtler for more general fluids, which are modeled classically by a 2-tensor on the underlying manifold.

Let us start by recalling the Newtonian and Lorentzian cases. In Classical Mechanics, one starts working in an orthonormal basis of Euclidean space to obtain the components T_{ij} of the Cauchy stress tensor, which give the flux of i-momentum (or force) across the j-surface in the background. [2]. The laws of conservation of linear momentum and static equilibrium of forces imply that these components give truly a 2-tensor (linear in each variable), and the conservation of linear momentum implies that this tensor is symmetric.

In the relativistic setting, each observer will determine some symmetric components T^{ij} in its rest-space by essentially the same procedure as above. Additionally, it constructs T^{00}, T^{0i}, and T^{i0} as the density energy and the energy flux across i-surface and i-momentum density, resp. The interpretation of these magnitudes completes the symmetry. [3] $T^{0i} = T^{i0}$ as well as the linearity in the 0-component. However, the bilinearity in the components $T^{\mu\nu}$ has been only ensured for vectors in the rest-space of the observer. In relativity, one can claim Lorentz invariance in order to complete the reasons justifying that, finally, the components $T^{\mu\nu}$ will transform as a tensor[4].

Nevertheless, it is not clear in Lorentz–Finsler geometry why the transformation of the components T_{ij} from an observer u to a second one u' must be linear, taking into account that they apply to space-like coordinates in distinct Euclidean subspaces and no Lorentz-invariance is assumed. Indeed, the following simple academic example shows that this is not the case.

Example 1. *Assume that (M, L) is an affine space with a Lorentz norm with domain A and consider the anisotropic tensor $\mathbf{T} = \phi\, C \otimes C$, where C is the canonical (Liouville) vector field, and $\phi : \Sigma \to R$ is a smooth function, which is extended as a 0-homogeneous function on A. Then, for each $u \in \Sigma$ and $w \in T_u\Sigma$, one has $\mathbf{T}_u(u, u) = \phi(u)$, $\mathbf{T}_u(w, w) = 0$, and $\mathbf{T}_u(u, w) = 0$. In this case, each $\mathbf{T}_u$ is a symmetric 2-tensor, but the information on $\mathbf{T}$ requires the knowledge of $\phi(u)$ for all possible $u \in \Sigma$. Recall that this example holds even if (M, L) is the Lorentz–Minkowski spacetime regarded as a Finsler spacetime (but no Lorentz-invariance is assumed for $\mathbf{T}$).*

Therefore, the following issues about T appear:

(a) Observer dependence: even if we assume that the components $T^{\mu\nu}$ measured by any observer u are bilinear, and, then, it is a standard tensor, the components measured by a second observer u' may transform by a linear map, which depends on Σ as well as the experimental method of measuring (as in the case of the energy–momentum vector).

(b) Nonlinearity: it is not clear even why such a linear transformation must exist, as bilinearity is only ensured in the direction of u and of its rest-space. Thus, the tensor T_u measured by a single observer u would not be enough to grasp the physics of the fluid at each event $p \in M$, as in the example above.

(c) Contribution of the anisotropies of Σ: as an additional possibility, the local geometry of Σ at u underlies the measurements of this observer and might provide a contribution for the stress–energy tensor itself.

Summing up, Lorentz–Finsler geometry leads to assume that the measurements by u are not enough to determine the state of the fluid, and the stress-energy tensor should be regarded as a non-isotropic tensor field, determined by the measurements of all the observers.

Formally, this means an *anisotropic tensor* $T \in \mathcal{T}_0^2(M_A)$ (see [20] for a summary of the formal approach), which can be expressed locally as

$$T_v = T^{\mu\nu}(v)\, \partial_\mu\big|_x \otimes \partial_\nu\big|_x \qquad \forall v \equiv (x,y) \equiv y^\mu \frac{\partial}{\partial x^\mu}\bigg|_x \in A \subset TM,$$

where $T^{\mu\nu}(\lambda v) = T^{\mu\nu}(v)$ for all $\lambda > 0$ (i.e., T_v depends only on the direction of v). As a first approach , we can assume $T^{\mu\nu} = T^{\nu\mu}$ (recall Appendix A.5). Consistently, we will assume that there exists a Lorentz–Finsler metric L on M with indicatrix $\Sigma \subset TM$, and, so, indexes can be raised and lowered by using its fundamental tensor g. The fact that T has order 2 is important to establish classical analogies. However, other tensors might appear as more fundamental energy-momentum tensors, and, then, one would try to derive a semi-classical 2-tensor as in Section 3.3.

In principle, the intuitive relativistic interpretations would be transplanted directly to each v, whenever $v \in \Sigma$. That is, given two g_v-unit vectors u, w, the value $T_v(u, w)$ of the 2-covariant stress–energy tensor perceived by the observer v (at $x = \pi(v)$) is obtained as the flux of w-energy-momentum per unit of g_v-volume orthogonal to u. More precisely, let $B(u)$ be a small coordinate 3-cube in a hypersurface g_v-orthogonal to u, and P_B is the total flux of the energy-momentum of particles crossing $B(u)$ (being positive from the $-u$ side to the u side and negative the opposite direction), then the w-energy-momentum per unit of g_v-volume is

$$\epsilon\, T_v(u, w) := \lim_{Vol_{g_v}(B(u)) \to 0} \frac{g_v(P_B, w)}{Vol_{g_v}(B(u))}.$$

where $\epsilon = g_v(w, w)$. As a Finslerian subtlety, recall that g_v is only defined in $T_v(T_x M)$ and then in $T_x M$ (i.e., it is trivially extended to $B(u)$ in a coordinate-depending way), but the above limit depends only on the value of g_v. Namely, if one considers two semi-Riemannian metrics g and $\tilde{g}$ in a neighborhood of p such that $g_p = \tilde{g}_p$ and B_n are open subsets with p in the interior of B_m for all $n \in N$ and $\lim_{n \to +\infty} vol_g(B_m) = 0$, then

$$\lim_{m \to +\infty} \frac{vol_g(B_m)}{vol_{\tilde{g}}(B_m)} = 1.$$

In particular, we have the interpretations (recall signature $(+, -, -, -)$):

1. $T_v(v, v)$ is the energy density measured by $v \in \Sigma$,

$$T_v(v, v) := \lim_{Vol_{g_v}(B(v)) \to 0} \frac{g_v(P_B, v)}{Vol_{g_v}(B(v))} = \lim_{Vol_{g_v}(B(v)) \to 0} \frac{E_B}{Vol_{g_v}(B(v))},$$

being $E_B := g_v(P_B, v)$ the measured energy.

2. If w is g_v-orthogonal to v and g_v-unit, $T_v(w, v)$ measures the flow of energy per unit of g_v-volume in a surface g_v-orthogonal to v and w (i.e., some small surface of area A flowing a lapse Δt), while $T_v(v, u)$ measures the w-momentum density,

$$T_v(w, v) := \lim_{Vol_{g_v}(B(w)) \to 0} \frac{g_v(P_B, v)}{Vol_{g_v}(B(w))} = \lim_{Vol_{g_v}(A) \to 0} \frac{1}{A} \left\{ \lim_{\Delta t \to 0} \frac{E_B}{\Delta t} \right\}.$$

$$-T_v(v, w) := \lim_{Vol_{g_v}(B(v)) \to 0} \frac{g_v(P_B, w)}{Vol_{g_v}(B(v))}.$$

3. If z, w are g_v-orthogonal to v and g_v-unit, $T_v(z, w)$ measures the flow of w-momentum per unit of g_v-volume in a surface g_v-orthogonal to v and z,

$$-T_v(z, w) := \lim_{Vol_{g_v}(B(z)) \to 0} \frac{g_v(P_B, w)}{Vol_{g_v}(B(z))} = \lim_{Vol_{g_v}(A) \to 0} \frac{1}{A} \left\{ \lim_{\Delta t \to 0} \frac{g_v(P_B, w)}{\Delta t} \right\}.$$

3.3. Lagrangian Viewpoint

In the Lagrangian approach for Special Relativity, the background spacetime is assumed to be endowed with a flat metric η. So, the Lagrangian $\mathcal{L}$ is constructed by using the prescribed η and some matter fields ϕ_α. The stress–energy tensor coincides with the *canonical* energy–momentum tensor associated with the Lagrangian, in most cases (the exceptions include theories involving spin). This canonical tensor appears as the Noether current associated with the invariance by spacetime translations (i.e., when $\mathcal{L}(\phi_\alpha, \partial_\mu \phi_\alpha, x^\mu) \equiv \mathcal{L}(\phi_\alpha, \partial_\mu \phi_\alpha)$), namely[5]

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_\alpha)} \partial^\nu \phi_\alpha - \eta^{\mu\nu} \mathcal{L}. \tag{8}$$

In principle, these interpretations would hold unaltered for the case of an affine space with a Lorentz norm, including the case of Very Special Relativity.

In General Relativity, however, the Lagrangian formulation introduces a background Lagrangian independent of matter fields (the Einstein–Hilbert one, eventually with a cosmological constant) and, then, a matter Lagrangian $\mathcal{L}_{matter}$, which includes a constant of coupling with the background. Then, the safest way to define the stress–energy tensor is the canonical one obtained as the corresponding action term $\delta S_{matter} / \delta g^{\mu\nu}$ in the Euler–Lagrange equations[6],

$$T_{\mu\nu} = -2 \frac{\delta \mathcal{L}_{matter}}{\delta g^{\mu\nu}} + g_{\mu\nu} \mathcal{L}_{matter}. \tag{9}$$

Any tensor obtained in this way will have some advantages to play the role of a stress–energy tensor, because it will be automatically symmetric (in contrast to (8)) and will have vanishing divergence.

In the Finslerian setting, the variational viewpoint has been systematically studied in a very recent study by Hohmann, Pfeifer, and Voicu [16]. Previously, the background Lagrangian closest to the Einstein–Hilbert functional in the Finslerian setting had been studied in [15,29]. Such a functional is obtained as the integral of the Ricci scalar function on the indicatrix of the Lorentz–Finsler metric[7] L. Taking into account this background functional, they define the energy-momentum scalar function by taking the corresponding variational action term (Formula (84) in [16]),

$$\mathfrak{T} = -2 \frac{L^3}{|g|} \frac{\delta \mathcal{L}_{matter}}{\delta L}.$$

Notice that, here, the functional coordinate for the Lagrangian is L, and, thus, an (anisotropic) function rather than a 2-tensor is obtained. However, starting at this function

some tensors become useful (Formulas (88) and (91) in [16]), in particular a canonically associated (anisotropic Liouville) 2-tensor

$$\Theta^\mu_\nu = \mathfrak{T}\, C^\mu\, C_\nu$$

as in Example 1. Notice that, essentially, the information of these tensors is codified in $\mathfrak{T}$. Even though such a tensor is justified by the procedure of Gotay–Mardsen in [30], some issues as the following ones might deserve interest for a further discussion:

1. This is not the unique natural possibility to construct an anisotropic 2-tensor starting at $\mathfrak{T}$. For example, an alternative would be the vertical Hessian,

$$T_{\mu\nu} = \dot{\partial}_{\mu,\nu}\mathfrak{T} \equiv \frac{\partial^2 \mathfrak{T}}{\partial y^\mu \partial y^\nu}. \tag{10}$$

It is natural to wonder about the choice closer to the relativistic intuitions about the stress–energy tensor.

2. Recently, the Palatini approach has also been studied for the Finslerian setting [17]. There, the dynamic variables are L and the components of an (independent) non-linear connection. Thus, a similar Lagrangian procedure would lead to a higher-order tensor. In the relativistic setting, this approach supports classical relativity, as it recovers both equations and (in the symmetric case) the Levi–Civita connection. However, the Palatini approach is no longer equivalent in the Finslerian case, as it yields non-equivalent connections, and it shows a variety of possibilities for the non-linear connections. So, it is natural to wonder about the most natural choice of a Lagrangian-based stress–energy tensor in this setting.

Finally, let us discuss an example analyzed from the Lagrangian viewpoint in [1,16] taking into account also the observers' one in Section 3.2.

Example 2. *The gravitational field sourced by a kinetic gas has been deeply studied in [1,16]. In the relativistic setting, this is derived from the Einstein–Vlasov equations in terms of a 1-particle distribution function (1PDF) $\phi(x,\dot{x})$, which encodes how many gas particles at a given spacetime point x propagate on worldlines with a normalized 4-velocity $\dot{x}$. Specifically, the stress energy tensor is:*

$$T^{\mu\nu}(x) = \int_{\Sigma_x} \dot{x}^\mu \dot{x}^\nu \phi(x,\dot{x})\, dvol_{g_x}, \qquad x \in M,$$

being Σ_x the indicatrix (future-directed unit vectors of the Lorentz metric) and $dvol_{g_x}$ the volume element induced by the scalar product g_x at each x. In [1], they propose to derive the gravitational field of a kinetic gas directly from the 1PDF without averaging, i.e., taking into account the full information on the velocity distribution. This leads to consider the function $\phi : \Sigma \to R$, $u \equiv (x,\dot{x}) \mapsto \phi(u) \geq 0$ as an energy–momentum function, which plays the role of a stress–energy tensor (even though it is a scalar rather than a 2-tensor). Moreover, the original Lorentz metric is naturally allowed to be Lorentz–Finsler, which permits to obtain more general cosmological models (Section III in [1]).

Indeed, up to a coupling constant, ϕ is regarded directly as the matter source in the Finslerian Einstein–Hilbert equation (i.e., it is placed at the right-hand side of this equation, (Equation (7) in [1])). It is worth pointing out:

*   *ϕ can be reobtained as a Lagrangian energy-momentum by inserting it directly as a term in the background Lagrangian (Equation (75) in [16]). However, the Lagrangian is not natural then, as ϕ is written in terms of $x,\dot{x}$ (recall (Appendix 3, Section (a) in [16])).*
*   *As discussed above, such a function allows one to construct several tensors, in particular, the vertical Hessian $\partial^2 \phi / \partial \dot{y}^\mu \partial \dot{y}^\nu$ (as in (10)), which also might play a role to compare with the relativistic $T^{\mu\nu}(x)$.*

Anyway, starting at the 1PDF ϕ, another Finslerian interpretation would be possible. In particular, one can define the energy–momentum distribution $\phi(u)u$. *Then, given an observer $v \in \Sigma$ and a g_v-unit vector, the* w-energy–momentum *might be defined as*

$$g_v(u, w)\phi(u).$$

n particular, when $w = v$, this would be the energy perceived by v, *and when w is unit and g_v-orthogonal to v would be (minus) the* momentum in the direction w *(compare with the discussion at the end of Section 3.2). So, an alternative stress–energy tensor perceived by each observer $v \in \Sigma$ might be defined as the anisotropic tensor:*

$$T_v(w, z) = \int_{\Sigma_{\pi(v)}} g_v(u, w)g_v(u, z)\phi(u)dvol_{g_v},$$

where the integration in u is carried out with the volume form of $(Sectionigma_{\pi(v)}, g_v)$, denoted by $dvol_{g_v}$.

