University of Naples Federico II - Italy

Department of Structural Engineering

Continuum Mechanics on Manifolds

Giovanni Romano

with the collaboration of

Raffaele Barretta

March 21, 2009

Contents

Figure 1: Giovanni Romano (1941 -). Full professor of Structural Mechanics, Department of Structural Engineering, University of Naples Federico II - e-mail: romano@unina.it. A trip from Helsinki to Tallin in 2004

Premise

This book collects the results of a research activity developed by the first author with the aim to provide a modern presentation of the basics of Continuum Mechanics. Our variational approach is based on the paradigmatic assumption that kinematics provides the primary description of the mathematical model while statics stems out by duality. To develop a sufficiently general formulation a differential geometric approach is compelling. Since a knowledge of this fascinating field of mathematics is not in the tool box of most graduated engineers and applied mathematicians, we have tried to provide a sufficiently exhaustive presentation of the subject in the first chapter. The treatment is however limited to foundational concepts and to results and methods that have found more direct application in the presentation of Continuum Mechanics. Many original results contributed in the course of this research activity have been included in the book. Some of them concern the geometric preliminaries and among them we quote the noteworthy general derivation of the curvature formula for a general nonlinear connection in a fibre bundle. Another chapter in which the geometric method has a primary role is the one dedicated to Dynamics and Geometric Optics. New results provided in this field include the very formulation of the geometric action principle, in which the fixed end condition has been ruled out, a new, more general, formulation of FERMAT principle of least optical lenght and of MAUPERTUIS' least action priciple which are considered as the prototypes of variational formulations in Mathematical Physics. A precise on some statements concerning the HAMILTON-JACOBI equation for non differentiable Lagrangians is also contributed. The subsequent chapters of the book include the treatment of some basic issues in Continuum Mechanics that have been investigated in greater detail by the first author. The collaboration of the second author has been valuable and fruitful on the whole of the topics dealt with and this book would not have been written without his precious assistance. The first more or less complete edition may be fixed around 2007 but much material may be dated much earlier and some is still under construction. Any comments, suggestions and corrections are very welcome.

Naples, March 2009

Giovanni Romano

Chapter 1

Calculus on manifolds

1.1 Introduction

In this chapter we introduce basic elements and results of differential geometry which provide the essential tools for the analysis of continuous bodies whose kinematics is defined on submanifolds of a larger ambient manifold. In the mechanics of continuous bodies the kinematical aspects are the fundamental issues on which the subsequent theoretical developments are built up by introducing dual quantities such as force systems and stress fields. Duality means that the interaction is a virtual work. The material behavior is also described in terms of kinematical quantities, which provide a suitable geometric measure of the deformation, and of their dual counterparts.

A general approach to the geometric description of the kinematics in terms of differentiable manifolds is of the utmost importance to deal with continuous models even for the classical CAUCHY $3D$-continuum. Lower dimensional continuous models have been formulated for the analysis of cables, membranes and shells, whose placements in the ambient eucldean space are described by one or two-dimensional differentiable submanifolds. Other important and useful models, mainly adopted for computational purposes, are the polar models of beams, shells, and $3D$-polar continua, whose placements are described by a special kind of manifold, a fibre bundle. The analysis of these models requires a deeper knowledge of the elements of differential geometry. We present here an organized collection of items in differential geometry for subsequent reference in the development of mechanical models.

The treatment begins with the introduction of the concept of a differentiable manifolds, which generalizes the classical idea of a regular curve or surface, and of the relevant tangent and cotangent vector bundles. We then introduce the push forward and pull back transformations of scalar, vector, co-vector and tensor fields, according to a diffeomorphic correspondence between manifolds, and the notion of the flow generated by a vector field on a manifold.

Existence,uniqueness and continuous dependence on the initial condition is proven by relying on two classical tools of the theory of ordinary differential equations, the PICARD-BANACH's fixed point method and the GRÖNWALL inequality. In this context a special kind of derivative for functionals on vector bundles is defined, the *fibre-derivative.* It provides a connection between tangent and cotangent bundles and extends to differentiable manifolds a classical tool in physics: the LEGENDRE-FENCHEL transform of convex potentials.

A basic kind of derivation, unfortunately not included in standard courses of calculus, emerges as a cornerstone for a proper understanding of the kinematics of continua: the LIE derivative, also dubbed the *fisherman derivative*, introduced by MARIUS SOPHUS LIE in the last decades of the eighteenth century.

The LIE derivative or convective derivative plays a basic role in classical physics and its magic properties will fascinate the interested reader.

Another basic kind of differentiation emerges in the study of the modern theory of integration on manifolds: the *exterior derivative* of a differential form. This notion stems out from a direct extension of the fundamental theorem of calculus on the real line to the multidimensional case. It provides a compact and expressive formula for the transformation of the integral of a volume form on the boundary of a chain into an integral of a volume form on the chain, the celebrated GABRIEL STOKES fomula.

Other important and useful results of calculus on manifolds are then illustrated such as the FUBINI's theorem, the POINCARÉ lemma and the homotopy formula, also called the HENRI CARTAN's magic formula, which provides a relation between the LIE derivative and the exterior derivative of a differential form. It is also shown how STOKES fomula generates all the classical integral transformation formulas and how the notion of exterior derivative is an effective tool to get the expressions of gradient, curl and divergence in general curvilinear coordinates. Last but not least, general formulations of OSBORNE REYNOLDS transport theorem for flowing manifolds are provided.

The basic notion of a connection on a manifold is then introduced and illustrated in the general setting of fibre bundles due to CHARLES EHRESMANN who introduced it in 1950 . Here a third basic kind of derivative is introduced: the *covariant derivative.*

In this context an original proof, of the result which provides the expression of curvature of a connection on a fibre bundle in terms of covariant derivatives, is contributed.

The special properties of connections and of covariant derivatives on vector bundles and on tangent bundles are then illustrated and the notion of torsion is introduced.

The treatment then turns to a presentation of the wonderful idea conceived by BERNHARD RIEMANN who, in his dissertation for Habilitation *Über die Hypothesen welche der Geometrie zu Grunde liegen* (On the hypotheses that lie at the foundations of geometry), delivered on 10 June 1854, at the presence of GAUSS, introduced the notion of a differential geometric object, a differentiable manifold, endowed with a regular field of metric tensors providing lenght measurements of the vectors of each of its tangent spaces.

This is in fact the basic concept for the general description and the investigation of the deformation of continuous bodies. The end of the chapter is then devoted to illustrate a generalization of the euclidean translation to differentiable manifold: the connection between the tangent spaces.

The most effective representation of this concept is due to GREGORIO RICCI-CURBASTRO and TULLIO LEVI-CIVITA who, at the very beginning of the nineteenth century, introduced the idea of the parallel transport on a riemannian manifold. The LEVI-CIVITA connection between tangent spaces of a riemannian manifold enjoys peculiar properties that provide a relation between the LIE derivative and the covariant derivative.

The RIEMANN-CHRISTOFFEL curvature tensor field yields the test to discover if a $3D$ riemannian manifold can be embedded in the euclidean space and provides the answer to the question of the kinematic compatibility of the metric changes induced by elastic and anelastic strains in a continuous body. Comprehensive treatments of differentiable and riemannian manifolds can be found in [81], [82] [180], [3], [106], [143], [80], [91] and in the references therein.

1.1.1 Duality and metric tensor

The response of a continuous material body to given actions (such as force systems or temperature changes) is locally described by the changes in length between the material line-elements at the points of a reference placement and the corresponding ones at the corresponding points of the current placement. Material line-elements are geometrically described by tangent vectors to material lines drawn in the body thru the point of interest.

To define tangent vectors at a point we have to consider an arbitrary regular parametric representation of material lines and take the derivative at that point. The result will depend on the orientation of material lines so induced and on the travel speed along the lines at that point induced by the parametrization. Since orientation and speed are arbitrarily choosen, the mechanical interpretations of the geometrical construct should be checked to be independent of the choice.

A continuous body is described, in geometrical terms, as a fibre bundle having as base manifold the placement of the material particles of the body and as fibres the tangent spaces at each point.

Then the placement of a body is described by a tangent bundle whose elements are pairs made of a point and of a tangent vector at that point.

The proper geometrical tool, in measuring the length of material lines at a point, is a metric tensor. For completeness sake, we provide hereafter a summary of the basic properties of a metric tensor and of related issues.

Although it could seem not appropriate to present these concepts and definitions in a course on classical continuum mechanics, since them should be familiar to anyone sufficiently trained in geometry, the lack of precision and the introduction of undefined geometrical objects is often a main source of troubles in theoretical treatments. Indeed the mathematical modelling of the deformation of a continuous body requires a clear definition of the suitable geometrical tools and a precise statement of their basic properties.

1.1.2 The metric tensor

Mathematics tells us that the knowledge of the length of the vectors, in a linear space V of dimension n, leads to the definition of the inner product between vectors. In turn the symmetry and the bilinearity of the inner product reveals that a finite number of length informations suffices to have a complete metric description of the linear space: we need $C_2^n = (n+1)\,n/2$ independent length informations. To be precise, we know that the norm (or length) of a vector meets the classical axioms:

$$
\begin{aligned}
&i) && \|\mathbf{a}\| \geq 0\,, \quad \|\mathbf{a}\| = 0 \iff \mathbf{a} = 0\,,\\
&ii) && \|\mathbf{a}+\mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|\,, \quad \text{triangle inequality}\\
&iii) && \|\alpha\mathbf{a}\| = |\alpha|\|\mathbf{a}\|\,,
\end{aligned}
$$

for any $\mathbf{a}, \mathbf{b} \in V$ and $\alpha \in \Re$. A linear space V, even non finite dimensional, endowed with a norm and complete with respect to the induced topology is called a Banach space.

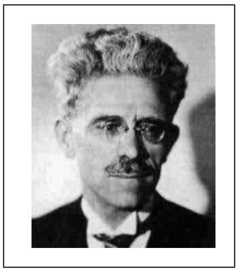

Figure 1.1: Maurice René Fréchet (1878 - 1973)

Figure 1.2: Pascual Jordan (1902 - 1980)

A noteworthy theorem by M. FRÉCHET, J. von NEUMANN and P. JORDAN (see [192], theorem I.5.1) states that, if the norm fulfills the *parallelogram law*:

$$\|\mathbf{a}+\mathbf{b}\|^2 + \|\mathbf{a}-\mathbf{b}\|^2 = 2\,(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2)\,,$$

the inner product of $\mathbf{a},\mathbf{b} \in V$ may be defined by the *polarization formula*:

$$\mathbf{g}(\mathbf{a},\mathbf{b}) := \frac{1}{4}(\|\mathbf{a}+\mathbf{b}\|^2 - \|\mathbf{a}-\mathbf{b}\|^2)\,,$$

which is indeed a symmetric, bilinear and positive definite *metric tensor* $\mathbf{g} \in BL\,(V^2\,;\Re)$.

By eliminating $\|\mathbf{a}+\mathbf{b}\|$ or $\|\mathbf{a}-\mathbf{b}\|$ between the parallelogram law and the polarization formula, we may rewrite the latter as

$$\mathbf{g}(\mathbf{a},\mathbf{b}) := \frac{1}{2}(\|\mathbf{a}+\mathbf{b}\|^2 - \|\mathbf{a}\|^2 - \|\mathbf{b}\|^2) = \frac{1}{2}(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - \|\mathbf{a}-\mathbf{b}\|^2)\,.$$

Figure 1.3: John von Neumann (1903 - 1957)

The converse implication is trivial. Indeed it is clear that, if a metric tensor $\mathbf{g} \in BL(V^2;\Re)$ is given, the norm defined by $\|\mathbf{a}\|^2 = \mathbf{g}(\mathbf{a},\mathbf{a})$ fulfils the parallelogram law.

The metric tensor $\mathbf{g} \in BL(V^2;\Re)$ provides the informations concerning the length of any vector and the cosinus of the angle between any two (non-zero) vectors, according to the definition

$$\cos(\mathbf{a},\mathbf{b}) := \mathbf{g}(\mathbf{a},\mathbf{b})/(\|\mathbf{a}\|\|\mathbf{b}\|)\,.$$

By virtue of the CAUCHY-SCHWARZ inequality:

$$\|\mathbf{a} + \lambda\,\mathbf{b}\| \geq 0\,, \quad \forall\,\lambda \in \Re \iff |\mathbf{g}(\mathbf{a},\mathbf{b})| \leq \|\mathbf{a}\|\|\mathbf{b}\|\,,$$

the absolute value of the cosinus is not greater than unity and equal to unity if and only if the two vectors are proportional one another. A non finite dimensional linear space V with a metric tensor is called a pre-HILBERT space. If complete with respect to the induced topology, it is said a HILBERT space.