4. Divergence of Anisotropic Vector Fields

After studying the basic properties of the Finslerian stress–energy tensor T, our next aim was to analyze the meaning and significance of the *infinitesimal conservation law* $\mathrm{div}(T) = 0$. Along this and the next section, we will always consider an anisotropic tensor $T \in \mathcal{T}_1^1(M_A)$ interpreted as an endomorphism of anisotropic vector fields. $T^\flat \in \mathcal{T}_2^0(M_A)$ and $T^\sharp \in \mathcal{T}_2^0(M_A)$ will be defined on vectors and 1-forms by $T^\flat(X, Y) := g(X, T(Y))$ and $T^\sharp(\theta, \eta) := g^*(T^*(\theta), \eta)$ resp., where g^* is the inverse fundamental tensor, and T^* is the transpose of T. They will have components $\left(T^\flat\right)_{ij} = g_{il}T_j^l =: T_{ij}$ and $\left(T^\sharp\right)^{ij} = T_l^i g^{lj} =: T^{ij}$, and in principle we will not even assume that these are symmetric. We will be assuming that M is orientable and oriented. This is not restrictive: one could always reduce the theory to this case by pulling back all the objects (the fibered manifold $A \to M$ included) to the oriented double cover of M (Chapter 15 in [31]).

Let us briefly recall the mathematically precise meaning of the conservation laws in classical General Relativity (g, T, and X isotropic). One has

$$\mathrm{div}(T(X)) = \nabla_i(T_j^i X^j) = \nabla_i T_j^i X^j + T_j^i \nabla_i X^j = \mathrm{div}(T)(X) + \mathrm{trace}(T(\nabla X)) \quad (11)$$

with ∇ the Levi–Civita connection. The first contribution vanishes due to $\mathrm{div}(T) = 0$, and there are different situations in which the second one vanishes as well. For instance, if $T^\flat(-, \nabla_- X)$ is antisymmetric, then

$$\mathrm{trace}(T(\nabla X)) = T_j^i \nabla_i X^j = g^{il} T_{lj} \nabla_i X^j = \frac{1}{2}g^{il}\left(T_{lj}\nabla_i X^j + T_{ij}\nabla_l X^j\right) = 0, \quad (12)$$

and if $T^\flat$ is symmetric and $\nabla X^\sharp$ is antisymmetric (equiv., X is a Killing vector field), then also

$$\mathrm{trace}(T(\nabla X)) = g^{il} T_{lj} \nabla_i X^j = \frac{1}{2}T_{lj}\left(g^{li}\nabla_i X^j + g^{ji}\nabla_i X^l\right) = 0. \quad (13)$$

Anyway, whenever $\mathrm{trace}(T(\nabla X)) = 0$, one can integrate (11) and apply the pseudo-Riemannian divergence theorem to get the *integral conservation law*

$$\int_{\partial D} \imath_{T(X)}(d\mathrm{Vol}) = 0, \quad (14)$$

where $\overline{D}$ is a domain of appropriate regularity, $\imath$ is the interior product operator, and $d\mathrm{Vol}$ is the metric volume form. In a sense that will be made more precise in §5, this is expressing that the total amount of X-momentum in a space region only changes along time as much as it flows across the spatial boundary of the region. In this way, there is no "creation" nor "destruction" of X-momentum in any space region.

Extending the infinitesimal or the integral conservation laws poses, first and foremost, the problem of appropriately defining the divergence of an anisotropic T. Observe that a priori it is not clear even how to define the divergence of a vector field Z, isotropic or not, as one could consider trace(∇Z) for different anisotropic connections ∇, mainly Chern's and Berwald's. An alternative is to seek for a more geometric, hence, *unbiased*, definition. For instance, the *metric (anisotropic) volume form of L*,

$$d\text{Vol} = \sqrt{|\det g_{ab}(x,y)|}\,dx^1 \wedge \cdots \wedge dx^n \in \Omega_n(M_A) \tag{15}$$

for $(x^1, \ldots, x^n)$ positively oriented, is well-defined, and when $Z \in \mathfrak{X}(M)$ (i.e., Z is isotropic), so is the Lie derivative

$$\mathfrak{L}_Z \colon \mathcal{T}(M_A) \to \mathcal{T}(M_A)$$

(see § 5 in [21]). So, by analogy with the classical case, one could think of $\mathfrak{L}_Z(d\text{Vol})$ for defining div(Z).

It turns out that the unbiased definition, including all $Z \in \mathfrak{X}(M_A)$, is achieved with a modification of this Lie derivative that we will regard as an extension of the classical Lie bracket. We devote the next subsection to the technical mathematical foundations of such an *anisotropic Lie bracket*, which needs of a nonlinear connection on $A \to M$ to be well-defined. All the maps $\mathcal{T}(M_A) \to \mathcal{T}(M_A)$ that will appear in Section 4.1 will be *(anisotropic) tensor derivations* in the sense of (Definition 2.6 in [21]), and their local nature will be apparent, so we will not explicitly discuss it. For example, the Lie derivative along $Z \in \mathfrak{X}(M)$ is the only tensor derivation such that for $X \in \mathfrak{X}(M)$ and $f \in \mathcal{F}(A)$,

$$\mathfrak{L}_Z X = [Z, X], \qquad \mathfrak{L}_Z f = Z^c(f) := Z^k \frac{\partial f}{\partial x^k} + y^k \frac{\partial Z^i}{\partial x^k}\frac{\partial f}{\partial y^i}. \tag{16}$$

4.1. Mathematical Formalism of the Anisotropic Lie Bracket

During this subsection, we fix an arbitrary nonlinear connection given by $TA = HA \oplus VA$ or by the nonlinear covariant derivative D (keep in mind (1) and (2)), and also an anisotropic vector field $Z \in \mathfrak{X}(M_A)$

For $X \in \mathfrak{X}(M_A)$, it is very natural to consider the commutator of the horizontal lifts of Z and X:

$$\left[Z^H, X^H\right] = \left[Z^j\delta_j, X^k\delta_k\right] = \left(Z^j\delta_j X^i - X^j\delta_j Z^i\right)\delta_i + Z^j X^k\left[\delta_j, \delta_k\right] \in \mathfrak{X}(A).$$

We recall that $Z^j X^k\left[\delta_j, \delta_k\right]$ is always vertical. Indeed, $\left[\delta_j, \delta_k\right] = \mathcal{R}^i_{jk}\dot\partial_i$, where $\mathcal{R}$ is the curvature tensor of the nonlinear connection (see [17], where this curvature is regarded as an anisotropic tensor and the homogeneity of the connection is not really required). This means that the horizontal part of $\left[Z^H, X^H\right]$ has coordinates $Z^j\delta_j X^i - X^j\delta_j Z^i$, and this corresponds to a globally well-defined A-anisotropic vector field:

$$\mathfrak{l}^H_Z X := \left(Z^j\delta_j X^i - X^j\delta_j Z^i\right)\partial_i \in \mathfrak{X}(M_A). \tag{17}$$

Definition 1. $\mathfrak{l}^H_Z X$ *is the* anisotropic Lie bracket of Z and X with respect to the nonlinear connection HA.

Remark 1. *The word "anisotropic" could be omitted in the previous definition, in the sense that for $Z, X \in \mathfrak{X}(M_A)$, there is no other Lie bracket, isotropic or not, defined in general. Nonetheless, (17) makes apparent that when $Z, X \in \mathfrak{X}(M)$ (i.e., when Z and X are isotropic), $\mathfrak{l}^H_Z X$ coincides with the standard Lie bracket $[Z, X]$ regardless of the connection.*

Lemma 1. *Given a nonlinear connection* HA, $V \in \mathfrak{X}^A(U)$, $f \in \mathcal{F}(A)$ *and anisotropic vector fields* $X, Z \in \mathfrak{X}(M_A)$, *it holds that*

$$Z^{\mathrm{H}}(f) = Z(f(V)) - \dot{\partial}_{\mathrm{D}_Z V} f, \tag{18}$$

$$\left(\mathfrak{l}_Z^{\mathrm{H}} X \right)_V = [Z_V, X_V] - \left(\dot{\partial}_{\mathrm{D}_Z V} X \right)_V + \left(\dot{\partial}_{\mathrm{D}_X V} Z \right)_V. \tag{19}$$

Proof. Observe that

$$
\begin{aligned}
Z(f(V)) - \dot{\partial}_{\mathrm{D}_Z V} f &= Z^i \left(\frac{\partial f}{\partial x^i}(V) + \frac{\partial f}{\partial y^j}(V) \frac{\partial V^j}{\partial x^i} \right) \\
&\quad - \frac{\partial f}{\partial y^j}(V) Z^k \left(\frac{\partial V^j}{\partial x^k} - N_k^j(V) \right) \\
&= Z^i \left(\frac{\partial f}{\partial x^i}(V) - \frac{\partial f}{\partial y^j}(V) N_i^j(V) \right) \\
&= Z^{\mathrm{H}}(f),
\end{aligned}
$$

which concludes (18). In particular, $\delta_i f(V) = \partial_i(f(V)) - \left(\dot{\partial}_{\mathrm{D}_{\partial_i} V} f \right)(V)$, and using this in (17), (19) follows. $\quad\square$

We also recall that the *torsion* of an *A*-anisotropic connection ∇ (18 in [21]), (Definition 5 in [20]) is the anisotropic tensor Tor $\in \mathcal{T}_2^1(M_A)$ defined first on isotropic fields $Z, X \in \mathfrak{X}(M)$ by $\mathrm{Tor}(Z, X) = \nabla_Z X - \nabla_X Z - [Z, X]$ and then extended by $\mathcal{F}(A)$-bilinearity. Therefore, it can be regarded as and $\mathcal{F}(A)$-bilinear map Tor: $\mathfrak{X}(M_A) \times \mathfrak{X}(M_A) \to \mathfrak{X}(M_A)$ and it has coordinates

$$\mathrm{Tor}_{jk}^i = \Gamma_{jk}^i - \Gamma_{kj}^i, \tag{20}$$

where the Γ_{jk}^i's are the Christoffel symbols of ∇[8].

Theorem 1. *Let a nonlinear connection* T$A = HA \oplus VA$ *and an anisotropic vector field* $Z \in \mathfrak{X}(M_A)$ *be fixed.*
(A) If ∇ *is any A-anisotropic connection whose underlying nonlinear connection is* HA, *then for any* $X \in \mathfrak{X}(M_A)$,

$$\mathrm{Tor}(Z, X) = \nabla_Z X - \nabla_X Z - \mathfrak{l}_Z^{\mathrm{H}} X \tag{21}$$

(where Tor *is the torsion of* ∇*).*
(B) By imposing the Leibniz rule with respect to tensor products and the commutativity with contractions, the map $X \mapsto \mathfrak{l}_Z^{\mathrm{H}} X$ *extends unequivocally to an (anisotropic) tensor derivation* $\mathfrak{l}_Z^{\mathrm{H}} : \mathcal{T}_s^r(M_A) \to \mathcal{T}_s^r(M_A)$ *given by*

$$
\begin{aligned}
\mathfrak{l}_Z^{\mathrm{H}} T(\theta^1, \ldots, \theta^r, X_1, \ldots, X_s) &= Z^{\mathrm{H}}(T(\theta^1, \ldots, \theta^r, X_1, \ldots, X_s)) \\
&\quad - \sum_{\mu=1}^r T(\theta^1, \ldots, \mathfrak{l}_Z^{\mathrm{H}} \theta^\mu, \ldots, \theta^r, X_1, \ldots, X_s) \\
&\quad - \sum_{\nu=1}^s T(\theta^1, \ldots, \theta^r, X_1, \ldots, \mathfrak{l}_Z^{\mathrm{H}} X_\nu, \ldots, X_s)
\end{aligned} \tag{22}
$$

for $\theta^\mu \in \Omega_1(M)$ *and* $X_\nu \in \mathfrak{X}(M)$. *In coordinates, if*

$$T = T_{j_1, \ldots, j_s}^{i_1, \ldots, i_r}(x, y) \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s},$$

then

$$\left(\mathfrak{l}_Z^{\mathrm{H}} T \right)_{j_1, \ldots, j_s}^{i_1, \ldots, i_r} = Z^k \frac{\delta T_{j_1, \ldots, j_s}^{i_1, \ldots, i_r}}{\delta x^k} - \sum_{\mu=1}^r \frac{\delta Z^{i_\mu}}{\delta x^k} T_{j_1, \ldots, j_s}^{i_1, \ldots, k, \ldots, i_r} + \sum_{\nu=1}^s \frac{\delta Z^k}{\delta x^{j_\nu}} T_{j_1, \ldots, k, \ldots, j_s}^{i_1, \ldots, i_r}. \tag{23}$$

(C) The map

$$\mathfrak{L}_Z^H := \mathfrak{l}_Z^H - \dot{\partial}_{\mathfrak{l}_Z^H C} : \mathcal{T}(M_A) \to \mathcal{T}(M_A)$$

is also a tensor derivation. When $Z \in \mathfrak{X}(M)$,

$$\mathfrak{L}_Z^H T = \mathfrak{L}_Z T \tag{24}$$

for all $T \in \mathcal{T}(M_A)$, *where* $\mathfrak{L}_Z$ *is the Lie derivative* (16), *regardless of the nonlinear connection.*
(D) Given $V \in \mathfrak{X}^A(U)$ *and* $\omega \in \Omega_n(M_A)$ $(n = \dim M)$, *it holds that*

$$\left(\mathfrak{l}_Z^H \omega\right)_V = \mathfrak{L}_{Z_V}(\omega_V) - \dot{\partial}_{D_Z V}\omega - \text{trace}(\dot{\partial}_{DV} Z)\omega. \tag{25}$$

Proof. (A) It is straightforward to compute that the right-hand side of (21) is $\mathcal{F}(A)$-multilinear. Moreover, the identity is trivial on isotropic vector fields $X, Z \in \mathfrak{X}(M)$, as $\mathfrak{l}_Z^H X = [X, Z]$ in this case, which concludes.
   (B) Given $f \in \mathcal{T}_0^0(M_A) = \mathcal{F}(A)$, for $X \in \mathcal{T}_0^1(M_A) = \mathfrak{X}(M_A)$ it follows from (17) that

$$\mathfrak{l}_Z^H(fX) = Z^H(f)X + f\,\mathfrak{l}_Z^H X.$$

Thus, in order to respect the Leibniz rule, the only possibility is to define

$$\mathfrak{l}_Z^H f = Z^H(f) = Z^k \frac{\delta f}{\delta x^k}. \tag{26}$$

Now, given $\theta \in \mathcal{T}_1^0(M_A) = \Omega_1(M_A)$, in order to respect again the Leibniz rule and the commutativity with contractions, the only possibility is to define $\mathfrak{l}_Z^H \theta$ on every $X \in \mathfrak{X}(M_A)$ by

$$\left(\mathfrak{l}_Z^H \theta\right)(X) = Z^H(\theta(X)) - \theta(\mathfrak{l}_Z^H X) = \left(Z^k \frac{\delta \theta_j}{\delta x^k} + \frac{\delta Z^k}{\delta x^j}\theta_k\right)X^j. \tag{27}$$

(26), (17), and (27) make apparent that $\mathfrak{l}_Z^H$ is already local on functions, vector fields, and 1-forms, and they allow to compute

$$\mathfrak{l}_Z^H(\partial_i) = -\frac{\delta Z^k}{\delta x^i}\partial_k, \qquad \mathfrak{l}_Z^H(dx^j) = \frac{\delta Z^j}{\delta x^k}dx^k. \tag{28}$$

Finally, given $T \in \mathcal{T}_s^r(M_A)$, one is led to define $\mathfrak{l}_Z^H T$ by (22). Clearly, this indeed provides a tensor derivation and (23) follows from the evaluation of (22) at $(dx^{i_1}, \ldots, dx^{i_r}, \partial_{j_1}, \ldots, \partial_{j_s})$ together with (26) and (28).
   (C) $\dot{\partial}_X : \mathcal{T}(M_A) \to \mathcal{T}(M_A)$ is a tensor derivation for any $X \in \mathfrak{X}(M_A)$, in particular for

$$X = \mathfrak{l}_Z^H C = \left(Z^j \delta_j y^i - y^j \delta_j Z^i\right)\partial_i = -\left(Z^j N_j^i + y^j \delta_j Z^i\right)\partial_i \tag{29}$$

(see (17)). Thus, the difference $\mathfrak{L}_Z^H = \mathfrak{l}_Z^H - \dot{\partial}_{\mathfrak{l}_Z^H C}$ is again a derivation. As for the last assertion, where $Z \in \mathfrak{X}(M)$, we are going to use (Proposition 2.7 in [21]). For $X \in \mathfrak{X}(M)$, we have

$$\mathfrak{L}_Z^H X = \mathfrak{l}_Z^H X = [Z, X] = \mathfrak{L}_Z X \tag{30}$$

(recall Remark 1). For $f \in \mathcal{F}(A)$, we have

$$\mathfrak{L}_Z^{\mathrm{H}} f = \mathfrak{l}_Z^{\mathrm{H}} f - \dot{\partial}_{\mathfrak{l}_Z^{\mathrm{H}} C} f = Z^j \delta_j f + \left(Z^j N_j^i + y^j \delta_j Z^i \right) \dot{\partial}_i f$$

$$= Z^j \left(\partial_j f - N_j^i \dot{\partial}_i f \right) + \left(Z^j N_j^i + y^j \delta_j Z^i \right) \dot{\partial}_i f$$

$$= Z^j \partial_j f + y^j \delta_j Z^i \dot{\partial}_i f$$

$$= \mathfrak{L}_Z f$$

(see (26), (29), (1), and (16)). As $\mathfrak{L}_Z^{\mathrm{H}}$ and $\mathfrak{L}_Z$ act the same on isotropic vector field and anisotropic functions, they are equal.

(D) Observe that for $X \in \mathfrak{X}(M)$, the term $\dot{\partial}_{\mathrm{D}_Z V} X$ vanishes in (19). Moreover, if $Z \in \mathfrak{X}(M_A)$ and $f \in \mathcal{F}(A)$, then $Z^{\mathrm{H}}(f)_V = Z_V(f(V)) - (\dot{\partial}_{\mathrm{D}_Z V} f)(V)$. Given a local reference frame $E_1, \ldots, E_n \in \mathfrak{X}(U)$, and taking into account the last two identities and the definitions of $\mathfrak{l}^{\mathrm{H}}$ and $\mathfrak{L}$, it follows that

$$\left(\mathfrak{l}_Z^{\mathrm{H}} \omega \right)_V (E_1, \ldots, E_n) - \mathfrak{L}_{Z_V}(\omega_V)(E_1, \ldots, E_n) = -\dot{\partial}_{\mathrm{D}_Z V} \omega(E_1, \ldots, E_n)$$

$$- \sum_{i=1}^n \omega(E_1, \ldots, \dot{\partial}_{\mathrm{D}_{E_i} V} Z, \ldots, E_n).$$

As $\omega(E_1, \ldots, \dot{\partial}_{\mathrm{D}_{E_i} V} Z, \ldots, E_n) = E_i^*(\dot{\partial}_{\mathrm{D}_{E_i} V} Z) \omega_V(E_1, \ldots, E_n)$, (25) follows. $\square$

Definition 2. *The tensor derivation* $\mathfrak{l}_Z^{\mathrm{H}} \colon \mathcal{T}(M_A) \to \mathcal{T}(M_A)$ *defined in Theorem* 1 *(B) is the (anisotropic) Lie bracket with* Z, *while* $\mathfrak{L}_Z^{\mathrm{H}} \colon \mathcal{T}(M_A) \to \mathcal{T}(M_A)$ *is the (anisotropic) Lie derivative along* Z, *both of them with respect to the connection* $\mathrm{H}A$.

Remark 2 (Anisotropic Lie bracket and Lie derivative)**.** *The derivation* $\mathfrak{L}_Z^{\mathrm{H}}$ *defined in Theorem* 1 *(C) would be the* Lie derivative along Z *with respect to* $\mathrm{H}A$. *Analogously to the discussion of Remark* 1, *what makes this name consistent is* (24): *whenever the Lie derivative along* Z *was already defined,* $\mathfrak{L}_Z^{\mathrm{H}}$ *coincides with it. Even though the Lie bracket and the Lie derivative are equal in the classical regime, it is heuristically useful to regard* $\mathfrak{l}^{\mathrm{H}}$ *as the anisotropic generalization of the former and* $\mathfrak{L}^{\mathrm{H}}$ *as that of the latter, in order to distinguish them. It is actually* $\mathfrak{l}^{\mathrm{H}}$, *and not* $\mathfrak{L}$, *which will be relevant for the definition of divergence. The reason is that the former, as we will see below, has a clear geometric interpretation in terms of flows, while the latter would just add the term* $\dot{\partial}_{\mathfrak{l}_Z^{\mathrm{H}} C}$ *to that interpretation. Moreover, Theorem* 1 *(D) actually corresponds to a Cartan formula for* $\mathfrak{L}_Z$ *whose full development we postpone for a future work. Thus,* $\mathfrak{L}_Z(d\mathrm{Vol}) = \mathfrak{L}_Z^{\mathrm{H}}(d\mathrm{Vol})$ *can be regarded as an initial guess for the divergence of* Z, *but we will not employ* $\mathfrak{L}^{\mathrm{H}}$ *from now on.*

Let us observe that given a diffeomorphism $\psi_t \colon M \to M$ that is the flow of an isotropic vector field Z, we can define the pullback $\psi_t^*(\omega)$ of an anisotropic differential form $\omega \in \Omega_s(M_A)$ as the anisotropic form given by $\psi_t^*(\omega)_v(u_1, \ldots, u_s) := \omega_{P_t(v)}(\mathrm{d}\psi_t(u_1), \ldots, \mathrm{d}\psi_t(u_s))$, where $P_t(v)$ is the $\mathrm{H}A$-parallel transport of v along the integral curve of Z and $u_1, \ldots, u_s \in \mathrm{T}_{\pi(v)} M$.

Proposition 1. *If* $Z \in \mathfrak{X}(M)$ *and* $\omega \in \Omega_s(M_A)$, *then*

$$\mathfrak{l}_Z^{\mathrm{H}} \omega = \lim_{t \to 0} \frac{\psi_t^*(\omega) - \omega}{t}, \tag{31}$$

where ψ_t *is the (possibly local) flow of* Z.

Proof. Observe that $\psi_t^*(\omega)_v$ can be obtained as $\psi_t^*(\omega_V)$ with V an extension of v such that $\mathrm{D}_Z V = 0$. Then (25) and the classical formula for the Lie derivative in terms of the flow imply (31). $\square$

Remark 3. *Even though, for convenience, we stated the previous geometrical interpretation for an s-form ω, it should be clear that it holds true for any r-contravariant s-covariant A-anisotropic tensor.*

4.2. Lie Bracket Definition of Divergence

Finally, in this and the next subsections a pseudo-Finsler metric L defined on A is fixed again. In its presence, and in view of the Riemannian case and Proposition 1, the most natural way of defining the divergence of an anisotropic vector field Z is by $\mathfrak{l}_Z^H(d\mathrm{Vol})$. Here there is a canonical choice for HA: the metric nonlinear connection of L. The definition obtained this way is unbiased, in that one does not choose any anisotropic connection *a priori*. Notwithstanding, it will turn out to be most conveniently expressed in terms of the Chern connection.

Definition 3. *For $Z \in \mathfrak{X}(M_A)$, its divergence with respect to the pseudo-Finsler metric L is the anisotropic function $\mathrm{div}(Z) \in \mathcal{F}(A)$ defined by*

$$\mathfrak{l}_Z^H(d\mathrm{Vol}) =: \mathrm{div}(Z)d\mathrm{Vol},$$

where HA and dVol are the metric nonlinear connection (4) and the metric volume form (15) of L, resp.

Remark 4. *Even though we will keep assuming it for simplicity, the hypothesis of M being orientable is not really needed for this definition. As in pseudo-Riemannian geometry, on small enough open sets $U \subseteq M$ it is always possible to choose an orientation, define $d\mathrm{Vol}_U \in \Omega_n(M_A)$ with respect to it and put $\mathrm{div}(Z)|_{A \cap TU}d\mathrm{Vol}_U := \mathfrak{l}_Z^H(d\mathrm{Vol}_U)$. The different definitions will be coherent because when the orientation changes, $d\mathrm{Vol}_U$ changes to $-d\mathrm{Vol}_U$ and*

$$\mathfrak{l}_Z^H(-d\mathrm{Vol}_U) = -\mathfrak{l}_Z^H(d\mathrm{Vol}_U) = -\mathrm{div}(Z)|_{A \cap TU}d\mathrm{Vol}_U = \mathrm{div}(Z)_{A \cap TU}(-d\mathrm{Vol}_U).$$

In particular, when M is orientable, $\mathrm{div}(Z)$ is independent of the orientation choice.

Proposition 2. *Let L be a fixed pseudo-Finsler metric defined on A, and let $Z \in \mathfrak{X}(M_A)$. If ∇ is any symmetric A-anisotropic connection such that its underlying nonlinear connection is the metric one and $\nabla_Z(d\mathrm{Vol}) = 0$, then*

$$\mathrm{div}(Z) = \mathrm{trace}(\nabla Z), \tag{32}$$

or in coordinates,

$$\mathrm{div}(Z) = \frac{\delta Z^i}{\delta x^i} + \Gamma_{ik}^i Z^k \tag{33}$$

This, in particular, is true for the (Levi-Civita)–Chern anisotropic connection of L, so one can take the Christoffel symbols to be those of (5).

Proof. One expresses the Z-Lie bracket of the volume form in terms of the anisotropic connection, analogously to the isotropic case. From (15) and the fact that $\mathfrak{l}_Z^H$ is a tensor derivation, we obtain

$$\begin{aligned}
\mathrm{div}(Z)\sqrt{|\det g_{ab}|} &= \mathrm{div}(Z)d\mathrm{Vol}(\partial_1, \ldots, \partial_n) \\
&= \mathfrak{l}_Z^H(d\mathrm{Vol})(\partial_1, \ldots, \partial_n) \\
&= \mathfrak{l}_Z^H(d\mathrm{Vol}(\partial_1, \ldots, \partial_n)) - \sum_{i=1}^n d\mathrm{Vol}(\partial_1, \ldots, \mathfrak{l}_Z^H\partial_i, \ldots, \partial_n).
\end{aligned}$$

(26) and the fact that HA is the underlying nonlinear connection of ∇ give

$$\mathfrak{l}_Z^H(d\mathrm{Vol}(\partial_1, \ldots, \partial_n)) = Z^H(d\mathrm{Vol}(\partial_1, \ldots, \partial_n)) = \nabla_Z(d\mathrm{Vol}(\partial_1, \ldots, \partial_n)).$$

(21) and Tor $= 0$

$$d\mathrm{Vol}(\partial_1,\ldots,\iota_Z^{\mathrm{H}}\partial_i,\ldots,\partial_n) = d\mathrm{Vol}(\partial_1,\ldots,\nabla_Z\partial_i,\ldots,\partial_n) - d\mathrm{Vol}(\partial_1,\ldots,\nabla_{\partial_i}Z,\ldots,\partial_n).$$

From these and $\nabla_Z(d\mathrm{Vol}) = 0$,

$$\begin{aligned}
\mathrm{div}(Z)\sqrt{|\det g_{ab}|} &= \nabla_Z(d\mathrm{Vol}(\partial_1,\ldots,\partial_n)) - \sum_{i=1}^n d\mathrm{Vol}(\partial_1,\ldots,\nabla_Z\partial_i,\ldots,\partial_n) \\
&\quad + \sum_{i=1}^n d\mathrm{Vol}(\partial_1,\ldots,\nabla_{\partial_i}Z,\ldots,\partial_n) \\
&= \nabla_Z(d\mathrm{Vol})(\partial_1,\ldots,\partial_n) + \sum_{i=1}^n d\mathrm{Vol}(\partial_1,\ldots,\nabla_{\partial_i}Z,\ldots,\partial_n) \\
&= \sum_{i=1}^n d\mathrm{Vol}(\partial_1,\ldots,\nabla_{\partial_i}Z,\ldots,\partial_n) \\
&= \mathrm{trace}(\nabla Z)\sqrt{|\det g_{ab}|},
\end{aligned} \tag{34}$$

where the last equality is reasoned analogously as in the proof of (25).