In n dimensional spaces the knowledge of the $C_2^n = (n+1)\,n/2$ independent components of the metric tensor in a local frame is all that one needs to get a complete information on the geometric properties of the tangent space. If only length measurements are allowed, the complete information about the metric requires the knowledge of the length of the sides of a *nondegenerated symplex* at the point of interest.

A *symplex* in a n-dimensional linear space is the convex envelope of $n+1$ points and is *nondegenerated* if its volume is non zero. In the euclidean three-space length measurements of the sides of a tetrahedron suffice while in a two-dimensional space the sides of a triangle make the job.

In a n-dimensional linear space we need again $C_2^n = (n+1)\,n/2$ length measurements.

Indeed the edges of a nondegenerated symplex are formed by a basis of the n-dimensional linear space and by the differences between any (non-ordered) pair of basis vectors. The number of edges is thus again $n+n\,(n-1)/2 = (n+1)\,n/2$. If, for any pair of basis vector we know the lengths $\|\mathbf{a}\|$, $\|\mathbf{b}\|$ and $\|\mathbf{a}-\mathbf{b}\|$, by the parallelogram law we may compute

$$\|\mathbf{a}+\mathbf{b}\|^2 = 2\,(\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2) - \|\mathbf{a}-\mathbf{b}\|^2\,,$$

and hence $\mathbf{g}(\mathbf{a},\mathbf{b}) := \frac{1}{4}(\|\mathbf{a}+\mathbf{b}\|^2 - \|\mathbf{a}-\mathbf{b}\|^2)$ by the polarization formula.

1.1.3 Volume forms and invariants

Once a metric tensor $\mathbf{g}$ is at hand, related volume measurements can be performed by evaluating the inner product between the sides $\{\mathbf{e}_i\}$, $i,j = 1,\dots,n$ of an oriented parallelepiped, forming the corresponding GRAM *matrix* with entries $\mathrm{GRAM}_{ij} := \mathbf{g}(\mathbf{e}_i,\mathbf{e}_j)$ with $i,j = 1,\dots,n$, and taking the square root of its determinant. This formula for the n-linear alternating *volume form* $\boldsymbol{\mu}_{\mathbf{g}} \in BL\,(V^n\,;\Re)$ induced by the metric can be deduced by considering the GRAM *operator* which to any metric tensor $\mathbf{g}$ relates a matrix-valued mapping $\mathrm{GRAM}(\mathbf{g})$ acting on pairs of n-tuples $\{\mathbf{a}_i\},\{\mathbf{b}_j\}$, $i,j = 1,\dots,n$, which is n-linear and alternating separately in the two n-tuples:

$$\begin{aligned}\det \mathrm{GRAM}(\mathbf{g})\cdot\{\mathbf{a}_1,\dots,\mathbf{a}_n\}\cdot\{\mathbf{b}_1,\dots,\mathbf{b}_n\} &:= \det \mathbf{g}(\mathbf{a}_i,\mathbf{b}_j)\\ &= \boldsymbol{\mu}_{\mathbf{g}}\{\mathbf{a}_1,\dots,\mathbf{a}_n\}\,\boldsymbol{\mu}_{\mathbf{g}}\{\mathbf{b}_1,\dots,\mathbf{b}_n\}\,.\end{aligned}$$

Setting $\mathbf{a}_i = \mathbf{b}_i$, $i = 1,\dots,n$, we write $\det \mathrm{GRAM}(\mathbf{g}) = \boldsymbol{\mu}_{\mathbf{g}}^2$. The sign of the volume form is taken depending on the chosen positive orientation of the space.

If a standard signed volume form is fixed, all the others are proportional to it, as is easily verified. In a 3-dimensional space V, once the standard volume form has been chosen, the signed area of the parallelogram of sides $\mathbf{a},\mathbf{b} \in V$ is given by $\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{n},\mathbf{a},\mathbf{b})$ with $\mathbf{n}$ of unit length and orthogonal to $\mathbf{a},\mathbf{b} \in V$.

The sinus of the angle between any two (non-zero) vectors $\mathbf{a},\mathbf{b} \in V$ is then computed according to the relation

$$\sin(\mathbf{a},\mathbf{b}) := \frac{\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{n},\mathbf{a},\mathbf{b})}{\|\mathbf{a}\|\|\mathbf{b}\|}\,.$$

From the definition of the volume form it follows that

$$\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{n},\mathbf{a},\mathbf{b})^2 = \|\mathbf{a}\|^2\|\mathbf{b}\|^2 - \mathbf{g}(\mathbf{a},\mathbf{b})^2\,,$$

and we recover the well-known Pythagoras's theorem

$$\sin(\mathbf{a},\mathbf{b})^2+\cos(\mathbf{a},\mathbf{b})^2=1\,.$$

To any (bounded) linear operator $\mathbf{A}\in BL(V\,;V)$ we may associate n independent invariants which, for any volume form $\boldsymbol{\mu}$ on V, provide the ratios of the volumes of n sets of parallelepipeds in V, with respect to the volume of a given one, generated according to the rule (we set $n=3$):

$$J_1(\mathbf{A})\,\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3):=\boldsymbol{\mu}(\mathbf{A}\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)+\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{A}\mathbf{e}_2,\mathbf{e}_3)+\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{A}\mathbf{e}_3)\,,$$

$$J_2(\mathbf{A})\,\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3):=\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{A}\mathbf{e}_2,\mathbf{A}\mathbf{e}_3)+\boldsymbol{\mu}(\mathbf{A}\mathbf{e}_1,\mathbf{e}_2,\mathbf{A}\mathbf{e}_3)+\boldsymbol{\mu}(\mathbf{A}\mathbf{e}_1,\mathbf{A}\mathbf{e}_2,\mathbf{e}_3)\,,$$

$$J_3(\mathbf{A})\,\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3):=\boldsymbol{\mu}(\mathbf{A}\mathbf{e}_1,\mathbf{A}\mathbf{e}_2,\mathbf{A}\mathbf{e}_3)\,.$$

The volume ratios are called the *linear invariant* or *trace*, the *quadratic invariant* and the *cubic invariant* or *determinant* respectively, so that $\mathrm{tr}\mathbf{A}:=J_1(\mathbf{A})$ and $\det\mathbf{A}:=J_3(\mathbf{A})$.

For any linear isomorphism $\varphi\in BL(V\,;W)$, between two n-dimensional linear spaces, we have:

$$J_k(\mathbf{A})=J_k(\varphi\mathbf{A}\varphi^{-1})\,,\quad k=1,2,3\,,$$

which can be seen by choosing in W the volume form (we set $n=3$):

$$\boldsymbol{\mu}_\varphi(\varphi\,\mathbf{e}_1,\varphi\,\mathbf{e}_2,\varphi\,\mathbf{e}_3)=\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)\,,$$

and observing that (we consider only J_3):

$$\begin{aligned}(\det\mathbf{A})\,\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)=\boldsymbol{\mu}(\mathbf{A}\mathbf{e}_1,\mathbf{A}\mathbf{e}_2,\mathbf{A}\mathbf{e}_3)&=\boldsymbol{\mu}_\varphi(\varphi\mathbf{A}\mathbf{e}_1,\varphi\mathbf{A}\mathbf{e}_2,\varphi\mathbf{A}\mathbf{e}_3)\\&=\det(\varphi\mathbf{A}\varphi^{-1})\,\boldsymbol{\mu}_\varphi(\varphi\mathbf{e}_1,\varphi\mathbf{e}_2,\varphi\mathbf{e}_3)\\&=\det(\varphi\mathbf{A}\varphi^{-1})\,\boldsymbol{\mu}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)\,.\end{aligned}$$

Then the *invariants* of a linear operator $\mathbf{A}\in BL(V\,;V)$ are scalar valued homogeneous functions of $\mathbf{A}$ of degree $1,\dots,n$, which are invariant with respect to the choice of a volume form and with respect to linear isomorphic transformations of the linear space. The following properties can be easily verified:

$$\begin{aligned}\mathrm{tr}(\mathbf{ABC})&=\mathrm{tr}(\mathbf{BCA})=\mathrm{tr}(\mathbf{CAB})\,,\\J_2(\mathbf{A})&=\tfrac{1}{2}\,(J_1(\mathbf{A}^2)-J_1(\mathbf{A})^2)\,,\\\det(\mathbf{AB})&=(\det\mathbf{A})(\det\mathbf{B})\,.\end{aligned}$$

1.1.4 Transposition, isomorphisms and duality pairing

Let V be a BANACH space and V^* the dual space of bounded linear functions from V in $\Re$.

- A (p, q)-*tensor* on a linear space V is multilinear map from the cartesian product of p copies of V and q copies of V^* into a BANACH space E.

Most often the spaces V and V^* are finite dimensional and the space E is simply the real field $\Re$. The following are useful identifications.

A $(2, 0)$-tensor $\boldsymbol{\beta} \in BL(V^2; \Re)$ may be represented as an operator, denoted by $\boldsymbol{\alpha}^\flat \in BL(V; V^*)$, setting:

$$\boldsymbol{\alpha}(\mathbf{a}, \mathbf{b}) = \langle \boldsymbol{\alpha}^\flat(\mathbf{a}), \mathbf{b} \rangle , \quad \forall\, \mathbf{a}, \mathbf{b} \in V ,$$

By the reflexivity of the finite dimensional space V, we may identify the bidual linear space V^{**} with V. It follows that a tensor $\boldsymbol{\alpha}^* \in BL(V^{*2}; \Re)$ is identified with the operator $\boldsymbol{\alpha}^*_\sharp \in BL(V^*; V)$ by setting:

$$\boldsymbol{\alpha}^*(\mathbf{a}^*, \mathbf{b}^*) = \langle \boldsymbol{\alpha}^*_\sharp(\mathbf{a}^*), \mathbf{b}^* \rangle , \quad \forall\, \mathbf{a}^*, \mathbf{b}^* \in V^* .$$

A metric tensor $\mathbf{g} \in BL(V^2; \Re)$ provides a one-to-one correspondence between a tensor $\boldsymbol{\alpha} \in BL(V^2; \Re)$ and a pair of linear operators, $\mathbf{A}, \mathbf{A}^T \in BL(V; V)$, one the $\mathbf{g}$-*transpose* of the other, according to the relation

$$\boldsymbol{\alpha}(\mathbf{a}, \mathbf{b}) = \mathbf{g}(\mathbf{A}\,\mathbf{a}, \mathbf{b}) = \mathbf{g}(\mathbf{a}, \mathbf{A}^T\mathbf{b}) , \quad \forall\, \mathbf{a}, \mathbf{b} \in V .$$

The $\mathbf{g}$-transposition is an involutive relation, i.e. $(\mathbf{A}^T)^T = \mathbf{A}$ and transposed linear operarors have the same invariants: $J_k(\mathbf{A}^T) = J_k(\mathbf{A})$ for $k = 1, \ldots, n$.

Note however that, by changing the metric, the pair $\mathbf{A}, \mathbf{A}^T \in BL(V; V)$ and the invariants change too, so that invariants cannot be associated with a tensor $\boldsymbol{\alpha} \in BL(V^2; \Re)$, unless a metric tensor is specified.

A metric tensor $\mathbf{g} \in BL(V^2; \Re)$ induces a *linear isomorphism* between the space V and its dual V^*. Indeed to any vector $\mathbf{a} \in V$ we may associate uniquely the covector $\mathbf{a}^* \in V^*$ defined by

$$\mathbf{a}^* = \mathbf{g}\,\mathbf{a} \iff \langle \mathbf{a}^*, \mathbf{b} \rangle = \mathbf{g}(\mathbf{a}, \mathbf{b}) , \quad \forall\, \mathbf{b} \in V ,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between vectors and covectors, that is the value taken by the covector on the vector. We shall denote by the same symbol $\mathbf{g} = \mathbf{g}^\flat \in BL(V; V^*)$ the linear isomorphism that associates the covector

$\mathbf{g}\,\mathbf{a} \in V^*$ with the vector $\mathbf{a} \in V$. Conversely to any given covector $\mathbf{a}^* \in V^*$ there corresponds a unique vector $\mathbf{g}^{-1}\mathbf{a}^*$ which is the unique solution of the linear problem $\mathbf{g}\,\mathbf{a} = \mathbf{a}^*$. In fact the linear operator $\mathbf{g} \in BL(V\,;V^*)$ has a degenerated kernel and hence is surjective.

These isomorphisms are often denoted by the musical symbols *flat*: $\flat = \mathbf{g}$ and *sharp*: $\sharp = \mathbf{g}^{-1}$ but we will not follow this aesthetically pleasant symbolism because it doesn't keep track of the underlying metric.