For the Chern connection, it can be checked that $\nabla(d\mathrm{Vol}) = 0$ by considering a parallel orthonormal basis with respect to a parallel observer V along the integral curves of any vector field. The coordinate expression of $\mathrm{trace}(\nabla Z)$ in this case concludes (33). $\square$

4.3. Divergence Theorem and Boundary Term Representations

Our Lie bracket derivation allows us to obtain a statement of the Finslerian divergence theorem that subsumes both Rund's (3.17 in [18]) and Minguzzi's (Theorem 2 in [19]). This way, it does not need of computations in coordinates from the beginning nor of the "pullback metric" (g_V in our notation). Naturally, our statement does not include Shen's (Theorem 2.4.2 in [32]), as this one is an independent generalization of the Riemannian theorem not dealing with anisotropic differential forms nor vector fields.

Lemma 2. *For $X \in \mathfrak{X}(M_A)$, the vertical derivative of $d\mathrm{Vol}$ is given by*

$$\dot{\partial}_X(d\mathrm{Vol}) = C^{\mathrm{m}}(X)d\mathrm{Vol}, \tag{35}$$

where C^{m} is the mean Cartan tensor of L (see (3)).

Proof. Let $E_1(t), \ldots, E_n(t)$ be a positively oriented g_{v+tX}-orthonormal basis for every $t \in [0,\varepsilon]$ for a certain $\varepsilon > 0$. Then $d\mathrm{Vol}_{v+tX}(E_1(t),\ldots,E_n(t)) = 1$ for all $t \in [0,\varepsilon]$. This implies that

$$\dot{\partial}_X(d\mathrm{Vol})_v(E_1(0),\ldots,E_n(0)) + \sum_{i=1}^n d\mathrm{Vol}_v(E_1(0),\ldots,\dot{E}_i(0),\ldots,E_n(0)) = 0.$$

Moreover, as $g_{v+tX}(E_i(t),E_i(t)) = \pm 1$,

$$2C_v(E_i(0),E_i(0),X) + 2g_v(\dot{E}_i(0),E_i(0)) = 0.$$

Using this relation above, we conclude (35). $\square$

In the present article, by a *domain* $\overline{D}$ we understand a nonempty connected set that coincides with the closure of its interior D; then, its boundary is $\partial\overline{D} = \partial D$. Physically, it is very important to include examples in which different parts of ∂D have different causal characters, and this tipically leads to the boundary not being totally smooth. Hence, we will make a weaker regularity assumption that still allows one to apply Stokes' theorem

on $\overline{D}$. A subset of M has 0 *m-dimensional measure* if its intersection with any embedded m-dimensional submanifold $\sigma \subseteq M$ is of 0 measure in the smooth manifold σ. Finally, the interior product of an s-form ω with a vector field X will be

$$\iota_X\omega := \omega(X, -, \ldots, -).$$

Theorem 2. *Let L be a fixed pseudo-Finsler metric defined on A. If*

(i) $Z \in \mathfrak{X}(M_A)$ *is an anisotropic vector field,*

(ii) $V \in \mathfrak{X}^A(U)$ *is an A-admissible field with $U \subseteq M$ open, and*

(iii) $\overline{D} \subseteq U$ *is a domain with ∂D smooth up to subset of 0 $(n-1)$-dimensional measure on M and $\mathrm{Supp}(Z_V) \cap \overline{D}$ compact,*

then

$$\int_D \mathrm{div}(Z)_V d\mathrm{Vol}_V + \int_D \left\{ \mathrm{C}^{\mathbf{m}}(\mathrm{D}_Z V) + \mathrm{trace}(\dot{\partial}_{\mathrm{D}V} Z) \right\} d\mathrm{Vol}_V \tag{36}$$
$$= \int_{\partial D} \iota_{Z_V}(d\mathrm{Vol}_V),$$

where $\mathrm{C}^{\mathbf{m}}$ is the mean Cartan tensor, and $\mathrm{D}V$ is computed with the metric nonlinear connection (4).

Proof. The idea is to apply Stokes' theorem to $\mathfrak{L}_{Z_V}(d\mathrm{Vol}_V)$. However, taking into account (25) and Lem. 2, it follows that

$$\mathfrak{L}_{Z_V}(d\mathrm{Vol}_V) = \iota_Z^{\mathrm{H}}(d\mathrm{Vol})_V + \left\{ \mathrm{C}^{\mathbf{m}}(\mathrm{D}_Z V) + \mathrm{trace}(\dot{\partial}_{\mathrm{D}V} Z) \right\} d\mathrm{Vol}_V,$$

concluding (36). $\square$

Remark 5 (Riemannian and Finslerian unit normals). *Let $i \colon \Gamma \hookrightarrow M$ be the inclusion of a smooth open subset $\Gamma \subseteq \partial D$.*

(i)   *Even though we do not use the pseudo-Riemannian metric g_V to derive Theorem 2, from our physical viewpoint it is natural to use it to re-express the boundary term. If Γ is non-g_V-lightlike, then for a g_V-normal field $\widehat{N}_V$ and a transverse field X along i, the form*

$$d\sigma_V := \mathrm{sgn}(g_V(\widehat{N}_V, \widehat{N}_V)) \frac{\sqrt{\left| g_V(\widehat{N}_V, \widehat{N}_V) \right|}}{g_V(\widehat{N}_V, X)} i^*(\iota_X(d\mathrm{Vol}_V)) \in \Omega_{n-1}(\Gamma) \tag{37}$$

is nonvanishing and independent of X. In particular,

$$d\sigma_V = \frac{1}{\sqrt{\left| g_V(\widehat{N}_V, \widehat{N}_V) \right|}} i^*(\iota_{\widehat{N}_V}(d\mathrm{Vol}_V))$$

is independent of the scale of $\widehat{N}_V$, which we will always assume to be g_V-unitary and D-salient, so

$$d\sigma_V = i^*(\iota_{\widehat{N}_V}(d\mathrm{Vol}_V))$$

coincides with the hypersurface g_V-volume form of Γ. Taking into account that $i^*(\iota_{Z_V}(d\mathrm{Vol}_V))$ vanishes wherever Z_V is tangent to Γ and that $g_V(\widehat{N}_V, \widehat{N}_V) = \pm 1$, (37) allows us to represent and the right-hand side of (36) as

$$\int_\Gamma \iota_{Z_V}(d\mathrm{Vol}_V) = \int_\Gamma g_V(\widehat{N}_V, \widehat{N}_V) g_V(\widehat{N}_V, Z_V) d\sigma_V. \tag{38}$$

In fact, this is how Rund's divergence theorem follows from Theorem 2.

(ii) *There is another way that one can try to represent the boundary term. Namely, assume that there exists a smooth $\xi \colon p \in \Gamma \to \xi_p \in A \cap T_p M$ with $T_p \Gamma = \operatorname{Ker} g_{\xi_p}(\xi_p, -)$ and $L(\xi_p) = \pm 1$ (in the Lorentz–Finsler case, it will necessarily be $L(\xi) = 1$). This is called a Finslerian unit normal along Γ. Analogously as in (i), one can put*

$$d\Sigma_V^{\xi} := L(\xi) \frac{1}{g_{\xi}(\xi, X)} i^*(\imath_X(d\mathrm{Vol}_V)) = i^*(\imath_\xi(d\mathrm{Vol}_V)),$$

$$\int_{\Gamma} \imath_{Z_V}(d\mathrm{Vol}_V) = \int_{\Gamma} \epsilon_{\xi} L(\xi) g_{\xi}(\xi, Z_V) d\Sigma_V^{\xi}; \tag{39}$$

here, due to the possible orientation difference between both sides,

$$\epsilon_{\xi} = \begin{cases} 1, & \text{where } \xi \text{ is } D\text{-salient,} \\ -1 & \text{where } \xi \text{ is } D\text{-entering.} \end{cases}$$

In fact, this is how Minguzzi deduces his divergence theorem (Theorem 2 in [19]). Note, however, that he does it under the hypothesis of vanishing mean Cartan tensor ($C^{\mathrm{m}} = 0$), which implies that $d\Sigma_V^{\xi}$ is independent of V. As we do not require this, Theorem 2 is more general statement than Minguzzi's.

(iii) *The Finslerian unit normal presents some issues in the general case, as we are not taking $A = \mathrm{T}M \setminus 0$. In our physical interpretation, with L Lorentz–Finsler, A consists of timelike vectors, so asking for a Finslerian unit normal is only reasonable when Γ is L-spacelike, that is, $T_p \Gamma \cap (A \cap \partial A) = \varnothing$ for $p \in \Gamma$. In such a case, the strong concavity of the indicatrix $\{v \in A_p \colon L(v) = 1\}$ guarantees the existence and uniqueness of ξ: one defines ξ_p to be the unique vector such that $T_p \Gamma + \xi_p$ and the indicatrix are tangent at ξ_p.*

(iv) *Of course, if L comes from a pseudo-Riemannian metric on M, then $\xi = \epsilon_{\xi} \widehat{N}_V = \epsilon_{\xi} \widehat{N}$ and $d\Sigma_V^{\xi} = \epsilon_{\xi} d\sigma_V = \epsilon_{\xi} d\sigma$.*

(v) *It should be clear from this discussion that the form that one integrates on the right-hand side of (36) is always the same and that the only difference between Rund's and Minguzzi's divergence theorems is how each of them represents it. Notwithstanding, this is an important difference, for the boundary terms (38) and (39) could potentially have different physical interpretations.*

5. Divergence of Anisotropic Tensor Fields

Our developments of the previous section will allow us to obtain integral Finslerian conservation laws for a tensor T with $\operatorname{div}(T) = 0$. We obtain one for each $V \in \mathfrak{X}^A(U)$ satisfying certain hypotheses. Physically, T can be interpreted as an anisotropic stress–energy tensor and V as an observer field. We will also revisit two of the main examples with a clearer physical interpretation: Special Relativity and the conservation of the "total energy of the universe." In order to do all this, let us see how the Chern connection enters the Finslerian definition of $\operatorname{div}(T)$.

5.1. Definition of Divergence with the Chern Connection

Proposition 2 motivates the most natural definition of divergence of $T \in \mathcal{T}_1^1(M_A)$. Namely, by analogy with the classical case, we shall require (11) to hold for any anisotropic vector field $X \in \mathfrak{X}(M_A)$. This makes the Chern connection appear now: it is the only Finslerian connection ∇ for which one can assure that (32) holds independently of $Z := T(X)$. We shall also explore the conditions under which the term $\operatorname{trace}(\nabla Z)$ vanishes in the general Finslerian setting.

Proposition 3. *Let L be a fixed pseudo-Finsler metric defined on A with metric nonlinear connection $\mathrm{H}A$ and Chern anisotropic connection ∇. Also, let $S \in \mathcal{T}_2^0(M_A)$ be symmetric, $v \in A$, $T \in \mathcal{T}_1^1(M_A)$ and $X \in \mathfrak{X}(M_A)$.*
(A) The following are equivalent.

(Ai)　　　$S_v(-, \nabla^v_- X)$ *is antisymmetric.*
(Aii)　　$\nabla^v X$ *is anti-self-adjoint with respect to S_v, that is, $S_v(\nabla^v_- X, -) = -S_v(-, \nabla^v_- X)$.*
(Aiii)　　$\left(\mathfrak{l}^{\mathrm{H}}_X S\right)_v = \nabla^v_X S$.
(B) One has

$$\mathrm{div}(T(X)) - \mathrm{trace}(T(\nabla X)) = \mathbf{C}^1_2(\nabla T)(X),$$

where $\mathbf{C}^1_2$ is the operator that contracts the contravariant index with the covariant one introduced by ∇.
(C) One has $\mathrm{trace}(T(\nabla X))(v) = 0$ assuming any of the following conditions.

(Ci)　　$T^\flat_v(-, \nabla^v_- X)$ *is antisymmetric.*
(Cii)　　$T^\flat_v$ *is symmetric and $\left(\mathfrak{l}^{\mathrm{H}}_X g\right)_v = 0$.*

Proof. For (A), take $Y, W \in \mathfrak{X}(M)$. The antisymmetry of $S_v(-, \nabla^v_- X)$ reads

$$S_v(\nabla^v_Y X, W) = S_v(W, \nabla^v_Y X) = -S_v(Y, \nabla^v_W X),$$

which is exactly the anti-self-adjointness of $\nabla^v X$ with respect to S_v. Besides, (26) and (21) together with $\mathrm{Tor} = 0$ for the Chern connection give

$$
\begin{aligned}
& \mathfrak{l}^{\mathrm{H}}_X S(Y, W) \\
&= X^{\mathrm{H}}(S(Y, W)) - S(\mathfrak{l}^{\mathrm{H}}_X Y, W) - S(Y, \mathfrak{l}^{\mathrm{H}}_X W) \\
&= X^{\mathrm{H}}(S(Y, W)) - S(\nabla_X Y - \nabla_Y X, W) - S(Y, \nabla_X W - \nabla_W X) \\
&= \nabla_X S(Y, W) + S(\nabla_Y X, W) + S(Y, \nabla_W X),
\end{aligned}
\tag{40}
$$

which shows that $\left(\mathfrak{l}^{\mathrm{H}}_X S\right)_v = \nabla^v_X S$ also is equivalent to the anti-self-adjointness.

　　For (B), all the computations in (11) hold formally the same in the general Finslerian case due to Proposition 2.

　　As for the vanishing of $\mathrm{trace}(T(\nabla X))(v)$, it follows from (Ci) by the same computations as in (12). Indeed, the antisymmetry can be expressed as

$$T_{lj}(v)\nabla_i X^j(v) + T_{ij}(v)\nabla_l X^j(v) = 0.$$

It also follows from (Cii) by (13). Indeed, $\left(\mathfrak{l}^{\mathrm{H}}_X g\right)_v = 0$ is equivalent to $\nabla^v X$ being anti-self-adjoint with respect to g_v, and this can be expressed as

$$g^{li}(v)\nabla_i X^j(v) + g^{ji}(v)\nabla_i X^l(v) = 0.$$

□

Remark 6 ($\mathfrak{l}^{\mathrm{H}}_X g$ and Finslerian Killing fields)**.** *In classical relativity (g, T, and X isotropic), the second condition in (C ii) above would read $(\mathfrak{L}_X g)_{\pi(v)} = 0$, and $\mathfrak{L}_X g = 0$ would be equivalent to X being a Killing vector field. In the general case, X being Killing can be defined by the conditions $X \in \mathfrak{X}(M)$ and $\mathfrak{L}_X L = 0$ (Section 5 in [21]), but (using Theorem 1 (C), the facts that $\dot\partial C = \mathrm{Id}$ and $C(C, -, -) = 0$, and also (40))*

$$\mathfrak{L}_X L = \mathfrak{L}_X(g(C,C))$$
$$= \mathfrak{L}_X g(C,C) + 2g(\mathfrak{L}_X C, C)$$
$$= \left(\iota_X^{\mathrm{H}} g - \dot{\partial}_{\iota_X^{\mathrm{H}} C} g \right)(C,C) + 2g(\iota_X^{\mathrm{H}} C - \dot{\partial}_{\iota_X^{\mathrm{H}} C} C, C)$$
$$= \iota_X^{\mathrm{H}} g(C,C) - 2C(C,C,\iota_X^{\mathrm{H}} C) + 2g(\iota_X^{\mathrm{H}} C - \iota_X^{\mathrm{H}} C, C)$$
$$= \iota_X^{\mathrm{H}} g(C,C)$$
$$= \nabla g(C,C) + g(\nabla_C X, C) + g(C, \nabla_C X)$$
$$= 2g(C, \nabla_C X)$$

This way, we see that neither of X being Killing or $\iota_X^{\mathrm{H}} g = 0$ implies the other, and additionally we recover the characterization of (Proposition 6.1 (i) in [33]).

Definition 4. *Let L be a fixed pseudo-Finsler metric defined on A with (Levi-Civita–)Chern anisotropic connection ∇. For $T \in \mathcal{T}_1^1(M_A)$, its* divergence *with respect to L is defined as*

$$\mathrm{div}(T) := \mathbf{C}_2^1(\nabla T) \in \mathcal{T}_1^0(M_A) = \Omega_1(M_A),$$

where $\mathbf{C}_2^1$ is the operator that contracts the contravariant index with the covariant one introduced by ∇. In coordinates,

$$\mathrm{div}(T)_j = \nabla_i T_j^i = \delta_i T_j^i + \Gamma_{ik}^i T_j^k - \Gamma_{ij}^k T_k^i \tag{41}$$

for the Christoffel symbols of (5).

Remark 7 (Divergence vs. raising and lowering indices).

(i)   *First and foremost, by construction, (11) indeed holds for any $X \in \mathfrak{X}(M_A)$. At this point, it is important that the connection with which one defines $\mathrm{trace}(\nabla X)$ is the Chern one.*

(ii)   *Thanks to the fact that the Chern connection parallelizes g, namely, $\nabla_k g_{ij} = 0$ and $\nabla_k g^{ij} = 0$, the following hold:*

$$g^{ik} \nabla_k T_{ij} = g^{ik} g_{il} \nabla_k T_j^l = \nabla_k T_j^k = \mathrm{div}(T)_j, \tag{42}$$

$$\nabla_i T^{ij} = \nabla_i T_l^i g^{lj} = g^{jl} \mathrm{div}(T)_l. \tag{43}$$

This means that one could define the divergences of $S \in \mathcal{T}_2^0(M_A)$ and $R \in \mathcal{T}_0^2(M_A)$ straightforwardly,[9] $\mathrm{div}(S) = \mathbf{C}_{1,3}(\nabla S) \in \mathcal{T}_1^0(M_A) = \Omega_1(M_A)$ and $\mathrm{div}(R) = \mathbf{C}_1^1(\nabla R) \in \mathcal{T}_0^1(M_A) = \mathfrak{X}(M_A)$, and then (42) and (43) would read, respectively

$$\mathrm{div}(T^\flat) = \mathrm{div}(T),$$

$$\mathrm{div}(T^\sharp) = \mathrm{div}(T)^\sharp.$$

(iii)   *Regardless of this, in general we are not assuming the symmetry of $T^\flat$ or $T^\sharp$—we only did in Proposition 3 (Cii). Instead, at the beginning of §5 we fixed a convention for the order of the indices in T_{ij} and T^{ij} (for example, $T^\flat(X,Y) = g(X,T(Y)) \neq g(T(X),Y)$)—in the remainder of §4 and with said condition (Cii) only.*

5.2. Chern vs. Berwald

One needs to keep in mind a discussion present in [20]. The metric connection HA is the underlying nonlinear connection of an infinite family of A-anisotropic connections ∇. One of them is the (Levi–Civita)–Chern connection of L, which is the horizontal part of Chern–Rund's and Cartan's classical connections and has Christoffel symbols (5). All the others are this one plus an anisotropic tensor $Q \in \mathcal{T}_2^1(M_A)$ with $Q(-,C) = 0$ when viewed as an $\mathcal{F}(A)$-bilinear map $\mathfrak{X}(M_A) \times \mathfrak{X}(M_A) \to \mathfrak{X}(M_A)$. In particular, for $Q = -\mathrm{Lan}^\sharp$, one

gets the Berwald anisotropic connection of *L*, which is the horizontal part of Berwald's and Hasiguchi's classical connections and has Christoffel symbols (6). We did not as a priory select any of these ∇'s.

In some of the previous literature [34–37], the Finslerian divergence of vector fields was chosen to be defined directly with the Chern connection. In [18,19], the quantity trace(∇Z), with ∇ the Chern anisotropic connection, was referred to as the divergence of *Z*, though only after it had appeared in the divergence theorem. We have proven that the most natural definition leads to this characterization, hence clarifying why using Chern's covariant derivative is not arbitrary. Moreover, we have seen that said derivative fulfills the natural requisite (11) and is compatible with the lowering and raising of indices; these are key properties when it comes to the stress–energy tensor *T*. Still, it is important to compare this with what happens when one uses the other most natural covariant derivative: Berwald's.

Remark 8 (Divergence in terms of the Berwald connection)**.** *Let ∇ be the Chern anisotropic connection of L, with Christoffel symbols (5), and $\widehat{\nabla}$ be the Berwald one, with symbols (6).*

(i)   (33) *and* (41) *read respectively*

$$\mathrm{div}(Z) = \widehat{\nabla}_i Z^i + \mathrm{Lan}_k Z^k = \mathrm{trace}(\widehat{\nabla} Z) + \mathrm{Lan}^{\mathrm{m}}(Z),$$

$$\mathrm{div}(T)_j = \widehat{\nabla}_i T_j^i + \mathrm{Lan}_k T_j^k - \mathrm{Lan}_{ij}^k T_k^i$$
$$= \mathbf{C}_2^1(\widehat{\nabla} T)_j + \mathrm{Lan}^{\mathrm{m}}(T)_j - \mathbf{C}_1^1(\mathrm{Lan}^{\sharp}(T(-),-))_j,$$

where $\mathrm{Lan}^{\mathrm{m}}$ *is the mean Landsberg tensor (see* (7)*) and the contraction operators have the obvious meanings. Moreover, for* $X \in \mathfrak{X}(M_A)$

$$\mathrm{trace}(T(\nabla X)) = T_j^i \nabla_i X^j = T_j^i \widehat{\nabla}_i X^j + T_j^i \mathrm{Lan}_{ik}^j X^k$$
$$= \mathrm{trace}(T(\widehat{\nabla} X)) + \mathrm{trace}(\mathrm{Lan}^{\sharp}(T(-),X)),$$

which makes (11) *consistent with the previous formulas.*

(ii)   *One sees that the vanishing of* $\mathrm{Lan}^{\mathrm{m}}$ *(or of the mean Cartan* C^{m}*, see* ((6.37) *in* [38]*)) implies that the divergence of elements of* $\mathfrak{X}(M_A)$ *coincides with the trace of their Berwald covariant derivative. However,* $\mathrm{Lan}^{\mathrm{m}} = 0$ *(or even* $C^{\mathrm{m}} = 0$*) is not enough if one wants to obtain the same characterization for elements of* $\mathcal{T}_1^1(M_A)$*.*

Remark 9 (Sufficient conditions for $\mathfrak{l}_X^{\mathrm{H}} g = 0$ and being Finselrian Killing)**.** *In Remark 13 one could see that* $X \in \mathfrak{X}(M)$ *together with* $\nabla_C X = 0$ *is sufficient for X to be Killing. This condition does not privilege the Chern connection* ∇ *against the Berwald* $\widehat{\nabla}$*:*

$$\nabla_C X = \widehat{\nabla}_C X + \mathrm{Lan}^{\sharp}(C,X) = \widehat{\nabla}_C X$$

(see ((38) *in* [21]*), where* $\mathfrak{L}^{\flat}$ *is what here we would denote* $\mathrm{Lan}^{\sharp}$*). However, when it comes to the stress–energy tensor, we have seen that the relevant condition is not this but rather* $\mathfrak{l}_X^{\mathrm{H}} g = 0$*. Proposition 3 (A) implies that* $\nabla^v X = 0$ *is sufficient for* $(\mathfrak{l}_X^{\mathrm{H}} g)_v = 0$*, and this does privilege* ∇ *against* $\widehat{\nabla}$*.*

5.3. Finslerian Conservation Laws and Main Examples

Compare the results here with the classical case (14) and also with [19].

Corollary 1. *Let L be a fixed pseudo-Finsler metric defined on A. If*

(i) $X \in \mathfrak{X}(M_A)$ *is an anisotropic vector field,*
(ii) $V \in \mathfrak{X}^A(U)$ *is an A-admissible field with* $U \subseteq M$ *open,*
(iii) $T \in \mathcal{T}_1^1(M_A)$ *is an anisotropic 2-tensor, and*
(iv) $\overline{D} \subseteq U$ *is a domain with* ∂D *smooth up to subset of* 0 $(n-1)$*-dimensional measure on M and* $\mathrm{Supp}(X_V) \cap \overline{D}$ *compact,*

then

$$\int_D \text{div}(T)(X)d\text{Vol}_V + \int_D \text{trace}(T(\nabla X))_V d\text{Vol}_V$$
$$+ \int_D \left\{ C^{\mathbf{m}}(D_{T(X)}V) + \text{trace}(\dot{\partial}_{DV}T(X)) \right\} d\text{Vol}_V = \int_{\partial D} \imath_{T(X)_V}(d\text{Vol}_V), \tag{44}$$

where $C^{\mathbf{m}}$ is the mean Cartan tensor, and DV is computed with the metric nonlinear connection (4).

Proof. Just take $Z = T(X)$ in Theorem 2 and use part (B) or Proposition 3 . $\square$

Remark 10. *Observe that (44) allows for an interpretation of the divergence of T in terms of the flow in the boundary. Consider a sequence of domains D_m such that their volumes go to zero when $m \to +\infty$ and consider an observer V such that is infinitesimally parallel at $p \in M$, namely, $DV = 0$ in $p \in M$ and X such that $\nabla^v X = 0$. Then (44) and the mean value theorem imply that*

$$\text{div}(T)_v(X) = \lim_{m \to +\infty} \frac{1}{\text{Vol}_V(D_m)} \int_{\partial D_m} \imath_{T(X)_V}(d\text{Vol}_V).$$

In particular, $\text{div}(T)_v = 0$ can be interpreted as that the observer v measures conservation of energy in its restspace.

Corollary 2. *In the ambient of the previous corollary, assume:*

(i) $\text{div}(T)_V = 0$.
(ii)   *Any of the conditions (Ci) or (Cii) of Proposition 3 holds for $T_V^\flat$.*
(iii) $C^{\mathbf{m}}(D_{T(X)}V) + \text{trace}\{\dot{\partial}_{DV}(T(X))\} = 0$.

Then

$$\int_{\partial D} \imath_{T_V(X_V)}(d\text{Vol}_V) = 0. \tag{45}$$

Proof. It follows from Corollary 1, taking into account that the hypotheses (*i*), (*ii*), and (*iii*) imply that the three first integrals in (44) vanish. $\square$

Remark 11 (Sufficient conditions for the hypotheses (i), (ii), and (iii)).

(i)   *Obviously, $\text{div}(T) = 0$ suffices, but we do not need to assume that the divergence vanishes for all observers.*
(ii)   *$X = C$ suffices. In fact, $\nabla C = 0$ (Proposition 2.9 in [17]), so (Ci) of Proposition 3 holds for $T_V^\flat$. Thus, assuming the other two hypotheses, we get*

$$\int_{\partial D} \imath_{T_V(V)}(d\text{Vol}_V) = 0.$$

(iii)   *Although the hypothesis may seem artificial as it stands, there are a number of natural situations in which it is guaranteed. First, in classical relativity (g, T, and X isotropic), because $C^{\mathbf{m}} = 0$ and $\dot{\partial}(T(X)) = 0$; the result is then independent of V. Second, when the observer field is parallel ($DV = 0$), trivially. Third, when $DV = \theta \otimes V$ for some 1-form V and $T(X)$ is 0-homogeneous, because of Euler's theorem. Fourth, in the situation described in (Section 5.1 in [19]) (Z is our $T(X)$, s is our V, and I is our $C^{\mathbf{m}}$).*

Remark 12 (Representations of (45)). *One needs to keep in mind Remark 5. For a smooth part Γ of ∂D, one can use the (salient) Riemannian unit normal to represent*

$$\int_\Gamma \imath_{T_V(X_V)}(d\text{Vol}_V) = \int_\Gamma g_V(\widehat{N}_V, \widehat{N}_V)g_V(\widehat{N}_V, T_V(X_V))d\sigma_V$$
$$= \int_\Gamma g_V(\widehat{N}_V, \widehat{N}_V)T_V^\flat(\widehat{N}_V, X_V)d\sigma_V \tag{46}$$

when Γ is non-g_V-lightlike, and the Finslerian unit normal to represent

$$\int_\Gamma \iota_{T_V(X_V)}(d\mathrm{Vol}_V) = \int_\Gamma \epsilon_\xi L(\xi)g_\xi(\xi, T_V(X_V))d\Sigma_V^\xi$$

when L is Lorentz–Finsler and Γ is L-spacelike. This makes it possible to have the very same conservation law (45) *written in distinct ways, and in the examples below we will see that different expressions are preferable in different situations.*

In the remainder of the section, we analyze the Finslerian conservation laws in two settings in which L is Lorentz–Finsler. In particular, g has signature $(+, -, \ldots, -)$, A determines a time orientation, $L > 0$ on A, and (A, L) is maximal with these properties. We also have regularity conditions at ∂A, and in fact one sees that Theorem 2 and Corollary 2 still hold when allowing that $Z, X \in \mathfrak{X}(M_{\overline{A}})$, $T \in \mathcal{T}_1^1(M_{\overline{A}})$ and $V \in \mathfrak{X}^{\overline{A}}(U)$. Despite this, in both settings it will be necessary to take V as L-timelike, so the regularity at ∂A will not be used.