The isomorphism between V and V^* becomes an isometry by endowing the dual space V^* with the metric tensor $\mathbf{g}^* \in BL(V^{*2}\,;\Re)$ defined as

$$\mathbf{g}^*(\mathbf{g}\,\mathbf{a}, \mathbf{g}\,\mathbf{b}) := \mathbf{g}(\mathbf{a}, \mathbf{b})\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in V\,.$$

Given a tensor $\boldsymbol{\gamma} \in BL(V^2\,;\Re)$ and a linear operator $\mathbf{A} \in BL(V\,;V)$ we define the tensor $\boldsymbol{\gamma}\mathbf{A} \in BL(V^2\,;\Re)$ by the formula

$$(\boldsymbol{\gamma}\mathbf{A})(\mathbf{a}, \mathbf{b}) := \boldsymbol{\gamma}(\mathbf{A}\mathbf{a}, \mathbf{b}) = \langle\, (\boldsymbol{\gamma}^\flat \circ \mathbf{A})\mathbf{a}, \mathbf{b}\,\rangle\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in V\,,$$

so that the metric-induced correspondence between the tensor $\boldsymbol{\alpha} \in BL(V^2\,;\Re)$ and the linear operator $\mathbf{A} \in BL(V\,;V)$ may be written simply as

$$\boldsymbol{\alpha}^\flat = \mathbf{g}\,\mathbf{A} \iff \mathbf{A} = \mathbf{g}^{-1}\boldsymbol{\alpha}^\flat\,.$$

For any $\mathbf{L} \in BL(V\,;V)$ is the following relation holds:

$$\langle\, (\mathbf{g}\mathbf{a}) \circ \mathbf{L}, \mathbf{b}\,\rangle = \langle\, \mathbf{g}\mathbf{a}, \mathbf{L}\mathbf{b}\,\rangle = \mathbf{g}(\mathbf{a}, \mathbf{L}\mathbf{b}) = \mathbf{g}(\mathbf{L}^T\mathbf{a}, \mathbf{b}) = \langle\, \mathbf{g}(\mathbf{L}^T\mathbf{a}), \mathbf{b}\,\rangle\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in V\,,$$

so that

$$(\mathbf{g}\mathbf{a}) \circ \mathbf{L} = \mathbf{g}(\mathbf{L}^T\mathbf{a})\,, \quad \forall\, \mathbf{a} \in V\,.$$

The metric tensor $\mathbf{g}^* \in BL(V^{*2}\,;\Re)$ may be identified with the linear isomorphism $\mathbf{g}^* \in BL(V^*\,;V)$ which is in fact $\mathbf{g}^{-1} \in BL(V^*\,;V)$. Indeed, by definition:

$$\mathbf{g}^*(\mathbf{g}\,\mathbf{a}, \mathbf{g}\,\mathbf{b}) = \langle\, \mathbf{g}^*(\mathbf{g}\,\mathbf{a}), \mathbf{g}\,\mathbf{b}\,\rangle = \mathbf{g}(\mathbf{a}, \mathbf{b}) = \langle\, \mathbf{a}, \mathbf{g}\,\mathbf{b}\,\rangle\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in V\,,$$

$$\iff \mathbf{g}^*(\mathbf{g}\,\mathbf{a}) = \mathbf{a}\,, \quad \forall\, \mathbf{a} \in V \iff \mathbf{g}^* = \mathbf{g}^{-1}\,.$$

To tensors $\boldsymbol{\alpha}^* \in BL(V^{*2}\,;\Re) = BL(V^*\,;V)$ and $\boldsymbol{\beta} \in BL(V^2\,;\Re) = BL(V\,;V^*)$ we may associate the linear operators $\mathbf{A} \in BL(V\,;V)$ and $\mathbf{B} \in BL(V\,;V)$ according to the correspondences:

$$\boldsymbol{\alpha}^* \circ \mathbf{g} = \mathbf{A} \iff \boldsymbol{\alpha}^* = \mathbf{A} \circ \mathbf{g}^{-1}$$

$$\mathbf{g}^{-1} \circ \boldsymbol{\beta} = \mathbf{B} \iff \boldsymbol{\beta} = \mathbf{g} \circ \mathbf{B}\,.$$

The metric tensor $\mathbf{g}$ induces a metric $\langle \cdot, \cdot \rangle_{\mathbf{g}}$ in the linear space $BL(V\,;V)$ of (bounded) linear operators by setting

$$\langle \mathbf{A}, \mathbf{B} \rangle_{\mathbf{g}} := \mathrm{tr}(\mathbf{A}^T\mathbf{B})\,,$$

and the following properties hold:

$$\langle \mathbf{A}, \mathbf{B} \rangle_{\mathbf{g}} = \mathrm{tr}(\mathbf{B}\mathbf{A}^T) = \langle \mathbf{B}^T, \mathbf{A}^T \rangle_{\mathbf{g}} = \mathrm{tr}(\mathbf{B}^T\mathbf{A}) = \langle \mathbf{B}, \mathbf{A} \rangle_{\mathbf{g}}\,,$$

$$\langle \mathbf{A}, \mathbf{A} \rangle_{\mathbf{g}} = \mathrm{tr}(\mathbf{A}^T\mathbf{A}) > 0 \quad \text{if} \quad \mathbf{A} \neq 0\,,$$

$$\langle \mathbf{ABC}, \mathbf{D} \rangle_{\mathbf{g}} = \langle \mathbf{B}, \mathbf{A}^T\mathbf{D}\mathbf{C}^T \rangle_{\mathbf{g}}\,.$$

A basic duality exists between the linear spaces of covariant and contravariant second order tensors on a vector space V. The *duality pairing* between two tensors $\boldsymbol{\alpha}^* \in BL(V^{*2}\,;\Re)$ and $\boldsymbol{\beta} \in BL(V^2\,;\Re)$ is then defined by:

$$\langle \boldsymbol{\alpha}^*, \boldsymbol{\beta} \rangle := \mathrm{tr}(\boldsymbol{\alpha}^* \circ \boldsymbol{\beta}) = \langle \boldsymbol{\alpha}^*\mathbf{g}, \mathbf{g}^{-1}\boldsymbol{\beta} \rangle_{\mathbf{g}}\,.$$

where $\boldsymbol{\alpha}^* \circ \mathbf{g} \in BL(V\,;V)$, $\mathbf{g}^{-1} \circ \boldsymbol{\beta} \in BL(V\,;V)$ and $\boldsymbol{\alpha}^* \circ \boldsymbol{\beta} \in BL(V\,;V)$.

The definition of the duality pairing is then well-posed since it is independent of the choice of the metric tensor. as is apparent from the formula above.

Let us now consider two HILBERT spaces $\{V\,, \mathbf{g}_V\}$ and $\{W\,, \mathbf{g}_W\}$ and their duals V^* and W^*. The bounded linear operator $\mathbf{A} \in BL(V\,;W)$ and its dual $\mathbf{A}^* \in BL(W^*\,;V^*)$, are related by

$$\langle \mathbf{w}^*, \mathbf{A}\mathbf{v} \rangle = \langle \mathbf{A}^*\mathbf{w}^*, \mathbf{v} \rangle\,, \quad \begin{cases} \forall\, \mathbf{v} \in V\,, \\ \forall\, \mathbf{w}^* \in W^*\,, \end{cases}$$

and the operator $\mathbf{A} \in BL(V\,;W)$ and its transpose $\mathbf{A}^T \in BL(W\,;V)$, are related by

$$\mathbf{g}_W\,(\mathbf{w}, \mathbf{A}\mathbf{v}) = \mathbf{g}_V\,(\mathbf{A}^T\mathbf{w}, \mathbf{v})\,, \quad \begin{cases} \forall\, \mathbf{v} \in V\,, \\ \forall\, \mathbf{w} \in W\,. \end{cases}$$

Then, being $\mathbf{g}_V \in BL(V\,;V^*)$ and $\mathbf{g}_W \in BL(W\,;W^*)$, it is:

$$\langle \mathbf{A}^*\mathbf{g}_W\mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{g}_W\mathbf{w}, \mathbf{A}\mathbf{v} \rangle = \mathbf{g}_W\,(\mathbf{w}, \mathbf{A}\mathbf{v}) = \mathbf{g}_V\,(\mathbf{A}^T\mathbf{w}, \mathbf{v})\,.$$

Then we have the commutative diagrams

$$\begin{array}{ccc} W & \xrightarrow{\ \mathbf{A}^T\ } & V \\ {\scriptstyle \mathbf{g}_W}\big\downarrow & & \big\downarrow{\scriptstyle \mathbf{g}_V} \\ W^* & \xrightarrow{\ \mathbf{A}^*\ } & V^* \end{array} \iff \quad \mathbf{A}^* \circ \mathbf{g}_W = \mathbf{g}_V \circ \mathbf{A}^T\,,$$

and

$$\begin{array}{ccc} W & \xrightarrow{\mathbf{A}^T} & V \\ \mathbf{g}_W^{-1}\Big\uparrow & & \mathbf{g}_V^{-1}\Big\uparrow \\ W^* & \xrightarrow{\mathbf{A}^*} & V^* \end{array} \qquad \Longleftrightarrow \qquad \mathbf{g}_V^{-1} \circ \mathbf{A}^* = \mathbf{A}^T \circ \mathbf{g}_W^{-1} \,,$$

so that $\mathbf{A}^* = \mathbf{g}_V \circ \mathbf{A}^T \circ \mathbf{g}_W^{-1}$.

1.1.5 Derivative and gradient of tensor functions

Let $f \in \mathrm{C}^1(BL\,(V^{*2}\,;\Re)\,;\Re)$ be a potential on the linear space of twice contravariant tensors on V and let $f_{\mathbf{g}} \in \mathrm{C}^1(BL\,(V\,;V)\,;\Re)$ be the associated potential on the linear space of linear operators on V , according to the relation:

$$f_{\mathbf{g}}(\boldsymbol{\alpha}^*\mathbf{g}) := f(\boldsymbol{\alpha}^*)\,, \quad \forall\, \boldsymbol{\alpha}^* \in BL\,(V^{*2}\,;\Re)\,,$$

where $\boldsymbol{\alpha}^*\mathbf{g} \in BL\,(V\,;V)$. We then have that

$$\begin{aligned} \langle Tf(\boldsymbol{\alpha}^*), \boldsymbol{\tau}^* \rangle &= \langle Tf_{\mathbf{g}}(\boldsymbol{\alpha}^*\mathbf{g}), \boldsymbol{\tau}^*\mathbf{g} \rangle = \langle \operatorname{grad} f_{\mathbf{g}}(\boldsymbol{\alpha}^*\mathbf{g}), \boldsymbol{\tau}^*\mathbf{g} \rangle_{\mathbf{g}} \\ &= \langle \mathbf{g} \operatorname{grad} f_{\mathbf{g}}(\boldsymbol{\alpha}^*\mathbf{g}), \boldsymbol{\tau}^* \rangle\,, \quad \forall\, \boldsymbol{\tau}^* \in BL\,(V^{*2}\,;\Re)\,. \end{aligned}$$

The derivative $Tf(\boldsymbol{\alpha}^*) \in BL\,(V^{*2}\,;\Re)^* = BL\,(V^2\,;\Re)$ is a twice covariant tensor on V and the gradient $\operatorname{grad} f_{\mathbf{g}}(\boldsymbol{\alpha}^*\mathbf{g}) \in BL\,(V\,;V)$ is a linear operator on V .

They are related by:

$$Tf(\boldsymbol{\alpha}^*) = \mathbf{g} \circ \operatorname{grad} f_{\mathbf{g}}(\boldsymbol{\alpha}^*\mathbf{g})\,.$$

An analogous result holds for a potential $h \in \mathrm{C}^1(BL\,(V^2\,;\Re)\,;\Re)$ on the linear space of twice covariant tensors and the associated potential on the linear space of linear operators on V , according to the relation:

$$h_{\mathbf{g}}(\mathbf{g}^{-1}\boldsymbol{\alpha}) := h(\boldsymbol{\alpha})\,, \quad \forall\, \boldsymbol{\alpha} \in BL\,(V^2\,;\Re)\,,$$

being

$$T_{\boldsymbol{\alpha}} h = \operatorname{grad} h_{\mathbf{g}}(\mathbf{g}^{-1}\boldsymbol{\alpha}) \circ \mathbf{g}^{-1}\,,$$

with $T_{\boldsymbol{\alpha}} h \in BL\,(V^2\,;\Re)^* = BL\,(V^{*2}\,;\Re)$ and $\operatorname{grad} h_{\mathbf{g}}(\mathbf{g}^{-1}\boldsymbol{\alpha}) \in BL\,(V\,;V)$.

1.1.6 Categories, Morphisms and Functors

The concepts of *category* was introduced in modern geometry to provide a unifying framework for many basic concepts [91]. The category theory was founded by EILENBERG and MACLANE about 1945.