5.3.1. Example: Lorentz Norms on an Affine Space

In this example, we shall particularize Corollary 2 to the easiest Finslerian setting in which we can assure that its hypothesis (iii) holds. Namely, the structure of an affine space automatically provides an infinite number of parallel observer fields, $V \in \mathfrak{X}^A(M)$ with $DV = 0$.

To be preicse, suppose that $M = E$ is an affine space equipped with a *Lorentz norm* on an open conic subset $A_* \subseteq \vec{E} \setminus 0$ (a positive pseudo-Minkowski norm with Lorentzian signature in (Definition 2.11 in [23])). Under the usual identifications, such a norm can be seen as a Lorentz–Finsler L on $A \subseteq TE \setminus \mathbf{0} \equiv E \times (\vec{E} \setminus 0)$ that is independent of the first factor. Consequently, its fundamental tensor is nothing more than a Lorentzian scalar product g_v for each $v \in A_*$. The metric nonlinear connection of L coincides with the canonical connection of E, hence so do the Chern and Berwald anisotropic connections[10]. This is what implies that the parallel $V \in \mathfrak{X}^A(E)$ correspond exactly to the elements $v \in A_*$.

Let us introduce some notation. Given $(p_0, v) \in A$ with $L(v) = 1$, we can consider the Lorentzian scalar product g_v and the orthogonal hyperplane $\mathscr{R} := p_0 + \vec{\mathscr{R}} := p_0 + \left\{ w \in \vec{E} : g_v(v, w) = 0 \right\}$. We get an isometry $(t, p) \in R \times \mathscr{R} \mapsto p + tv \in E$, where $\mathscr{R}$ is equipped with $-g_v|_{\mathscr{R}}$ (a Euclidean scalar product), $R \times \mathscr{R}$ with $dt^2 + g_v|_{\mathscr{R}}$ (a Lorentzian one) and E with g_v. Let $\overline{\Omega}$ be a compact domain of $\mathscr{R}$ with $\partial\Omega \subseteq \mathscr{R}$ smooth up to a null $(n-2)$-dimensional measure set, and let $\widehat{n}_v$ be its salient unit $(-g_v|_{\mathscr{R}})$-normal. Then, for $t_0 < t_1$, the compact domain $\overline{D} \equiv [t_0, t_1] \times \overline{\Omega} \subseteq E$ has the required smoothness to apply Corollary 2, its boundary is $\partial D = \{t_1\} \times \overline{\Omega} \cup [t_0, t_1] \times \partial\Omega \cup \{t_0\} \times \overline{\Omega}$, and its salient g_v-normal is given by

$$\widehat{N}_v\Big|_{\{t_1\}\times\Omega} = v, \qquad \widehat{N}_v\Big|_{]t_0,t_1[\times\partial\Omega} = \widehat{n}_v, \qquad \widehat{N}_v\Big|_{\{t_0\}\times\Omega} = -v;$$

$$g_v(-v, -v) = g_v(v, v) = L(v) = 1,$$

$$g_v(\widehat{n}_v, \widehat{n}_v) = -(-g_v|_{\mathscr{R}})(\widehat{n}_v, \widehat{n}_v) = -1.$$

Remark 13. *For a $V \in \mathfrak{X}^A(E)$ identifiable with $v \in A_*$, we know that the hypothesis (iii) of Corollary 2 holds automatically. If (i) and (ii) hold too, then we get* (45), *for which we can use the representation* (46). *However, given the nature of the metric "nonlinear" and Chern "anisotropic" connections, it is easy to convince oneself that evaluating the result of anisotropic computations on this V is the same as first evaluating on V and then computing with isotropic tensors. For instance* $\mathrm{div}(T)_V = \mathrm{div}(T_V)$ *and* $\left(\mathfrak{l}_X^{\mathrm{H}} g\right)_V = \mathfrak{L}_{X_V}(g_V)$. *As a consequence, mathematically we get exactly the same conservation laws as if we just were in the Lorentzian affine space (E, g_v). Physically, though, different observers will measure different momenta.*

Corollary 3. *Let $V \in \mathfrak{X}^A(E)$ parallelly identifiable with an $v \in A_*$. If $T \in \mathcal{T}_1^1(E_A)$ is such that $\mathrm{div}(T_V) = 0$ and $X \in \mathfrak{X}(E_A)$ is such that $T_V^\flat(-, \nabla_-^V X)$ is antisymmetric, or $T_V^\flat$ is symmetric and $\mathfrak{L}_{X_V}(g_V) = 0$, then*

$$
\begin{aligned}
0 = & \int_{\{t_1\}\times\Omega} T_V^\flat(V, X_V) d\sigma_V - \int_{\{t_0\}\times\Omega} T_V^\flat(V, X_V) d\sigma_V \\
& - \int_{]t_0,t_1[\times\partial\Omega} T_V^\flat(\widehat{n}_V, X_V) d\sigma_V,
\end{aligned}
\tag{47}
$$

where $d\sigma_V$ is identifiable with the volume form of $-g_v|_\Omega$ on $\{t_\mu\} \times \Omega$ and coincides with the volume form of $g_v|_{]t_0,t_1[\times\partial\Omega}$ on $]t_0, t_1[\times \partial\Omega$.

Physically, even though Lorentz norms generalize Very Special Relativity [7], the classical interpretations of Special Relativity are still valid; we list them for completeness: v is an instantaneous observer at an event p_0, $\vec{\mathscr{R}}$ is its restspace and $\mathscr{R}$ is the *simultaneity hyperplane of v*, namely, the "universe at an instant, say $t = 0$, as seen by v." The affine space structure allows for a canonical propagation of v to all of the spacetime. Hence, if $\overline{\Omega}$ is a space region at $t = 0$, then $\overline{D}$ is the "evolution of $\overline{\Omega}$ along the time interval $[t_0, t_1]$ as witnessed by v." (47) expresses that *the variation after some time of the total amount of X_v-momentum in Ω is exactly equal to the amount of it that flowed across $\partial\Omega$*.

5.3.2. Example: Cauchy Hypersurfaces in a Finsler Spacetime

Here we present a construction which manifestly generalizes that of the previous example, again with straightforward physical interpretations, and we find an estimate that allows us to interpret (47) when $\partial\Omega$ is "at infinity". We will take $V \in \mathfrak{X}^A(U)$ with $U \subseteq M$ open, and we recall that we will assume the hypotheses of Corollary 2.

Suppose that the Finsler spacetime (M, L) is *globally hyperbolic*. By this, we mean that there is some (smooth, for simplicity) *L-Cauchy hypersurface $\mathscr{S} \subseteq M$*: every inextensible L-timelike curve $\gamma\colon I \to M$ (thus $\dot{\gamma}(t) \in A$) meets $\mathscr{S}$ exactly once. Let us assume that there are two L-spacelike Cauchy hypersurfaces $\mathscr{S}_0, \mathscr{S}_1 \subseteq U$ which do not intersect[11]. Then the results of [39] can be automatically transplanted: there exists a foliation by spacelike Cauchy hypersurfaces $M \equiv R \times \mathscr{S}$ such that $\mathscr{S}_0 \equiv \{t_0\} \times \mathscr{S}$ and $\mathscr{S}_1 \equiv \{t_1\} \times \mathscr{S}$. Taking the Finslerian unit normal ξ to each level $\{t\} \times \mathscr{S}$ produces an L-timelike field $\xi \in \mathfrak{X}^A(M)$. We can take this ξ to be our V, but we will not do so for the most part of this example.

Suppose also that $\{\overline{\Omega_{0,m}}\}$ is an exhaustion by compact domains of $\mathscr{S}_0$, namely, $\overline{\Omega_{0,m}} \subseteq \Omega_{0,m+1}$ and $\bigcup_{m\in\mathbb{N}} \Omega_{0,m} = \mathscr{S}_0$, such that $\partial\Omega_{0,m} \subseteq \mathscr{S}_0$ is smooth a. e. For $p \in \mathscr{S}_0$, let γ_p be the integral curve of V starting at p, which necessarily meets $\mathscr{S}_1$ at a unique instant $t_p \in R$. Put

$$
\Omega_{1,m} := \bigcup_{p\in\Omega_{0,m}} \gamma_p(\{t_p\}) \subseteq \mathscr{S}_1, \qquad \Gamma_p := \gamma_p\big[\min\{0, t_p\}, \max\{0, t_p\}\big],
$$

$$
D_m := \bigcup_{p\in\Omega_{0,m}} \Gamma_p \subseteq U, \qquad \Gamma_m := \bigcup_{p\in\partial\Omega_{0,m}} \Gamma_p.
$$

Remark 14. *By construction,*

(i) $\{\overline{\Omega_{1,m}}\}$ *is again an exhaustion by compact domains of $\mathscr{S}_1$ such that $\partial\Omega_{1,m} = \bigcup_{p\in\partial\Omega_{0,m}} \gamma_p(\{t_p\}) \subseteq \mathscr{S}_1$ is smooth a. e.*

(ii) $\overline{D_m}$ *is a compact domain of U with $\partial D_m = \overline{\Omega_{1,m}} \cup \Gamma_m \cup \overline{\Omega_{0,m}} \subseteq U$ smooth a. e. We do not really need to consider the union of all the D_m's.*

Next, for $Z \in \mathfrak{X}(M_A)$, we shall give the quantitative decay condition on (some components of) Z_V so that the integral

$$\int_{\Gamma_m} \imath_{Z_V}(d\mathrm{Vol}_V)$$

vanishes in the limit. The key fact for it will be that V is everywhere tangent to Γ_m (this is composed of γ_p's). In particular, as V is g_V-timelike, so must be Γ_m.

Remark 15. *The presence of V allows us to define an auxiliar Riemannian metric h_V on U with norm $\|-\|_V$, which gives a very natural way of quantifying. Namely, if $\{e_0 = V_p/F(V_p), e_1, \ldots, e_n\}$ is an orthonormal basis for g_{V_p}, then we prescribe it to be also h_{V_p}-orthonormal; equivalently,*

$$h_{V_p}(u, w) = 2g_{V_p}(u, \frac{V_p}{F(V_p)})g_{V_p}(w, \frac{V_p}{F(V_p)}) - g_{V_p}(u, w).$$

Then, by construction:
(i) *The volume form of h_V coincides with that of g_V, namely $d\mathrm{Vol}_V$.*
(ii) *The salient unit h_V-normal to Γ_m coincides with the corresponding g_V-normal. We denote it by $\widehat{N}_V$, as in Remark 12.*
(iii) *The hypersurface volume form of Γ_m with respect to h_V coincides with the one computed with g_V, namely $d\sigma_V = i_m^*(\imath_{\widehat{N}_V}(d\mathrm{Vol}_V))$ with $i_m : \Gamma_m \hookrightarrow U$ the inclusion. Hence we speak just of the hypersurface volume of Γ_m, namely $\sigma_V(\Gamma_m)$. As $\widehat{N}_V$ is g_V-orthogonal to V, and hence g_V-spacelike, we can use the representation*

$$\begin{aligned}
\int_{\Gamma_m} \imath_{Z_V}(d\mathrm{Vol}_V) &= \int_{\Gamma_m} g_V(\widehat{N}_V, \widehat{N}_V)g_V(\widehat{N}_V, Z_V)d\sigma_V \\
&= -\int_{\Gamma_m} g_V(\widehat{N}_V, Z_V)d\sigma_V.
\end{aligned} \tag{48}$$

Thanks to (48) and the fact that $g_V(\widehat{N}_V, V) = 0$, we intuitively see that if Z_V is proportional to V at infinity and the hypersurface volume does not grow too much, then the integral will be negligible. To be precise, we require that

$$K_m \sigma_V(\Gamma_m) \longrightarrow 0 \quad (m \longrightarrow \infty), \tag{49}$$

where

$$\begin{aligned}
K_m &:= \max_{\Gamma_m} \left\| Z_V - g_V(Z_V, \frac{V}{F(V)})\frac{V}{F(V)} \right\|_V \\
&= \max_{\Gamma_m} \left\{ \sqrt{g_V(Z_V, \frac{V}{F(V)})^2 - g_V(Z_V, Z_V)} \right\}.
\end{aligned}$$

Corollary 4. *In the above set-up, let $T \in \mathcal{T}_1^1(M_A)$, $X \in \mathfrak{X}(M_A)$, and $V \in \mathfrak{X}^A(U)$ be such that the hypotheses of Corollary 2 hold on all the D_m's, and put $Z := T(X)$. If the decay condition (49) holds too, then*

$$\int_{\Omega_{1,m}} \imath_{Z_V}(d\mathrm{Vol}_V) + \int_{\Omega_{0,m}} \imath_{Z_V}(d\mathrm{Vol}_V) \longrightarrow 0 \quad (m \longrightarrow \infty), \tag{50}$$

where $\Omega_{1,m}$ is constructed from $\Omega_{0,m}$ by intersecting the integral curves of V with $\mathscr{S}_1$.