Definition 1.1.1 *A category* CAT *is a family of objects* $\{\mathbf{A}, \mathbf{B}, \ldots\}$ *such that to any ordered pair* $\{\mathbf{A}, \mathbf{B}\}$ *of objects there corresponds a set* $\mathrm{MOR}(\mathbf{A}, \mathbf{B})$ *of morphisms and for any ordered triplet* $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ *there corresponds an associative law of composition:*

$$\mathrm{MOR}(\mathbf{A}, \mathbf{B}) \times \mathrm{MOR}(\mathbf{B}, \mathbf{C}) \mapsto \mathrm{MOR}(\mathbf{A}, \mathbf{C}),$$

expressed as

$$\mathbf{f} \in \mathrm{MOR}(\mathbf{A}, \mathbf{B}), \quad \mathbf{g} \in \mathrm{MOR}(\mathbf{B}, \mathbf{C}) \implies \mathbf{f} \circ \mathbf{g} \in \mathrm{MOR}(\mathbf{A}, \mathbf{C}),$$

fulfilling the properties:

$$\mathbf{A} = \overline{\mathbf{A}} \text{ and } \mathbf{B} = \overline{\mathbf{B}} \implies \mathrm{MOR}(\mathbf{A}, \mathbf{B}) = \mathrm{MOR}(\overline{\mathbf{A}}, \overline{\mathbf{B}}),$$
$$\mathbf{A} \neq \overline{\mathbf{A}} \text{ or } \mathbf{B} \neq \overline{\mathbf{B}} \implies \mathrm{MOR}(\mathbf{A}, \mathbf{B}) \cap \mathrm{MOR}(\overline{\mathbf{A}}, \overline{\mathbf{B}}) = \emptyset,$$

where $\mathbf{A}, \overline{\mathbf{A}}, \mathbf{B}, \overline{\mathbf{B}} \in \mathrm{CAT}$.

Moreover for each $\mathbf{A} \in \mathrm{CAT}$ *there is an identity morphism*

$$\mathbf{id}_{\mathbf{A}} \in \mathrm{MOR}(\mathbf{A}, \mathbf{A}).$$

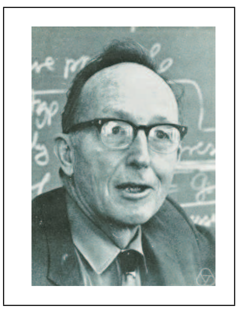

Figure 1.4: Saunders MacLane (1909 - 2005)

Figure 1.5: Samuel Eilenberg (1913 - 1998)

Definition 1.1.2 (Covariant and contravariant functors) *A covariant functor* $F : \mathrm{CAT} \mapsto \overline{\mathrm{CAT}}$ *is a map which associates with each object* $\mathbf{A} \in \mathrm{CAT}$ *an object* $F(\mathbf{A}) \in \overline{\mathrm{CAT}}$ *and with each morphism* $\mathbf{f} \in \mathrm{MOR}(\mathbf{A}, \mathbf{B})$ *, with* $\mathbf{A}, \mathbf{B} \in \mathrm{CAT}$ *, a morphism* $F(\mathbf{f}) \in \mathrm{MOR}(F(\mathbf{A}), F(\mathbf{B}))$ *which preserves the identity and the composition law:*

$$F(\mathbf{id}_{\mathbf{A}}) = \mathbf{id}_{F(\mathbf{A})} \,, \qquad F(\mathbf{g} \circ \mathbf{f}) = F(\mathbf{g}) \circ F(\mathbf{f}) \,,$$

so that the following is a commutative diagram

$$\begin{array}{ccc} \mathbf{A} & \xrightarrow{F} & F(\mathbf{A}) \\ {\scriptstyle \mathbf{f}}\big\downarrow & & \big\downarrow{\scriptstyle F(\mathbf{f})} \\ \mathbf{B} & \xrightarrow{F} & F(\mathbf{B}) \\ {\scriptstyle \mathbf{g}}\big\downarrow & & \big\downarrow{\scriptstyle F(\mathbf{g})} \\ \mathbf{C} & \xrightarrow{F} & F(\mathbf{C}) \end{array}$$

In contravariant functors the arrows are reversed so that the morphism $\mathbf{f} \in \mathrm{MOR}(\mathbf{A}, \mathbf{B})$ *transforms into* $F(\mathbf{f}) \in \mathrm{MOR}(F(\mathbf{B}), F(\mathbf{A}))$ *and the transformation of the composition law becomes* $F(\mathbf{g} \circ \mathbf{f}) = F(\mathbf{f}) \circ F(\mathbf{g})$ *, so that the following is* *a commutative diagram*

$$\begin{array}{ccc} \mathbf{A} & \xrightarrow{F} & F(\mathbf{A}) \\ \mathbf{f}\Big\downarrow & & \Big\uparrow F(\mathbf{f}) \\ \mathbf{B} & \xrightarrow{F} & F(\mathbf{B}) \\ \mathbf{g}\Big\downarrow & & \Big\uparrow F(\mathbf{g}) \\ \mathbf{C} & \xrightarrow{F} & F(\mathbf{C}) \end{array}$$

The functors $F : \textsc{Cat} \mapsto \overline{\textsc{Cat}}$ of the same variance from a category $\textsc{Cat}$ to a category $\overline{\textsc{Cat}}$ are themselves the objects of a category $\textsc{Fun}\{\textsc{Cat}\,, \overline{\textsc{Cat}}\}$.

Definition 1.1.3 (Functor morphism) *A functor morphism or natural transformation is a morphism of the category* $\textsc{Fun}\{\textsc{Cat}\,, \overline{\textsc{Cat}}\}$ *and is defined as follows. For any pair of covariant functors* $F1, F2 \in \textsc{Fun}\{\textsc{Cat}\,, \overline{\textsc{Cat}}\}$ *a natural transformations* $\textsc{Nat} : F1 \mapsto F2$ *is a collection of morphisms*

$$\textsc{Nat}_{\mathbf{A}} : F1(\mathbf{A}) \mapsto F2(\mathbf{A})\,,$$

with $\mathbf{A}$ *ranging in* $\textsc{Cat}$, *such that for any* $\mathbf{f} \in \textsc{Mor}(\mathbf{A}, \mathbf{B})$, *with* $\mathbf{A}, \mathbf{B} \in \textsc{Cat}$, *we have the commutative diagram:*

$$\begin{array}{ccccccc} \mathbf{A} & \xrightarrow{F1} & F1(\mathbf{A}) & \xrightarrow{\textsc{Nat}_{\mathbf{A}}} & F2(\mathbf{A}) & \xleftarrow{F2} & \mathbf{A} \\ \mathbf{f}\Big\downarrow & & F1(\mathbf{f})\Big\downarrow & & \Big\downarrow F2(\mathbf{f}) & & \mathbf{f}\Big\downarrow \\ \mathbf{B} & \xrightarrow{F1} & F1(\mathbf{B}) & \xrightarrow{\textsc{Nat}_{\mathbf{B}}} & F2(\mathbf{B}) & \xleftarrow{F2} & \mathbf{B} \end{array}$$

Definition 1.1.4 (Isomorphism)

An isomorphism in a category $\textsc{Cat}$ *is a morphism* $\mathbf{f} \in \textsc{Mor}(\mathbf{A}, \mathbf{B})$ *with the property that there exists a morphism* $\mathbf{g} \in \textsc{Mor}(\mathbf{B}, \mathbf{A})$ *such that* $\mathbf{f} \circ \mathbf{g} \in \textsc{Mor}(\mathbf{B}, \mathbf{B})$ *and* $\mathbf{g} \circ \mathbf{f} \in \textsc{Mor}(\mathbf{A}, \mathbf{A})$ *are identities.*

Definition 1.1.5 (Section of a morphism)

A section of a morphism $\mathbf{f} \in \textsc{Mor}(\mathbf{A}, \mathbf{B})$ *in a category* $\textsc{Cat}$ *is a morphism* $\mathbf{s} \in \textsc{Mor}(\mathbf{B}, \mathbf{A})$ *such that* $\mathbf{f} \circ \mathbf{s} \in \textsc{Mor}(\mathbf{B}, \mathbf{B})$ *is the identity.*

1.1.7 Flows

Evolution problems defined on a manifold are of the utmost importance in physics. They emerge, for example, in dynamics when studying the motion of a body subject to nonlinear kinematical contraints.

Motions are described by a two-parameter family of diffeomorphisms of the ambient manifold into itself. The two scalar parameters are the initial (or start) time and the final (or current) time and the diffeomorphisms of the family are called flows. When the start and the current times coincide the flow is the identity map. As we shall see, composed flows fulfill a determinism law.

The tangents to the trajectories of the flows provide a vector field of velocities on the manifold, and vice versa any assigned regular vector field is the velocity field of a flow that can be evaluated by a time integration. If the vector field is dependent on scalar parameters, the flow also will depend on these parameters.

The next section provides some basic results in the theory of ordinary differential equations, concerning existence, uniqueness and continuous dependence of the solution on the inital data.

These results are essential to get a proper definition of a flow associated with a velocity vector field and to illustrate its main properties.

1.1.8 Ordinary differential equations

Let $\mathbb{M}$ be a C^k differentiable manifold modeled on the BANACH space E.

- A vector field $v \in \mathrm{C}^0(E\,;E)$ is said to be LIPSCHITZ continuous if there is a constant $\mathrm{LIP} > 0$ such that

$$\|v(x) - v(y)\| \leq \mathrm{LIP}\,\|x - y\|\,, \quad \forall\, x, y \in E\,.$$

- To a vector field $\mathbf{v} \in \mathrm{C}^k(\mathbb{M}\,;\mathbb{TM})$ we may associate, in a local way, a vector field $v \in \mathrm{C}^k(E\,;E)$ in the model space by taking a chart $\{U, \varphi\}$ and pushing it forward according to $\varphi \in \mathrm{C}^1(\mathbb{M}\,;E)$:

$$v := \varphi\!\uparrow\!\mathbf{v}\,.$$

To any LIPSCHITZ continuous vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ there corresponds (at least locally) a unique integral curve $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{M})$ thru a point $\mathbf{x} \in \mathbb{M}$, solution of the differential equation

$$\partial_{\mu=\lambda}\,\mathbf{c}(\mu) = \mathbf{v}(\mathbf{c}(\lambda))\,, \quad \lambda \in I = [-\varepsilon, +\varepsilon]\,, \quad \varepsilon > 0\,,$$

under the initial condition $\mathbf{c}(0) = \mathbf{x}$.

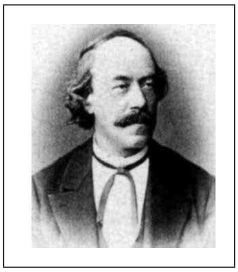

Figure 1.6: Rudolf Otto Sigismund Lipschitz (1832 - 1903)

The solution depends with continuity on the initial condition. If the vector field $\mathbf{v}$ is time-dependent, the LIPSCHITZ continuity is to be fulfilled uniformly in time, that is:

$$\sup_{\lambda \in I} \|v(x,\lambda) - v(y,\lambda)\| \leq \text{LIP} \, \|x - y\| \,, \quad \forall \, x, y \in E \,,$$

with $\text{LIP} > 0$ independent of time.

The differential equation then writes

$$\partial_{\mu=\lambda} \, \mathbf{c}(\mu) = \mathbf{v}(\mathbf{c}(\lambda), \lambda) \,, \quad \lambda \in I = [-\varepsilon, +\varepsilon] \,, \quad \varepsilon > 0 \,.$$

To prove this assertion we must rely upon two fundamental results in ordinary differential equation theory.

The former, presented in Proposition 1.1.1 and referred to in the literature as BANACH*'s fixed point theorem*, the *contraction lemma* or the *shrinking lemma*, provides existence and uniqueness of the solution.

The latter, known as the GRÖNWALL*'s lemma*, is presented in Proposition 1.1.3 and ensures continuous dependence of the solution on the initial condition.

Proposition 1.1.1 (Banach's fixed point theorem) *Let* $\{\mathcal{X}, \text{DIST}\}$ *be a complete metric space and* $\mathbf{T} : \mathcal{X} \mapsto \mathcal{X}$ *a contraction mapping:*

$$\text{DIST}(\mathbf{T}(\mathbf{x}), \mathbf{T}(\mathbf{y})) \leq \alpha \, \text{DIST}(\mathbf{x}, \mathbf{y}) \,, \quad \alpha < 1 \,, \quad \forall \, \mathbf{x}, \mathbf{y} \in \mathcal{X} \,.$$

Then the fixed point problem

$$\mathbf{T}(\mathbf{x}) = \mathbf{x} \,, \quad \mathbf{x} \in \mathcal{X} \,,$$

admits a unique solution.