Proof. Corollary 2 can be applied on $\overline{D_m}$, as $\mathrm{Supp}(Z_V) \cap \overline{D_m}$ is always compact. This and the representation (48) give

$$0 = \int_{\Omega_{1,m}} \iota_{Z_V}(d\mathrm{Vol}_V) + \int_{\Omega_{0,m}} \iota_{Z_V}(d\mathrm{Vol}_V) - \int_{\Gamma_m} g_V(\widehat{N}_V, Z_V) d\sigma_V. \tag{51}$$

Using the definition of h_V (Remark 15) and the Cauchy–Schwarz inequality,

$$
\begin{aligned}
0 &\le \left| \int_{\Gamma_m} -g_V(\widehat{N}_V, Z_V) d\sigma_V \right| \\
&\le \int_{\Gamma_m} \left| g_V(\widehat{N}_V, Z_V) \right| d\sigma_V \\
&= \int_{\Gamma_m} \left| g_V\left(\widehat{N}_V, Z_V - g_V\left(Z_V, \frac{V}{F(V)}\right) \frac{V}{F(V)}\right) \right| d\sigma_V \\
&= \int_{\Gamma_m} \left| -h_V\left(\widehat{N}_V, Z_V - g_V\left(Z_V, \frac{V}{F(V)}\right) \frac{V}{F(V)}\right) \right| d\sigma_V \\
&\le \int_{\Gamma_m} \left\| \widehat{N}_V \right\|_V \left\| Z_V - g_V\left(Z_V, \frac{V}{F(V)}\right) \frac{V}{F(V)} \right\|_V d\sigma_V \\
&= \int_{\Gamma_m} \left\| Z_V - g_V\left(Z_V, \frac{V}{F(V)}\right) \frac{V}{F(V)} \right\|_V d\sigma_V \\
&\le \int_{\Gamma_m} K_m d\sigma_V \\
&= K_m \sigma_V(\Gamma_m),
\end{aligned}
$$

so if $K_m \sigma_V(\Gamma_m)$ tends to 0, then so does the integral along Γ_m in (51). $\quad\square$

Remark 16. *In Corollary 4, if one of the integrals of $\iota_{Z_V}(d\mathrm{Vol}_V)$ along $\mathscr{S}_0$ or $\mathscr{S}_1$ exists in the Lebesgue sense, then so does the other and (50) reads*

$$\int_{\mathscr{S}_1} \iota_{Z_V}(d\mathrm{Vol}_V) + \int_{\mathscr{S}_0} \iota_{Z_V}(d\mathrm{Vol}_V) = 0.$$

Note that they could be $\pm\infty$, as we have not assumed, for instance, that Z_V is compactly supported in the union of all the D_m's. Rather, we have assumed the decay condition (49) alone.

Remark 17 (Sufficient conditions for (49)). *As for ensuring the decay condition, there are two possible scenarios.*

(i)   *The hypersurface volume $\sigma_V(\Gamma_m)$ stays bounded. Then, it is enough for (49) that $K_m \to 0$, and one could instead postulate the stronger condition that the maximum outside D_m tends to 0, which is independent of the concrete compact exhaustion.*

(ii)   *$\sigma_V(\Gamma_m)$ grows without bound. In this case, one can just postulate that the decay of K_m compensates the growth of $\sigma_V(\Gamma_m)$, but this does depend on the compact exhaustion.*

Notice that this is a purely Finslerian difficulty. Indeed, suppose that g, T, and X were isotropic and that $Z = T(X)$ was timelike. Then one could just set $V := Z$ and then carry out all the construction. Corollary 2 would be independent of the observer field (and its hypothesis (iii) would hold trivially), and $K_m = 0$ regardless of Γ_m. This is how we get the following statement of the classical law.

Corollary 5. *In the above se-up, suppose that L comes from a Lorentzian metric on M. Let $T \in \mathcal{T}_1^1(M)$ and $X \in \mathfrak{X}(M)$ be such that $\mathrm{div}(T) = 0$ and $T^\flat(-, \nabla_- X)$ is antisymmetric, or $T^\flat$ is symmetric and $\mathfrak{L}_X g = 0$. If $Z := T(X)$ is timelike, then*

$$\int_{\Omega_{1,m}} \iota_{Z_V}(d\mathrm{Vol}_V) + \int_{\Omega_{0,m}} \iota_{Z_V}(d\mathrm{Vol}_V) \longrightarrow 0 \quad (m \longrightarrow \infty),$$

where $\Omega_{1,m}$ is constructed from $\Omega_{0,m}$ by intersecting the integral curves of Z with $\mathscr{S}_1$.

Remark 18 (Conservation in terms of the Finslerian unit normal)**.**

(i)   *One could try to represent also the integrals of (50) in terms of $d\sigma_V$, as in Section 5.3.1. However, according to Remark 12, that would require assuming that $\mathscr{S}_\mu$ is non-g_V-lightlike, which is not very reasonable when all we know is that $\mathscr{S}_\mu$ L-spacelike and L-Cauchy.*

(ii)   *On the other hand, in terms of the Finslerian unit normal ξ, (50) reads*

$$\int_{\Omega_{1,m}} g_{\tilde{\xi}}(\xi, T_V(X_V)) d\Sigma_V^{\tilde{\xi}} - \int_{\Omega_{0,m}} g_{\tilde{\xi}}(\xi, T_V(X_V)) d\Sigma_V^{\tilde{\xi}} \longrightarrow 0 \qquad (52)$$

when $m \to \infty$. The sign in front of the second integral is explained as follows (see Remark 5 (ii)). $d\Sigma_V^{\tilde{\xi}}$ selects an orientation on each $\Omega_{\mu,m}$: the one for which $d\mathrm{Vol}_V(\xi, -, \dots, -)$ is positive. However, in (50) $\Omega_{1,m}$ already had an orientation $\mathfrak{D}_1$ and $\Omega_{0,m}$ had $\mathfrak{D}_0$: the D_m-salient ones. Necessarily[12], exactly one of these agrees with the $d\Sigma_V^{\tilde{\xi}}$-orientation: $\mathfrak{D}_1$ if $\mathscr{S}_1$ lays in the future of $\mathscr{S}_0$ and $\mathfrak{D}_0$ if it is the opposite. Notice that this, and hence (52), would fail if the Cauchy hypersurfaces crossed.

(iii)   *In the case $V = \xi$, (52) becomes*

$$\int_{\Omega_{1,m}} T_{\tilde{\xi}}^{\flat}(\xi, X_{\tilde{\xi}}) d\Sigma_{\tilde{\xi}} - \int_{\Omega_{0,m}} T_{\tilde{\xi}}^{\flat}(\xi, X_{\tilde{\xi}}) d\Sigma_{\tilde{\xi}} \longrightarrow 0,$$

a conservation law in which all the terms are purely Finslerian.

Summing up, in this example we have proven a Finslerian (observer-dependent) version of the classical law that *the total amount of X-momentum in the universe is conserved* (Corollary 4). Our formulation is asymptotic, so it is valid even for infinite total X_V-momentum (Remark 16). We have recovered the classical law (Corollary 5), which always holds under hypotheses on T and X alone, while in the general Finslerian case nontrivial difficulties appear in the regime of big separation between the Cauchy hypersurfaces (high $\sigma(\Gamma_m)$, Remark 17). Finally, we have expressed the law naturally in terms of the Finslerian unit normal (see (52)).

6. Conclusions

About the physical interpretation of T, Section 3:

1. *Heuristic interpretations from fluids*, Sections 3.1 and 3.2—Possible breakings of Lorentz-invariance lead to non-trivial transformations of coordinates between observers. Such transformations are still linear and permit a well-defined energy-momentum vector at each tangent space $\mathrm{T}_p M$, Section 3.1.

   However, the stress–energy–momentum T must not be regarded as a tensor on each $\mathrm{T}_p M$ but as an anisotropic tensor. This depends intrinsically on each observer $u \in \Sigma$ and may vary with u in a nonlinear way. Indeed, the breaking of Lorentz invariance does not permit to fully replicate the relativistic arguments leading to (isotropic) tensors on M, even though classical interpretations of the anisotropic T in terms of fluxes can be maintained, Section 3.2.

2. *Lagrangian viewpoint*, Section 3.3. In principle, the interpretations of Special Relativity about the canonical energy–momentum tensor associated with the invariance by translations remain for Lorentz norms and, thus, in Very Special Relativity. In the case of Lorentz–Finsler metrics, some issues to be studied further appear:

   (a)   The canonical stress–energy tensor in Relativity $\delta S_{matter}/\delta g^{\mu\nu}$ leads to different types of (anisotropic) tensors in the Finslerian setting (a scalar function $\delta S_{matter}/\delta L$ on $A \subseteq TM$ in the Einstein-Hilbert setting, higher order tensors in Palatini's). Starting at such tensors, different alternatives to recover the heuristic physical interpretations in terms of a 2-tensor appear.

(b)   In the particularly interesting case of a kinetic gas [1,16], the 1-PDF ϕ becomes naturally the matter source for the Euler-Lagrange equation of the Finslerian Einstein-Hilbert functional. However, the variational derivation of ϕ is obtained by means of a possibly non-natural Lagrangian. This might be analyzed by sharpening the framework of variational completion for Finslerian Einstein equations [15].

About the divergence theorem for anisotropic vector fields Z, Section 4:

1. Section 4.1: For any Lorentz–Finsler metric L, there is a natural definition of *anisotropic Lie bracket derivation along Z*, which depends only on the nonlinear connection HA and admits an interpretation by using flows.
2. Section 4.2: This bracket allows one to give a natural definition of $\text{div}(Z)$ which depends exclusively on HA and the volume form of L. This provides a geometric interpretation for the definition of divergence introduced by Rund [18].
3. A general divergence theorem is obtained (Theorem 2) so that Section 4.3:

    (a)   It can be seen as a conservation law for Z measured by each observer field V, even if the conserved quantity depends on V.

    (b)   The computation of the boundary term is intrinsically expressed in terms of forms. However, several metric elements can be used to re-express it, in particular the normal vector field for: (i) the pseudo-Riemannian metric g_V (Rund), or (ii) the pseudo-Finsler metric L, when L is defined on the whole TM (Minguzzi).

About the conservation of the stress–energy T Section 5:

1. Sections 5.1 and 5.2: The computation of $\text{div}(T)$ priviledges the Levi-Civita–Chern anisotropic connection, showing explicit equivalence with Rund's approach.
2. Corollarys 1 and 2: a vector field $T(X)_V$ on M is preserved assuming that some natural elements vanish on V for T, X and DV.
3. Section 5.3: Natural laws of conservation on Cauchy hypersurfaces under general conditions (including rates of decay for unbounded domains) can be obtained by a combination of the techniques (i) and (ii) in the item 3b above.

Appendix A. Kinematics: Observers and Relative Velocities

Here, we discusss a series of different possibilities for the notion of relative velocity between two observers, each one with a well-defined geometric construction. This is done as an academic exercise, because we do not discuss experimental issues (compare with [40,41]). However, it is worth emphasizing that all the possibilities studied here are intrinsic to the geometry of a flat model and, thus to any Finsler spacetime.

Start at an affine space endowed with a Lorentz norm let $u, u' \in \Sigma$ be two distinct observers and consider the plane $\Pi := \text{Span}\{u, u'\} \subset V$, which intersects transversally $\mathcal{C}$ and inherits a Lorentz–Finsler norm with indicatrix $\Sigma_\Pi := \Pi \cap \Sigma$. Recall that both tangent spaces $T_u \Pi$ and $T_{u'} \Pi$ inherit naturally a Lorentz scalar product by restricting the fundamental tensors g_u and $g_{u'}$, resp. Moreover, their (1-dimensional) restspaces $l := T_u \Sigma_\Pi$,

$l' := \mathrm{T}_{u'}\Sigma_\Pi$ also inherit a positive definite metric. In what follows, only the geometry of Π will be relevant.

Appendix A.1. The Lorentz Metric g_Π up to a Constant

Notice that $\Pi \cap \mathcal{C}_p$ is composed by two half-lines spanned by two $\mathcal{C}$-lightlike directions $w_\pm$; we will consider the orientation Π provided by the choice (w_+, w_-). One can determine a scalar product g_Π in Π (which is unique up to a positive constant), regarding both w_+ and w_- as g_Π-lightlike in the same causal cone. It is easy to check that Σ must be a strongly convex curve which converges asymptotically to the vector lines spanned by $w_\pm$. This implies both $u \in \Sigma$ will be timelike for g_Π and its restpace l will be g_Π-spacelike; we can assume also that the orientation l_+ in l is induced by the chosen w_+.

Notice that $g_u(u, w_\pm) \geq 0$ by the fundamental inequality, but $w_\pm$ might be timelike or spacelike for g_u (although $g_u(u, w_\pm) \to 0$ as $u \to w_\pm$). This possibility might be regarded as a possible measurement of the speed of light with respect to u by the observers in Π, namely, this velocity is in the orientation l_+ when w_+ is g_u-spacelike and smaller than 1 when it is timelike. However, a priori it is not clear an operational way to carry out such a measurement. Moreover such a measurement might be regarded as something non-intrinsic to the speed of light but to the way of measuring it.

Nevertheless, as pointed out in (Section 6 in [6]), there are several effects which might lead to a measurement of different speeds of light in different directions. So, we will consider that each Π has its own speeds of light $c_\Pi^\pm$ in each spacelike orientation $l_\pm$. Indeed, given u and an orientation l^+, the *speed of light* c_Π^+ will be defined as the the supremum of the relative velocities between u and all the observers u' such that $u' - u$ yields the orientation l^+. Next, we will explain several possible meanings of these velocities. To avoid cluttering, next we will write c_Π, assuming that the appropriate choice in $c_\Pi^\pm$ is done for each u'.

Appendix A.2. Simple Relative Velocity

As g_u determines naturally a Lorentz metric on V, we can define the *simple relative velocity* $v_u^s(u')$ of u' measured by u as the usual g_u-relativistic velocity between u, u' normalized to c_Π, i.e.,

$$v_u^s(u') = c_\Pi \tanh(\theta) \quad \text{where} \quad \cosh\theta = -g_u(u, u') > 1,$$

(the latter by the reversed fundamental inequality). Clearly, $v_{u'}^s(u) \neq v_u^s(u')$ in general, but this does not seem a drawback in the Finslerian setting.

A support for the physical plausibility of this velocity is that one could expect that each observer u will work as in Special Relativity just choosing an orthonormal frame of g_u. The possibility $g_u(v, v) \neq 1$ might seem ackward from a dynamical viewpoint (see below), but it seems harmless as far as only kinematics is being considered. In principle, the comparison between the measurements of the two observers would be geometrically possible by using the unique isometry of $(\mathrm{T}_u\Pi, g_u)$ to $(\mathrm{T}_{u'}\Pi, g_{u'})$ which maps u into u' and is consistent with orientations induced from Π. What is more, this isometry can also be extended to a natural isometry from (T_uV, g_u) to $(\mathrm{T}_{u'}V, g_{u'})$, namely, regard $(Section\;igma, g)$ as a Riemannian metric and use the parallel transport from u to u' along the segment of the curve $\Pi \cap \Sigma$ from u to u'. However, the following fact might suggest to explore further possibilities.