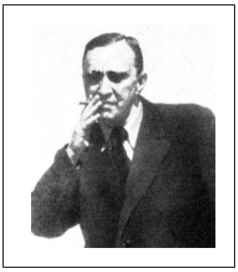

Figure 1.7: Stefan Banach (1892 - 1945)

Proof. The contraction property implies that the map $\mathbf{T}$ is continuous. Setting $\mathbf{x}_o \in \mathcal{X}$ we may define by induction the sequence

$$\mathbf{x}_{n+1} = \mathbf{T}(\mathbf{x}_n) \iff \mathbf{x}_{n+1} = \mathbf{T}^n(\mathbf{x}_o)\,,$$

where $\mathbf{T}^n$ is the n-th iterate of $\mathbf{T}$.

By induction we get

$$\mathrm{DIST}(\mathbf{x}_n, \mathbf{x}_{n+1}) \le \alpha^n\, \mathrm{DIST}(\mathbf{x}_o, \mathbf{x}_1)\,.$$

By the triangle inequality it follows that

$$\mathrm{DIST}(\mathbf{x}_n, \mathbf{x}_{n+k}) \le \mathrm{DIST}(\mathbf{x}_n, \mathbf{x}_{n+1}) + \ldots + \mathrm{DIST}(\mathbf{x}_{n+k-1}, \mathbf{x}_{n+k})$$

$$\le (\alpha^n + \ldots + \alpha^{n+k-1})\, \mathrm{DIST}(\mathbf{x}_o, \mathbf{x}_1)\,.$$

Being $\alpha < 1$ the series $\sum_{n=0}^{\infty} \alpha^n$ is convergent and therefore the partial sum $\alpha^n + \ldots + \alpha^{n+k-1}$ tends to zero as $n \to \infty$. The sequence $\{\mathbf{x}_n\}$ is then a CAUCHY sequence and the completeness of the space $\mathcal{X}$ ensures the existence of $\mathbf{x} \in \mathcal{X}$ such that

$$\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}\,.$$

Hence, by the continuity of $\mathbf{T}$, we get

$$\mathbf{T}(\mathbf{x}) = \lim_{n\to\infty} \mathbf{T}(\mathbf{x}_n) = \lim_{n\to\infty} \mathbf{x}_{n+1} = \mathbf{x}\,,$$

and the existence is proven. Uniqueness follows by observing that $\mathbf{T}(\mathbf{y}) = \mathbf{y}$ with $\mathbf{y} \in \mathcal{X}$ implies that

$$\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{T}(\mathbf{x}) - \mathbf{T}(\mathbf{y})\| \le \alpha\, \|\mathbf{x} - \mathbf{y}\| \implies \mathbf{x} = \mathbf{y}\,,$$

since $\alpha < 1$. ∎

The next result is based on a method of successive approximations due to Emile Picard and extended by Lindelöf and by Banach.

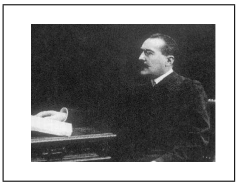

Figure 1.8: Charles Emile Picard (1856 - 1941)

Figure 1.9: Ernst Leonard Lindelöf (1870 - 1946)

Proposition 1.1.2 (Existence and uniqueness) *The differential equation*

$$\partial_{\mu=\lambda}\, \mathbf{c}(\mu) = \mathbf{v}(\mathbf{c}(\lambda), \lambda)\,, \quad \lambda \in I\,,$$

with the initial condition $\mathbf{c}(0) = \mathbf{x}$, *admits a unique solution in a neighborhood of* $\mathbf{x} \in \mathbb{M}$.

Proof. By a local chart $\{U, \varphi\}$ around $\mathbf{x} \in \mathbb{M}$ we may set

$$v := \varphi\uparrow\mathbf{v}\,, \qquad c := \varphi \circ \mathbf{c}\,,$$

and write the differential equation in the model Banach's space as

$$\partial_{\mu=\lambda}\, c(\mu) = v(c(\lambda), \lambda)\,, \quad \mu, \lambda \in I\,,$$

with the initial condition $c(0) = x$. An equivalent formulation is provided by the integral equation

$$c(\lambda) = x + \int_0^\lambda v(c(s), s)\, ds\,.$$

Let $X = \mathrm{C}^0(I\,;E) \cap \mathcal{B}(I\,;E)$ be the BANACH space of bounded continuous maps with the uniform convergence topology induced by the norm

$$\|c\|_X = \sup_{\lambda \in I} \|c(\lambda)\|_E \,,$$

and $T : X \mapsto X$ the map defined pointwise by

$$(T \circ c)(\lambda) = T(c(\lambda)) := x + \int_0^\lambda v(c(s), s)\, ds \,.$$

Then the solution of the differential equation is a fixed point of T. Now, by the uniform LIPSCHITZ continuity of the vector field, we have that

$$\begin{aligned}
\|T(c_2) - T(c_1)\|_X &= \sup_{\lambda \in I} \| \int_0^\lambda v(c_2(s), s) - v(c_1(s), s)\, ds \|_E \\
&\leq \sup_{\lambda \in I} \int_0^\lambda \|v(c_2(s), s) - v(c_1(s), s)\|_E\, ds \\
&\leq \sup_{\lambda \in I} \int_0^\lambda \mathrm{LIP}\, \|c_2(s) - c_1(s)\|_E\, ds \\
&\leq \mathrm{LIP} \int_I \sup_{s \in I} \|c_2(s) - c_1(s)\|_E\, ds \leq \mathrm{LIP} \cdot \mathrm{meas}\,(I)\, \|c_2 - c_1\|_X \,.
\end{aligned}$$

Taking I such that $\mathrm{LIP} \cdot \mathrm{meas}\,(I) < 1$, the map $T : X \mapsto X$ has the contraction property in the BANACH space X and we may apply Proposition 1.1.1 to get existence and uniqueness. ■

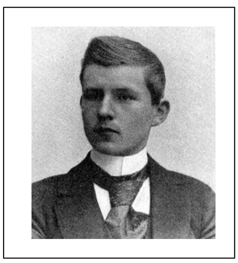

Figure 1.10: Thomas Hakon Grönwall (1877 - 1932)

Proposition 1.1.3 (Grönwall's lemma) *Let* $f, g \in \mathrm{C}^0(I\,; \Re)$ *be continuous and nonnegative on* $I = [a, b]$ *. If for some constant* $k > 0$ *it is:*

$$f(\lambda) \leq k + \int_a^\lambda f(s)\, g(s)\, ds \qquad \forall\, \lambda \in I\,,$$

then the following inequality holds

$$f(\lambda) \leq k\, \mathrm{EXP}\, (\int_a^\lambda g(s)\, ds) \qquad \forall\, \lambda \in I\,.$$

Proof. Setting

$$\alpha(\lambda) := k + \int_a^\lambda f(s)\, g(s)\, ds\,,$$

we have that $\alpha(\lambda) > 0$ and $\alpha'(\lambda) = g(\lambda)\, f(\lambda)$ for all $\lambda \in I$.

By assumption $f(\lambda) \leq \alpha(\lambda)$ so that $\alpha'(\lambda) \leq g(\lambda)\, \alpha(\lambda)$. Since $\alpha(a) = k$, integrating we get

$$\alpha(\lambda) \leq k\, \mathrm{EXP}\, (\int_a^\lambda g(s)\, ds) \qquad \forall\, \lambda \in I\,,$$

and hence the result. ■

Proposition 1.1.4 (Continuous dependence on the initial conditions) *Let us denote by* $F_\lambda(x_0)$ *the flow of the vector field* $v \in \mathrm{C}^0(E\,; E)$ *passing thru* $x_0 \in E$ *, that is the solution of the differential equation*

$$\partial_{\mu=\lambda}\, c(\mu) = v(c(\lambda), \lambda)\,, \quad \lambda \in I\,,$$

with the initial condition $c(0) = x_0$ *. Then there exists a neighborhood* $U(x_0)$ *and a time interval* $I = [-\varepsilon, +\varepsilon]$ *such that*

$$\| F_\lambda(y) - F_\lambda(x) \|_E \leq \mathrm{EXP}\, (\lambda\, t)\, \| y - x \|_E\,, \quad \forall\, \lambda \in I\,.$$

Proof. The flow fulfils the integral equation

$$F_\lambda(x) = x + \int_0^\lambda v(F_s(x), s)\, ds\,.$$

Hence, setting $f(\lambda) := \| F_\lambda(y) - F_\lambda(x) \|_E$, by the uniform LIPSCHITZ continuity of the vector field, we have that

$$f(\lambda) \leq \| y - x \|_E + \mathrm{LIP} \int_0^\lambda f(s)\, ds\,,$$

and the result follows by GRÖNWALL's lemma. ■

1.2 Differentiable manifolds

We provide here some basic facts and definitions about differentiable manifolds.

- Let $\mathbb{M}$ be a set and E be a BANACH space. A *chart* $\{U, \varphi\}$ on $\mathbb{M}$ is a pair with $\varphi : U \mapsto E$ bijection between the subset $U \subset \mathbb{M}$ and an open set in E. A C^k-*atlas* $\mathcal{A}$ on $\mathbb{M}$ is a family of charts $\{\{U_i, \varphi_i\} \,|\, i \in I\}$ such that $\{\cup U_i \,|\, i \in I\}$ is a covering of $\mathbb{M}$ and that the overlap maps are C^k-diffeomorfisms.

- Two atlases are equivalent if their union is still a C^k-atlas. The union of all the atlases equivalent to a given one $\mathcal{A}$ is called the *differentiable structure* generated by $\mathcal{A}$.

- A C^k *differentiable manifold* modeled on the BANACH space E is a pair $\{\mathbb{M}, \mathcal{D}\}$ where $\mathcal{D}$ is an equivalence class of C^k-atlases on $\mathbb{M}$. The space E is called the model space.

- A subset $\mathcal{O}$ of a differentiable manifold $\mathbb{M}$ is said to be *open* if for each $\mathbf{x} \in \mathcal{O}$ there is a chart $\{U, \varphi\}$ such that $\mathbf{x} \in U$ and $U \subset \mathcal{O}$.

- A *submanifold* $\mathbb{P} \subset \mathbb{M}$ is a manifold such that for each $\mathbf{x} \in \mathbb{P}$ there is a chart $\{U, \varphi\}$ in $\mathbb{M}$, with $\mathbf{x} \in U$, fulfilling the *submanifold property*:

  $$\varphi : U \mapsto E = E_1 \times E_2 \,, \quad \varphi(U \cap \mathbb{P}) = \varphi(U) \cap (E_1 \times \{0\}) \,,$$

  that is, the restriction of the chart to a submanifold maps locally the submanifold into a component space of the model space. Every open subset of a manifold $\mathbb{M}$ is a submanifold.

- A *finite dimensional* differentiable manifold is a manifold modeled on a finite dimensional normed linear space. All the tangent spaces to a finite dimensional differentiable manifold are of the same dimension.

- Tangent vectors at $\mathbf{x} \in \mathbb{M}$ are most simply defined by considering a regular curve $\mathbf{c}(\lambda) \in \mathbb{M}$, with $\lambda \in I \subset \Re$ an open interval of the real line containing the zero, such that $\mathbf{x} = \mathbf{c}(0)$. We then define a tangent vector $\{\mathbf{x}\,, \mathbf{v}\} = \mathbf{c}'(0)$ as an equivalence class of the curves thru $\mathbf{x} = \mathbf{c}(0)$ which share, in the model space E, a common tangent vector $(\varphi \circ \mathbf{c})'(0)$ to the corresponding curves $(\varphi \circ \mathbf{c})(\lambda)$, generated by a local chart $\{U, \varphi\}$ such that $\mathbf{x} \in U$.

The set of tangent vectors at $\mathbf{x} \in \mathbb{M}$ is a linear space, said the *tangent space* and denoted by $\mathbb{T}_{\mathbb{M}}(\mathbf{x}) = \mathbb{T}_{\mathbf{x}}\mathbb{M} = \mathbb{T}\mathbb{M}(\mathbf{x})$.

Tangent vectors $\{\mathbf{x}, \mathbf{v}\} \in \mathbb{T}_{\mathbb{M}}(\mathbf{x})$ may also be uniquely defined by requiring that they fulfil the formal properties of a *point derivation*:

$$\begin{cases} (\mathbf{v}_1 + \mathbf{v}_2)\, f = \mathbf{v}_1\, f + \mathbf{v}_2\, f\,, & \textit{additivity} \\ \mathbf{v}(a\, f) = a\,(\mathbf{v}\, f)\,, \quad a \in \Re\,, & \textit{homogeneity} \\ \mathbf{v}(fg) = (\mathbf{v}f)\, g + f\,(\mathbf{v}g)\,, & \end{cases} \quad \begin{matrix} \left.\begin{matrix} \\ \\ \end{matrix}\right\} \Re\text{-linearity} \\ \textsc{Leibniz rule} \end{matrix}$$

where $f, g \in \mathrm{C}^r(\mathbf{x}, U)$ and $\mathbf{v}f$ denotes the scalar field result of the operation $\mathbf{v}$ on the scalar field f. This point of view, that identifies the tangent vectors at a point of a differentiable manifold with the directional derivatives of smooth scalar functions at that point, is the most convenient to get basic results of differential geometry. Accordingly we may define the *tangent space* $\mathbb{T}\mathbb{M}(\mathbf{x})$ at a point $\mathbf{x} \in \mathbb{M}$ as the linear space of *tangent vectors* $\{\mathbf{x}, \mathbf{v}\} : \mathrm{C}^r(\mathbf{x}, U) \mapsto \mathrm{C}^{r-1}(\mathbf{x}, U)$ where $\mathrm{C}^r(\mathbf{x}, U)$ is the germ of scalar functions which are r-times continuously differentiable in a neighborhood U of $\mathbf{x} \in \mathbb{M}$.

1.2.1 Tangent and cotangent bundles

The *tangent bundle* $\mathbb{T}\mathbb{M}$ to the manifold $\mathbb{M}$ is the disjoint union of the pairs $\{\mathbf{x}, \mathbb{T}_{\mathbf{x}}\mathbb{M}\}$ with $\mathbf{x} \in \mathbb{M}$. An element $\{\mathbf{x}, \mathbf{v}\} \in \{\mathbf{x}, \mathbb{T}_{\mathbf{x}}\mathbb{M}\}$ is called a tangent vector applied at the base point $\mathbf{x} \in \mathbb{M}$.

The manifols $\mathbb{M}$ is called the *base manifold* of the tangent bundle $\mathbb{T}\mathbb{M}$ and each tangent space $\mathbb{T}_{\mathbf{x}}\mathbb{M}$ is called the *fibre* over $\mathbf{x} \in \mathbb{M}$. The characteristic operation on the tangent bundle is the *projection* on base points $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}; \mathbb{M})$ defined by $\boldsymbol{\tau}_{\mathbb{M}}(\{\mathbf{x}, \mathbf{v}\}) = \mathbf{x} \in \mathbb{M}$. The tangent bundle $\mathbb{T}\mathbb{M}$ to a manifold $\mathbb{M}$ is itself a manifold whose atlas of charts is induced by the differentiable structure of $\mathbb{M}$ by taking the differential of its charts.

- A C^k-*vector field* is a map $\mathbf{v} \in \mathrm{C}^k(\mathbb{M}; \mathbb{T}\mathbb{M})$ which to any point $\mathbf{x} \in \mathbb{M}$ assigns a vector $\{\mathbf{x}, \mathbf{v}(\mathbf{x})\} \in \{\mathbf{x}, \mathbb{T}_{\mathbf{x}}(\mathbb{M})\}$ based at $\mathbf{x} \in \mathbb{M}$. A vector field is therefore characterized by the property that its left composition $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} \in \mathrm{C}^1(\mathbb{M}; \mathbb{M})$, with the projection $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}; \mathbb{M})$, is the identity map on $\mathbb{M}$:
$$\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} = \mathbf{id}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{M}; \mathbb{M})\,.$$

According to the definition in section 1.1.6, a vector field is a section of the morphism $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}; \mathbb{M})$.

The *cotangent bundle* $\mathbb{T}^*\mathbb{M}$ to the manifold $\mathbb{M}$, also denoted as $\mathbb{T}\mathbb{M}^*$, is the disjoint union of the pairs $\{\mathbf{x}\,,\mathbb{T}^*_{\mathbf{x}}(\mathbb{M})\}$ with $\mathbb{T}^*_{\mathbf{x}}(\mathbb{M})$ topological dual space of $\mathbb{T}_{\mathbf{x}}(\mathbb{M})$. The elements of the cotangent bundle are called *covectors*.

A C^k-*covector field* is a map $\mathbf{v}^* \in \mathrm{C}^k(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ which to any point $\mathbf{x} \in \mathbb{M}$ assigns a covector $\{\mathbf{x}\,,\mathbf{v}^*(\mathbf{x})\} \in \{\mathbf{x}\,,\mathbb{T}^*_{\mathbf{x}}(\mathbb{M})\}$ based at $\mathbf{x} \in \mathbb{M}$.

A covector field is therefore characterized by the property that the left composition $\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathbf{v}^* \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ with the projection $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$ is the identity map.

We will denote by $\mathbb{T}\mathbb{M}(\mathcal{P}) \subseteq \mathbb{T}\mathbb{M}$ the disjoint union of pairs $\{\mathbf{x}\,,\mathbb{T}_{\mathbf{x}}\mathbb{M}\}$ with $\mathbf{x} \in \mathcal{P} \subseteq \mathbb{M}$.

Higher order tangent and cotangent bundles can be conceived by regarding the starting tangent and cotangent bundles as base manifolds.

1.2.2 Tensor fields

- A (q,p) *multilinear form* on a manifold $\mathbb{M}$ is a map $\mathcal{M} : \mathbb{M} \mapsto \Re$ which depends in a multilinear way on a set of p vector fields and q covector fields, taken according to any chosen ordering.

- A *tensor field* on $\mathbb{M}$ is a multilinear form with the property that its point-values depend only on the corresponding point-values of the vector and covector fields. A (q,p) tensor field is said to be p times covariant and q times contravariant.

The standard tensoriality criterion is provided by the following statement enunciated, for simplicity, with reference to a $(0,p)$ multilinear form (for the proof see [180] or [80] Lemma 7.3 or [91] Lemma 2.3 of Ch. VIII).

Lemma 1.2.1 (Tensoriality criterion) *A multilinear form* $\mathcal{M} : \mathbb{M} \mapsto \Re$, *which is linear on the space* $\mathrm{C}^\infty(\mathbb{M})$, *in the sense that*

$$\mathcal{M}(\mathbf{v}_1,\dots,f\,\mathbf{v}_i,\dots,\mathbf{v}_p) = f\,\mathcal{M}(\mathbf{v}_1,\dots,\mathbf{v}_p)\,,\quad \forall\, i=1,\dots,p\,,\quad \forall\, f \in \mathrm{C}^\infty(\mathbb{M})\,,$$

can be pointwise represented by a unique tensor field $T_{\mathcal{M}}$ *on* $\mathbb{M}$, *i.e.:*

$$\mathcal{M}(\mathbf{v}_1 \dots \mathbf{v}_p)(\mathbf{x}) := T_{\mathcal{M}}(\mathbf{x})(\mathbf{v}_1(\mathbf{x}),\dots,\mathbf{v}_p(\mathbf{x}))\,,\quad \forall\, \mathbf{x} \in \mathbb{M}\,.$$

1.2.3 Manifold morphisms

- A C^k-*morphism* between two C^k-differentiable manifolds $\mathbb{M}_1$ and $\mathbb{M}_2$ is a C^k-differentiable map $\boldsymbol{\varphi} \in C^k(\mathbb{M}_1 ; \mathbb{M}_2)$. Differentiability means that the composition of the map with local charts on $\mathbb{M}_1$ and $\mathbb{M}_2$ defines a differentiable map from the model space E_1 of $\mathbb{M}_1$ to the model BANACH space E_2 of $\mathbb{M}_2$.
- A C^k-*diffeomorphism* $\boldsymbol{\varphi} \in C^k(\mathbb{M}_1 ; \mathbb{M}_2)$ is a C^k-morphism which is invertible with a smooth inverse $\boldsymbol{\varphi}^{-1} \in C^k(\mathbb{M}_2 ; \mathbb{M}_1)$.
- The *differential* $T_{\mathbf{x}}\boldsymbol{\varphi} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M} ; \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N})$ of a morphism $\boldsymbol{\varphi} \in C^1(\mathbb{M} ; \mathbb{N})$ at the point $\mathbf{x} \in \mathbb{M}$ is the linear map defined by

$$T_{\mathbf{x}}\boldsymbol{\varphi} \cdot \mathbf{v}(\mathbf{x}) = \partial_{\lambda=0}\, (\boldsymbol{\varphi} \circ \mathbf{c})(\lambda) \in \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N} \,,$$

  for any curve $\mathbf{c} \in C^1(I ; \mathbb{M})$ with $0 \in I$, $\mathbf{c}(0) = \mathbf{x}$ and velocity at $\mathbf{x} \in \mathbb{M}$ given by $\mathbf{v}(\mathbf{x}) = \partial_{\lambda=0}\, \mathbf{c}(\lambda) \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ or, equivalently, defined for any $f \in C^1(\mathbb{N} ; \Re)$ by the derivation rule

$$(T_{\mathbf{x}}\boldsymbol{\varphi} \cdot \mathbf{v}(\mathbf{x}))f = \mathbf{v}(\mathbf{x})(f \circ \boldsymbol{\varphi}) \,.$$

- The *tangent map* $T\boldsymbol{\varphi} \in C^0(\mathbb{TM} ; \mathbb{TN})$ is defined by

$$(T\boldsymbol{\varphi} \circ \mathbf{v})(\mathbf{x}) = T_{\mathbf{x}}\boldsymbol{\varphi} \cdot \mathbf{v}(\mathbf{x}) \,, \quad \forall\, \mathbf{v} \in C^0(\mathbb{M} ; \mathbb{TM}) \,,$$

  which may also be written $T\boldsymbol{\varphi} \circ \mathbf{v} = T_{\boldsymbol{\pi}(\mathbf{v})}\boldsymbol{\varphi} \cdot \mathbf{v}$. The field $T\boldsymbol{\varphi} \circ \mathbf{v} \in C^0(\mathbb{M} ; \mathbb{TM})$ fulfils the relation

$$\boldsymbol{\tau}_{\mathbb{N}} \circ T\boldsymbol{\varphi} \circ \mathbf{v} = \boldsymbol{\varphi} \,.$$

  and is called a *vector field along* $\boldsymbol{\varphi}$.

A morphism $\boldsymbol{\varphi} \in C^1(\mathbb{M} ; \mathbb{N})$ and the associated tangent map $T\boldsymbol{\varphi} \in C^0(\mathbb{TM} ; \mathbb{TN})$ are related by the commutative diagram

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\boldsymbol{\varphi}} & \mathbb{TN} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{M}}}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\tau}_{\mathbb{N}}} \\ \mathbb{M} & \xrightarrow{\boldsymbol{\varphi}} & \mathbb{N} \end{array} \iff \quad \boldsymbol{\tau}_{\mathbb{N}} \circ T\boldsymbol{\varphi} = \boldsymbol{\varphi} \circ \boldsymbol{\tau}_{\mathbb{M}} \in C^0(\mathbb{TM} ; \mathbb{N}) \,.$$

This diagram states that taking a base point of a vector and then mapping it into another manifold by a morphism, provides the base point of the vector transformed by the tangent morphism.

Definition 1.2.1 (Immersion) *An immersion* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *is a morphism whose tangent map* $T\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{TM}\,;\mathbb{TN})$ *is fibrewise injective, i.e.*

$$\boldsymbol{ker}(T_{\mathbf{x}}\boldsymbol{\varphi}) = \{0\}\,, \quad \forall\, \mathbf{x} \in \mathbb{M}\,.$$

Definition 1.2.2 (Submersion) *A submersion* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *is morphism whose tangent map* $T\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{TM}\,;\mathbb{TN})$ *is fibrewise surjective, i.e.*

$$\boldsymbol{im}(T_{\mathbf{x}}\boldsymbol{\varphi}) = \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N}\,, \quad \forall\, \mathbf{x} \in \mathbb{M}\,.$$

Definition 1.2.3 (Immersed submanifold) *An immersed submanifold of* $\mathbb{N}$ *is the range* $\boldsymbol{\varphi}(\mathbb{M})$ *of an injective immersion* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *.*

Definition 1.2.4 (Embedding) *An embedding* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *is an injective immersion which is a homeomorphism between* $\mathbb{M}$ *and* $\boldsymbol{\varphi}(\mathbb{M})$ *, that is* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\boldsymbol{\varphi}(\mathbb{M}))$ *is one to one and continuous with its inverse, the topology on* $\boldsymbol{\varphi}(\mathbb{M})$ *being the one induced by* $\mathbb{N}$ *.*

A detailed treatment of these notions of calculus on manifolds may be found in [3] .

1.2.4 Tangent and cotangent functors

Two basic examples of covariant and contravariant functors are provided by the *tangent functors* and *cotangent functors*. In the category of differentiable manifolds morphisms are smooth maps from one manifold to another one.