Remark A1. *Assume that Σ is modified into the indicatrix $\bar{\Sigma}$ of another Lorentz–Finsler norm so that (i) $\bar{\Sigma} = \Sigma$ around u and (ii) $u' \in \bar{\Sigma}$ but its $\bar{\Sigma}$ restspace $\bar{l}'$ is different from l'. Then, the simple velocity would remain unaltered, i.e., $\bar{v}_u^s(u') = v_u^s(u')$.*

Appendix A.3. Velocity as a Distance between Observers

Notice that Σ can be regarded as a Riemannian manifold with the restriction of the fundamental tensor g and, then, $\Sigma \cap \Pi$ can be regarded as a curve whose length can be computed. Then, the *observers' distance velocity* is defined as:

$$v^d(u, u') = c_\Pi \tanh\Big(\text{length}_g\{\text{segment of } \Sigma \cap \Pi \text{ from } u \text{ to } u'\}\Big).$$

Notice that this velocity is symmetric and it generalizes directly the one in Special Relativity providing a geometric interpretation for the addition of velocities. Recall that $v^d(u, u')$ has been defined essentially as a distance in $\Sigma \cap \Pi$, where Π depends of each pair of observers, thus, one might have $v^d(u, u') + v^d(u', u'') < v^d(u, u'')$ when $n > 2$. If one prefers to avoid such a possibility, it is enough to consider g-distance in the whole space of observers Σ (*observers' space distance velocity*), at least in the case that c_Π is regarded as independent of Π.

Remark A2. *In the case studied in Remark A1, one would have $\bar{v}^d(u, u') \neq v^d(u, u')$ in general. However, the relative position of the restspaces l and l' does not play any special role.*

Appendix A.4. Length-Contraction and Velocity

Consider a segment S of l with g_u-length ℓ and the strip of V obtained by translating S in the direction of u. Let S' be the intersection of this strip with l', which will be a new segment of $g_{u'}$-length ℓ'. Let $\lambda = \ell'/\ell$ be the *length-contraction parameter*. In the relativistic case, $\lambda < 1$ and $\lambda \to 0$ as $u' \to \mathcal{C}_\Pi$. The former property does not hold for a general Lorentz norm but the latter does. So, whenever $\lambda < 1$ holds, we can define the *length-contractive velocity* $v_u^c(u')$ of u' with respect to u as:

$$v_u^c(u') = c_\Pi \sqrt{1 - \lambda^2}.$$

Again, this velocity is not symmetric. Because of the strong convexity of Σ, a different observer u' will have a different restspace l', but this does not imply a different length ℓ' nor velocity $v_u^c(u')$. However, this velocity gives a comparison between restspaces which was absent in the previous two velocities.

Appendix A.5. Symmetric Lorentz Velocities in Π

Let us consider the Lorentzian scalar product g_Π en Π, uncfique up to a positive constant (which will be irrelevant for our purposes) introduced above. Recall that u and u' were timelike for g_Π and, moreover, both l and l' were spacelike. Now, we can define two velocities between u and u': the *simple Lorentz velocity*,

$$v^s(u, u') = c_\Pi \tanh(\theta) \quad \text{where} \quad \cosh\theta = -\frac{g_\Pi(u, u')}{\sqrt{g_\Pi(u, u)g_\Pi(u', u')}},$$

and the *length-contractive Lorentz velocity*,

$$v^c(u, u') = c_\Pi \tanh(\theta) \quad \text{where} \quad \cosh\theta = -\frac{|g_\Pi(n, n')|}{\sqrt{g_\Pi(n, n)g_\Pi(n', n')}},$$

where, in the latter, n, n' are g_Π-timelike vectors orthogonal to l, l', resp.

Clearly, both velocities are symmetric. Their appearance might be physically sound because the intrinsic Lorentz metric g_Π (up to a constant) can be regarded as an object available (or, at least, a compromise one) for all the observers, as it would depend directly on physical light rays.

Notes

[1] Berwald spaces [12,13] are an exception, as the parallel transport becomes an isometry between the Lorentz norms. Thus, in some sense, these spaces would admit a principle of equivalence with respect to a Lorentz-normed space (non-necessarily Lorentz–Minkowski).

[2] In this section, $i, j = 1, 2, 3$ and $\mu, \nu = 0, 1, 2, 3$, but in the others they will run freely from 1 to n ($= \dim M$).

[3] The symmetry of T is dropped for the case of theories with high spin because of its contribution to angular momentum.

[4] See for example (Section 4.5 in [24], Section 35 in [25]).

[5] See for example [26] (around formula (E.1.36)) or (Section 32 in [25]).

[6] See for example, (Section E.1 in [26]), (Section 4.3 in [27]), (Sections 21.2 and 21.3 in [28]).

[7] Some arguments which support strongly their choice are (see [1]): (a) the simplest analogous to the vacuum Einstein equation in the Finslerian approach Ricci= 0 (proposed by Rund [18], and satisfied by Finsler pp-waves [9]) is not a variational equation; (b) the Ricci scalar functional yields an Euler–Lagrange equation, which agrees with Einstein's in the vacuum Lorentz case, and (c) this Euler–Lagrange equation is the variational completion of the Finslerian Ricci= 0.

[8] This is not to be mistaken by the torsion of the nonlinear connection HA, which would have coordinates $N^i_{j\cdot k} - N^i_{k\cdot j}$ (even though this can be seen as a particular case of the torsion of some ∇ and hence it is also denoted by Tor in [17]).

[9] Here, $\mathbf{C}_{1,3}$ is the operator that (metrically) contracts the first index of S with the one introduced by ∇, and $\mathbf{C}^1_1$ is the operator that (naturally) contracts the first index of R with the one introduced by ∇.

[10] For instance, it is clear that in affine coordinates the components of the metric spray vanish, so the geodesics are the straight lines of E.

[11] The case when they interesect can be also conisdered by taking into account that, then, the open set $M \setminus J^+(S_1 \cup S_2)$ is still globally hyperbolic and a Cauchy hypersurface S_3 of this open subset will be also Cauchy for M (and it will not intersect any of the previous ones).

[12] Suppose, for instance, that $\mathscr{S}_1$ lays in the future of $\mathscr{S}_0$: the γ_p's departing from $\Omega_{0,m}$ reach points $\gamma_p(t_p) \in \Omega_{1,m}$ with $t_p > 0$. Take bases $(e_1, \ldots, e_{n-1})$ for $\mathrm{T}_p\Omega_{0,m}$ and $(e'_1, \ldots, e'_{n-1})$ for $\mathrm{T}_{\gamma_p(t_p)}\Omega_{1,m}$ such that $(V_p, e_1, \ldots, e_{n-1})$ and $(V_{\gamma_p(t_p)}, e'_1, \ldots, e'_{n-1})$ are dVol-positive. Then $(e_1, \ldots, e_{n-1})$ and $(e'_1, \ldots, e'_{n-1})$ are both $d\Sigma^\xi_V$-positive (ξ and V always lie in the same half-space), the former is $\mathfrak{O}_0$-negative (V is D_m-entering at $\mathscr{S}_0$) and the latter is $\mathfrak{O}_1$-positive (V is D_m-salient at $\mathscr{S}_1$).

References

1. Hohmann, M.; Pfeifer, C.; Voicu, N. Relativistic kinetic gases as direct sources of gravity. *Phys. Rev. D* **2020**, *101*, 024062.
2. Hohmann, M.; Pfeifer, C.; Voicu, N. Cosmological Finsler Spacetimes. *Universe* **2020**, *6*, 65.
3. Kouretsis, A.P.; Stathakopoulos, M.; Stavrinos, P.C. The General Very Special Relativity in Finsler Cosmology. *Phys. Rev. D* **2009**, *79*, 104011.
4. Li, X.; Chang, Z. Exact solution of vacuum field equation in Finsler spacetime. *Phys. Rev. D* **2014**, *90*, 064049.
5. Stavrinos, P.; Vacaru, O.; Vacaru, S. Modified Einstein and Finsler Like Theories on Tangent Lorentz Bundles. *Int. J. Mod. Phys. D* **2014**, *23*, 1450094.
6. Bernal, A.N.; Javaloyes, M.A.; Sánchez, M. Foundations of Finsler Spacetimes from the Observers' Viewpoint. *Universe* **2020**, *6*, 39.
7. Bogoslovsky, G. A special-relativistic theory of the locally anisotropic space-time. *Il Nuovo C. B Ser.* **1977**, *40*, 99–115.
8. Cohen, A.G.; Glashow, S.L. Very special relativity. *Phys. Rev. Lett.* **2006**, *97*, 021601.
9. Fuster, A.; Pabst, C. Finsler pp-waves. *Phys. Rev. D* **2016**, *94*, 104072.
10. Gibbons, G.W.; Gomis, J.; Pope, C.N. General very special relativity is Finsler geometry. *Phys. Rev. D* **2007** *76*, 081701.
11. Kostelecký, V.A. Riemann-Finsler geometry and Lorentz-violating kinematics. *Phys. Lett. B* **2011**, *701*, 137–143.
12. Fuster, A.; Heefer, S.; Pfeifer, C.; Voicu, N. On the non metrizability of Berwald Finsler spacetimes. *Universe* **2020**, *6*, 64.
13. Fuster, A.; Pabst, C.; Pfeifer, C. Berwald spacetimes and very special relativity. *Phys. Rev. D* **2018**, *98*, 084062.
14. Javaloyes, M.A.; Sánchez, M. Finsler metrics and relativistic spacetimes. *Int. J. Geom. Methods Mod. Phys.* **2014**, *11*, 1460032.
15. Hohmann, M.; Pfeifer, C.; Voicu, N. Finsler gravity action from variational completion. *Phys. Rev. D* **2019**, *100*, 064035.
16. Hohmann, M.; Pfeifer, C.; Voicu, N. Finsler-based field theory—A mathematical foundation. *arXiv* **2021**, arXiv:2106.14965v1.
17. Javaloyes, M.A.; Sánchez, M.; Villasenor, F.F. The Einstein-Hilbert-Palatini formalism in pseudo-Finsler geometry. *arXiv* **2021**, arXiv:2108.03197v2.
18. Rund, H. A divergence theorem for Finsler metrics. *Monatshefte Fur Math.* **1975**, *79*, 233–252.
19. Minguzzi, E. A divergence theorem for pseudo-Finsler spaces. *Rep. Math. Phys.* **2017**, *80*, 307–315.
20. Javaloyes, M.A.; Sánchez, M.; Villasenor, F.F. Anisotropic connections and parallel transport in Finsler spacetimes. *arXiv* **2021**, arXiv:2107.05986.
21. Javaloyes, M.A. Anisotropic tensor calculus. *Int. J. Geom. Methods Mod. Phys.* **2019** *16*, 1941001.
22. Javaloyes, M.A. Curvature computations in Finsler Geometry using a distinguished class of anisotropic connections. *Mediterr. J. Math.* **2020**, *17*, 123.

23. Javaloyes, M.A.; Sanchez, M.; On the definition and examples of cones and Finsler spacetimes. *Rev. Real Acad. Cienc. Exactas, FíSicas Nat. Ser. MatemáTicas* **2020**, *114*, 30.

24. Schutz, B.F. *A First Course in General Relativity*, 2nd ed.; Cambridge University Press: New York, NY, USA, 2009.

25. Landau, L.D.; Lifshitz, E.E. *The Classical Theory of Fields*, 3rd ed.; Pergamon Press Ltd., Headington Hill Hall, UK, 1971.

26. Wald, R.M.. *General Relativity*; University of Chicago Press: Chicago, IL, USA, 1984;

27. Carroll, S.M. *Spacetime and Geometry: An Introduction to General Relativity*; Addison-Wesley: San Francisco, CA, USA, 2004.

28. Misner, C.W.; Thorne, K.S.; Wheler, J.A. *Gravitation*; W. H. Freeman: San Francisco, CA, USA, 1973.

29. Pfeifer, C.; Wohlfarth, M. Causal structure and electrodynamics on Finsler space-times. *Phys. Rev. D* **2020**, *84*, 044039.

30. Gotay, M.J.; Marsden, J.E. Stress-Energy-Momentum Tensors and the Belinfante-Rosenfeld Formula. *Contemp. Math.* **1992**, *132*, 367–392.

31. Lee, J.M. *Introduction to Smooth Manifolds*; Springer: Berlin/Heidelberg, Germany, 2012.

32. Shen, Z. *Lectures on Finsler Geometry*; World Scientific: Singapore, 2001.

33. Herrera, J.; Javaloyes, M.A.; Piccione, P. On a monodromy theorem for sheaves of local fields and applications. *Rev. Real Acad. Cienc. Exactas, FíSicas Nat. Ser. MatemáTicas* **2017**, *111*, 999–1029.

34. Dragomir, S.; Larato, B. Harmonic functions on Finsler spaces. *Instanbul Univ. Fen Fak. Mat. Der.* **1991**, *48*, 67–76.

35. Mbatakou, J.S.; Todjihounde, L. Conformal change of Finsler-Ehresmann connections *Appl. Sci.* **2014**, *16*, 32–47.

36. Nibaruta, G.; Degla, S.; Todjihounde, L. Finslerian Ricci deformation and conformal metrics. *J. Appl. Math. Phys.* **2018**, *6*, 1522–1536.

37. Nibaruta, G.; Nibirantiza, A.; Karimumuryango, M.; Ndayirukiye, D. Divergence lemma and Hopf's theorem on Finslerian slit tangent bundle. *Balk. J. Geom. Its Appl.* **2020**, *25*, 93–103.

38. Shen, Z. *Differential Geometry of Spray and Finsler Spaces*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2001.

39. Bernal, A.N.; Sanchez, M. Further results on the smoothability of Cauchy hypersurfaces and Cauchy temporal functions. *Lett. Math. Phys.* **2006**, *77*, 183–197.

40. Lammërzahl, C.; Perlick, V. Finsler geometry as a model for relativistic gravity. *Int. J. Geom. Methods Mod. Phys.* **2018**, *15* (Supp. 1), 1850166.

41. Pfeifer, C. Finsler spacetime geometry in Physics. *Int. J. Geom. Methods Mod. Phys.* **2019**, *16* (Supp. 2), 1941004.