Definition 1.2.5 (Tangent functor) *The tangent functor, between the category of differentiable manifolds and the category of tangent bundles, is the covariant functor defined by associating with each manifold its tangent bundle and with each morphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *its tangent map* $T\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{TM}\,;\mathbb{TN})$ *.*

The *tangent map* $Tf \in \mathrm{C}^0(\mathbb{TM}\,;\mathbb{T}\Re)$ of a scalar-valued function $f \in \mathrm{C}^1(\mathbb{M}\,;\Re)$ is defined by:

$$Tf \circ \mathbf{v} = (f\,, \mathbf{v}f) \in \mathrm{C}^0(\mathbb{M}\,;\Re \times \Re)\,, \quad \forall\, \mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})\,,$$

where we have canonically identified $\mathbb{T}\Re$ and $\Re \times \Re$. Given a morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ and a function $f \in \mathrm{C}^1(\mathbb{N}\,;\Re)$ the tangent map of the composition $f \circ \boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\Re)$, by the chain rule, is given by:

$$Tf \circ T\boldsymbol{\varphi} \circ \mathbf{v} = T(f \circ \boldsymbol{\varphi}) \circ \mathbf{v}\,, \quad \forall\, \mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})\,.$$

An equivalent definition of $T\boldsymbol{\varphi} \in C^0(\mathbb{TM}\,;\mathbb{TN})$ can be given by requiring the relation above to hold for any $f \in C^1(\mathbb{N}\,;\Re)$. The tangent functor is a covariant functor since:

$$T\mathbf{id}_{\mathbb{M}} = \mathbf{id}_{\mathbb{TM}}\,, \qquad T(\mathbf{g}\circ\mathbf{f}) = T\mathbf{g}\circ T\mathbf{f}\,.$$

for all $\mathbf{f} \in C^1(\mathbb{M}\,;\mathbb{N})$ and $\mathbf{g} \in C^1(\mathbb{N}\,;\mathbb{Y})$. We have the following commutative diagrams:

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\mathbf{id}_{\mathbb{M}}=\mathbf{id}_{\mathbb{TM}}} & \mathbb{TM} \\ \boldsymbol{\tau}_{\mathbb{M}}\big\downarrow & & \big\downarrow\boldsymbol{\tau}_{\mathbb{M}} \\ \mathbb{M} & \xrightarrow{\mathbf{id}_{\mathbb{M}}} & \mathbb{M} \end{array} \iff \quad \boldsymbol{\tau}_{\mathbb{M}}\circ T\mathbf{id}_{\mathbb{M}} = \mathbf{id}_{\mathbb{M}}\circ\boldsymbol{\tau}_{\mathbb{M}} \in C^0(\mathbb{TM}\,;\mathbb{M})\,.$$

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\mathbf{v}} & \mathbb{T}^2\mathbb{M} \\ \boldsymbol{\tau}_{\mathbb{M}}\big\downarrow & & \big\downarrow\boldsymbol{\tau}_{\mathbb{TM}} \\ \mathbb{M} & \xrightarrow{\mathbf{v}} & \mathbb{TM} \end{array} \iff \quad \boldsymbol{\tau}_{\mathbb{TM}}\circ T\mathbf{v} = \mathbf{v}\circ\boldsymbol{\tau}_{\mathbb{M}} \in C^0(\mathbb{TM}\,;\mathbb{TM})\,.$$

Remark 1.2.1 *By acting with the tangent functor on a map* $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ *we get a map* $T\mathbf{v} \in C^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *which is not a section of the bundle* $\boldsymbol{\tau}_{\mathbb{TM}} \in C^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$*. Indeed, for any section* $\mathbf{u} \in C^1(\mathbb{M}\,;\mathbb{TM})$*, we have that* $\boldsymbol{\tau}_{\mathbb{M}}\circ\mathbf{u} = \mathbf{id}_{\mathbb{M}}$ *and hence* $\boldsymbol{\tau}_{\mathbb{TM}}\circ T\mathbf{v}\circ\mathbf{u} = \mathbf{v}\circ\boldsymbol{\tau}_{\mathbb{M}}\circ\mathbf{u} = \mathbf{v}$ *so that* $\boldsymbol{\tau}_{\mathbb{TM}}\circ T\mathbf{v} \neq \mathbf{id}_{\mathbb{TM}}$*. This was to be expected since otherwise we would have* $\boldsymbol{\tau}_{\mathbb{TM}}\circ T\mathbf{v} = \mathbf{v}\circ\boldsymbol{\tau}_{\mathbb{M}} = \mathbf{id}_{\mathbb{TM}}$ *and this is impossible since by acting with the projection* $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}\,;\mathbb{M})$ *all the informations on a tangent vector are lost, except its base point. Acting again with* $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ *all the information are overwritten by that pertaining to the image of this map.* *Concerning this issue see Lemma 1.2.6 on page 50.*

Definition 1.2.6 (Cotangent functor) *The cotangent functor, between the category of differentiable manifolds and the category of cotangent bundles, is the contravariant functor defined by associating with each manifold its cotangent bundle and, with any invertible morphism* $\boldsymbol{\varphi} \in C^1(\mathbb{M}\,;\mathbb{N})$ *the cotangent map* $T^*\boldsymbol{\varphi} \in C^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{T}^*\mathbb{M})$*, according to the identity:*

$$\langle T^*\boldsymbol{\varphi}\circ\mathbf{w}^*, \mathbf{v}\circ\boldsymbol{\varphi}^{-1}\rangle = \langle \mathbf{w}^*, T\boldsymbol{\varphi}\circ\mathbf{v}\circ\boldsymbol{\varphi}^{-1}\rangle\,, \quad \begin{cases} \forall\,\mathbf{v} \in C^0(\mathbb{M}\,;\mathbb{TM})\,, \\ \forall\,\mathbf{w}^* \in C^0(\mathbb{N}\,;\mathbb{T}^*\mathbb{N})\,. \end{cases}$$

The *cotangent map* $T^*\boldsymbol{\varphi} \in C^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{T}^*\mathbb{M})$ is defined pointwise by the linear map $T^*_{\mathbf{x}}\boldsymbol{\varphi} = (T_{\mathbf{x}}\boldsymbol{\varphi})^* \in BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N}\,;\mathbb{T}^*_{\mathbf{x}}\mathbb{M})$ which is the dual of the *tangent map*

$T_{\mathbf{x}}\boldsymbol{\varphi} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N})$, according to the relation:

$$\langle\, (T_{\mathbf{x}}\boldsymbol{\varphi})^* \cdot \mathbf{w}^*_{\boldsymbol{\varphi}(\mathbf{x})}, \mathbf{v}_{\mathbf{x}} \,\rangle = \langle\, \mathbf{w}^*_{\boldsymbol{\varphi}(\mathbf{x})}, T_{\mathbf{x}}\boldsymbol{\varphi} \cdot \mathbf{v}_{\mathbf{x}} \,\rangle\,,$$

where $\mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ and $\mathbf{w}^*_{\boldsymbol{\varphi}(\mathbf{x})} \in \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N}$. The cotangent functor is a contravariant functor, being:

$$T^*\mathbf{id}_{\mathbb{M}} = \mathbf{id}_{\mathbb{T}^*\mathbb{M}}\,,$$

$$T^*(\mathbf{g} \circ \mathbf{f}) = T^*\mathbf{f} \circ T^*\mathbf{g}\,.$$

for all invertible morphisms $\mathbf{f} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ and $\mathbf{g} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{Y})$.

An invertible morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ and the cotangent map $T^*\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{T}^*\mathbb{M})$ are related by the commutative diagram

$$\begin{array}{ccc} \mathbb{T}^*\mathbb{M} & \xleftarrow{\;T^*\boldsymbol{\varphi}\;} & \mathbb{T}^*\mathbb{N} \\ {\scriptstyle \boldsymbol{\tau}^*_{\mathbb{M}}}\big\downarrow & & {\scriptstyle \boldsymbol{\tau}^*_{\mathbb{N}}}\big\downarrow \\ \mathbb{M} & \xleftarrow{\;\boldsymbol{\varphi}^{-1}\;} & \mathbb{N} \end{array} \quad \iff \quad \boldsymbol{\varphi}^{-1} \circ \boldsymbol{\tau}^*_{\mathbb{N}} = \boldsymbol{\tau}^*_{\mathbb{M}} \circ T^*\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{N})\,,$$

where $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$ and $\boldsymbol{\tau}^*_{\mathbb{N}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{N}\,;\mathbb{N})$ are the projection from the cotangent bundles to the base manifolds. The above diagram and formula are quoted in [3], page 566 with a reversed arrow.

In the same way, the cotangent map $T^*\boldsymbol{\varphi} \in \mathrm{C}^0(\boldsymbol{\varphi}(\mathbb{T}^*\mathbb{M})\,;\mathbb{T}^*\mathbb{M})$ may be associated with any injective morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$. A more general definition of the cotangent map associated with any morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ will be provided in Section 1.2.8, Definition 1.2.14 on page 44.

1.2.5 Pull back and push forward

A morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ between two manifolds $\mathbb{M}$ and $\mathbb{N}$ induces a transformation of scalar, vector and that tensor fields defined on $\mathbb{N}$ into corresponding fields on $\mathbb{M}$. These transformations are called *push* and *pull* operations, or *direct* and *inverse images*, associated with the morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$.

For scalar fields the push forward by a morphism simply consists in a change of the base point which leaves invariant the value of the scalar field.

For vector fields the push forward by a morphism changes the base point and is accompanied by a linear transformation which describes the modification of the tangent space due to the action of the morphism. The push forward transformation of covector and tensor fields, which are linear and multilinear forms, is defined so that their scalar values remain invariant.

Let us provide the basic definitions.

Definition 1.2.7 (Relatedness) *Given a morphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$ *, two vector fields* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M};\mathbb{TM})$ *and* $\mathbf{X} \in \mathrm{C}^0(\mathbb{N};\mathbb{TN})$ *are said to be* $\boldsymbol{\varphi}$*-related if the following diagram commutes*

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\boldsymbol{\varphi}} & \mathbb{TN} \\ \mathbf{v}\Big\uparrow & & \Big\uparrow\mathbf{X} \\ \mathbb{M} & \xrightarrow{\boldsymbol{\varphi}} & \mathbb{N} \end{array} \quad \Longleftrightarrow \quad \mathbf{X} \circ \boldsymbol{\varphi} = T\boldsymbol{\varphi} \circ \mathbf{v} \in \mathrm{C}^0(\mathbb{M};\mathbb{TN})\,.$$

Given a morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{M})$, a vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M};\mathbb{TM})$ is said to be $\boldsymbol{\varphi}$-*invariant* if it is $\boldsymbol{\varphi}$-*related* to itself.

- The *pull back* of a *scalar field* $f \in \mathrm{C}^0(\mathbb{N};\Re)$ according to a morphism $\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{M};\mathbb{N})$ is the field $\boldsymbol{\varphi}\!\downarrow\! f \in \mathrm{C}^0(\mathbb{M};\Re)$ which takes at a point $\mathbf{x} \in \mathbb{M}$ the same value taken by $f \in \mathrm{C}^0(\boldsymbol{\varphi}(\mathbb{M});\Re)$ at the point $\boldsymbol{\varphi}(\mathbf{x}) \in \mathbb{N}$:

$$\boldsymbol{\varphi}\!\downarrow\! f := f \circ \boldsymbol{\varphi}\,.$$

- The *push forward* $\boldsymbol{\varphi}\!\uparrow\! f \in \mathrm{C}^0(\boldsymbol{\varphi}(\mathbb{M});\Re)$ of a *scalar field* $f \in \mathrm{C}^0(\mathbb{M};\Re)$, according to an invertible morphism $\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{M};\mathbb{N})$, is the scalar field which takes, at the point $\mathbf{y} \in \mathbb{N}$, the same value taken by $f \in \mathrm{C}^0(\mathbb{M};\Re)$ at the point $\boldsymbol{\varphi}^{-1}(\mathbf{y}) \in \mathbb{M}$:

$$\boldsymbol{\varphi}\!\uparrow\! f := f \circ \boldsymbol{\varphi}^{-1}\,.$$

- The *push forward* $\boldsymbol{\varphi}\!\uparrow\!\mathbf{v} \in \mathrm{C}^0(\mathbb{N};\mathbb{TN})$ of a *vector field* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M};\mathbb{TM})$ according to an invertible morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$ is defined by the commutative diagram:

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\boldsymbol{\varphi}} & \mathbb{TN} \\ \mathbf{v}\Big\uparrow & & \Big\uparrow\boldsymbol{\varphi}\uparrow\mathbf{v} \\ \mathbb{M} & \xleftarrow{\boldsymbol{\varphi}^{-1}} & \mathbb{N} \end{array} \quad \Longleftrightarrow \quad \boldsymbol{\varphi}\!\uparrow\!\mathbf{v} := T\boldsymbol{\varphi} \circ \mathbf{v} \circ \boldsymbol{\varphi}^{-1} \in \mathrm{C}^0(\mathbb{N};\mathbb{TN})\,.$$

  Being $\boldsymbol{\tau}_{\mathbb{N}} \circ T\boldsymbol{\varphi} = \boldsymbol{\varphi} \circ \boldsymbol{\tau}_{\mathbb{M}}$ and $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} = \mathbf{id}_{\mathbb{M}}$, it is

$$\boldsymbol{\tau}_{\mathbb{N}} \circ \boldsymbol{\varphi}\!\uparrow\!\mathbf{v} = \boldsymbol{\tau}_{\mathbb{N}} \circ T\boldsymbol{\varphi} \circ \mathbf{v} \circ \boldsymbol{\varphi}^{-1} = \boldsymbol{\varphi} \circ \boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} \circ \boldsymbol{\varphi}^{-1} = \boldsymbol{\varphi} \circ \boldsymbol{\varphi}^{-1} = \mathbf{id}_{\mathbb{N}}\,.$$

  The push $\boldsymbol{\varphi}\!\uparrow\!\mathbf{v} \in \mathrm{C}^0(\mathbb{N};\mathbb{TN})$ is also called the *image* of $\mathbf{v} \in \mathrm{C}^0(\mathbb{M};\mathbb{TM})$ according to $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$.

- The *push forward* $\boldsymbol{\varphi}{\uparrow}\boldsymbol{\omega} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{T}^*\mathbb{N})$ of a field $\boldsymbol{\omega} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ of *co-vectors*, according to an invertible morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$, is defined by requiring that the evaluation $\langle \boldsymbol{\omega}, \mathbf{v} \rangle$ be invariant when both the vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ and the co-vector field $\boldsymbol{\omega} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ are pushed forward:
$$\langle \boldsymbol{\varphi}{\uparrow}\boldsymbol{\omega}, \boldsymbol{\varphi}{\uparrow}\mathbf{v} \rangle := \boldsymbol{\varphi}{\uparrow}\langle \boldsymbol{\omega}, \mathbf{v} \rangle\,, \quad \forall\, \mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})\,.$$
The push forward is the covector field given by:
$$\langle \boldsymbol{\varphi}{\uparrow}\boldsymbol{\omega} \circ \boldsymbol{\varphi}, T\boldsymbol{\varphi} \circ \mathbf{v} \rangle = \langle \boldsymbol{\omega}, \mathbf{v} \rangle\,, \quad \forall\, \mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})\,.$$

- The *push forward* of a *tensor field* $\mathbf{A} : \mathbb{M} \mapsto BL(\mathbb{T}\mathbb{M}, \mathbb{T}^*\mathbb{M}\,;\Re)$ is defined in a similar way by
$$(\boldsymbol{\varphi}{\uparrow}\mathbf{A})(\boldsymbol{\varphi}{\uparrow}\mathbf{v}, \boldsymbol{\varphi}{\uparrow}\boldsymbol{\omega}) := \boldsymbol{\varphi}{\uparrow}\big(\mathbf{A}(\mathbf{v}, \boldsymbol{\omega})\big)\,, \quad \begin{cases} \forall\, \mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})\,, \\ \forall\, \boldsymbol{\omega} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})\,. \end{cases}$$

- The *pull back* $\boldsymbol{\varphi}{\downarrow}\mathbf{w} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ of a *vector field* $\mathbf{w} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}\mathbb{N})$ according to a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ is defined by the commutative diagram:
$$\begin{array}{ccc} \mathbb{T}\mathbb{M} & \xleftarrow{\;T\boldsymbol{\varphi}^{-1}\;} & \mathbb{T}\mathbb{N} \\ {\scriptstyle \boldsymbol{\varphi}{\downarrow}\mathbf{w}}\Big\uparrow & & \Big\uparrow{\scriptstyle \mathbf{w}} \\ \mathbb{M} & \xrightarrow{\;\boldsymbol{\varphi}\;} & \mathbb{N} \end{array} \qquad \Longleftrightarrow \qquad \boldsymbol{\varphi}{\downarrow}\mathbf{w} = T\boldsymbol{\varphi}^{-1} \circ \mathbf{w} \circ \boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})\,.$$
The pull $\boldsymbol{\varphi}{\downarrow}\mathbf{w} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ is a vector field on $\mathbb{M}$ with values in $\mathbb{T}\mathbb{M}$. Indeed, being $\boldsymbol{\tau}_{\mathbb{M}} \circ T\boldsymbol{\varphi}^{-1} = \boldsymbol{\varphi}^{-1} \circ \boldsymbol{\tau}_{\mathbb{N}}$ and $\boldsymbol{\tau}_{\mathbb{N}} \circ \mathbf{v} = \mathbf{id}_{\mathbb{N}}$, it is
$$\boldsymbol{\tau}_{\mathbb{M}} \circ \boldsymbol{\varphi}{\downarrow}\mathbf{w} = \boldsymbol{\tau}_{\mathbb{M}} \circ T\boldsymbol{\varphi}^{-1} \circ \mathbf{w} \circ \boldsymbol{\varphi} = \boldsymbol{\varphi}^{-1} \circ \boldsymbol{\tau}_{\mathbb{N}} \circ \mathbf{w} \circ \boldsymbol{\varphi} = \mathbf{id}_{\mathbb{M}}\,.$$
The pull $\boldsymbol{\varphi}{\downarrow}\mathbf{w} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ is also called the *inverse image* of $\mathbf{w} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}\mathbb{N})$ according to the diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$.

- The *pull back* of a *covector* field $\boldsymbol{\omega} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}^*\mathbb{N})$ according to a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ is defined by requiring that the evaluation $\langle \boldsymbol{\omega}, \mathbf{w} \rangle$ be invariant when both the vector field $\mathbf{w} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}\mathbb{N})$ and the covector field $\boldsymbol{\omega} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}^*\mathbb{N})$ are pulled back:
$$\langle \boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega}, \boldsymbol{\varphi}{\downarrow}\mathbf{w} \rangle := \boldsymbol{\varphi}{\downarrow}\langle \boldsymbol{\omega}, \mathbf{w} \rangle\,, \quad \forall\, \mathbf{w} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}\mathbb{N})\,.$$

The pull back $\boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ is then the covector field given by:

$$\langle\, \boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega} \circ \boldsymbol{\varphi}^{-1}, T\boldsymbol{\varphi}^{-1} \circ \mathbf{v}\,\rangle = \langle\, \boldsymbol{\omega}, \mathbf{w}\,\rangle\,, \quad \forall\, \mathbf{w} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{TN})\,.$$

Proposition 1.2.1 *Let* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *be an invertible morphism. Then the pull back* $\boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ *of a covector field* $\boldsymbol{\omega} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}^*\mathbb{N})$ *is given by*

$$\boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega} := T^*\boldsymbol{\varphi} \circ \boldsymbol{\omega} \circ \boldsymbol{\varphi}\,.$$

Proof. Let us recall that $\boldsymbol{\varphi}{\uparrow}\mathbf{v} \circ \boldsymbol{\varphi} = T\boldsymbol{\varphi} \circ \mathbf{v}$. Then, for any $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$, we have that:

$$\begin{aligned}\langle\, \boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega}, \mathbf{v}\,\rangle &:= \boldsymbol{\varphi}{\downarrow}\langle\, \boldsymbol{\omega}, \boldsymbol{\varphi}{\uparrow}\mathbf{v}\,\rangle = \langle\, \boldsymbol{\omega} \circ \boldsymbol{\varphi}, \boldsymbol{\varphi}{\uparrow}\mathbf{v} \circ \boldsymbol{\varphi}\,\rangle \\ &= \langle\, \boldsymbol{\omega} \circ \boldsymbol{\varphi}, T\boldsymbol{\varphi} \circ \mathbf{v}\,\rangle = \langle\, T^*\boldsymbol{\varphi} \circ \boldsymbol{\omega} \circ \boldsymbol{\varphi}, \mathbf{v}\,\rangle\,,\end{aligned}$$

and the statement is proved. ■

In the literature it is customary to denote push-forward and pull-back operations according to a morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ by the symbols $\boldsymbol{\varphi}_*$ and $\boldsymbol{\varphi}^*$ but then too many stars appear in the geometrical sky (duality, Hodge operator). So we decided to adopt a new, more expressive and peculiar notation.

If $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ is a diffeomorphism, the *pull back*, or *inverse image*, of scalar, vector, co-vector and tensor fields on $\mathbb{N}$ along $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ is the push forward along the inverse diffeomorphism $\boldsymbol{\varphi}^{-1} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{M})$.

The *pull back* is denoted by the symbol

$$\boldsymbol{\varphi}{\downarrow} = (\boldsymbol{\varphi}^{-1}){\uparrow}$$

and hence we have that $\boldsymbol{\varphi}{\downarrow} \circ \boldsymbol{\varphi}{\uparrow} = \mathbf{I}_{\mathbb{M}}\,, \quad \boldsymbol{\varphi}{\uparrow} \circ \boldsymbol{\varphi}{\downarrow} = \mathbf{I}_{\mathbb{N}}$ where $\mathbf{I}_{\mathbb{M}}$ and $\mathbf{I}_{\mathbb{N}}$ are identity maps acting on scalar, vector, co-vector and tensor fields on $\mathbb{M}$ and $\mathbb{N}$ respectively. Let us prove this property for scalar and vector fields.

- For scalar fields $f : \mathbb{M} \mapsto \Re$ and $g : \mathbb{N} \mapsto \Re$ we have that

  $$\boldsymbol{\varphi}{\uparrow}f \circ \boldsymbol{\varphi} = f\,, \qquad \boldsymbol{\varphi}{\downarrow}g = g \circ \boldsymbol{\varphi}\,,$$

  and hence

  $$\begin{aligned}\boldsymbol{\varphi}{\downarrow}(\boldsymbol{\varphi}{\uparrow}f) &= \boldsymbol{\varphi}{\uparrow}f \circ \boldsymbol{\varphi} = f\,, \\ \boldsymbol{\varphi}{\uparrow}(\boldsymbol{\varphi}{\downarrow}g) &= \boldsymbol{\varphi}{\downarrow}g \circ \boldsymbol{\varphi}^{-1} = g\,.\end{aligned}$$

- For vector fields $\mathbf{u} : \mathbb{M} \mapsto \mathbb{TM}$ and $\mathbf{v} : \mathbb{N} \mapsto \mathbb{TN}$ we have that

$$(\boldsymbol{\varphi}{\uparrow}\mathbf{u}) \circ \boldsymbol{\varphi} := T\boldsymbol{\varphi} \circ \mathbf{u} \,,$$
$$(\boldsymbol{\varphi}{\downarrow}\mathbf{v}) \circ \boldsymbol{\varphi}^{-1} := T\boldsymbol{\varphi}^{-1} \circ \mathbf{v} \,,$$

and hence

$$\boldsymbol{\varphi}{\downarrow}\boldsymbol{\varphi}{\uparrow}\mathbf{u} = T\boldsymbol{\varphi}^{-1} \circ (\boldsymbol{\varphi}{\uparrow}\mathbf{u}) \circ \boldsymbol{\varphi} = T\boldsymbol{\varphi}^{-1} \circ T\boldsymbol{\varphi} \circ \mathbf{u} = \mathbf{u} \,,$$
$$\boldsymbol{\varphi}{\uparrow}\boldsymbol{\varphi}{\downarrow}\mathbf{v} = T\boldsymbol{\varphi} \circ (\boldsymbol{\varphi}{\downarrow}\mathbf{v}) \circ \boldsymbol{\varphi}^{-1} = T\boldsymbol{\varphi} \circ T\boldsymbol{\varphi}^{-1} \circ \mathbf{v} = \mathbf{v} \,.$$

Proposition 1.2.2 *Given two morphisms* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *and* $\boldsymbol{\phi} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{Q})$ *, the push forward of a vector field* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *fulfils the direct chain rule:*

$$(\boldsymbol{\phi} \circ \boldsymbol{\varphi}){\uparrow}\mathbf{v} = (\boldsymbol{\phi}{\uparrow} \circ \boldsymbol{\varphi}{\uparrow})\mathbf{v} \,.$$

Proof. Being $\boldsymbol{\varphi}{\uparrow}\mathbf{v} \circ \boldsymbol{\varphi} = T\boldsymbol{\varphi} \circ \mathbf{v}$, we have that

$$\begin{aligned}(\boldsymbol{\phi} \circ \boldsymbol{\varphi}){\uparrow}\mathbf{v} \circ (\boldsymbol{\phi} \circ \boldsymbol{\varphi}) &= T(\boldsymbol{\phi} \circ \boldsymbol{\varphi}) \circ \mathbf{v} = T\boldsymbol{\phi} \circ T\boldsymbol{\varphi} \circ \mathbf{v} \\ &= \boldsymbol{\phi}{\uparrow}(\boldsymbol{\varphi}{\uparrow}\mathbf{v}) \circ \boldsymbol{\varphi}) \circ \boldsymbol{\phi} = (\boldsymbol{\phi}{\uparrow} \circ \boldsymbol{\varphi}{\uparrow})\mathbf{v} \circ (\boldsymbol{\phi} \circ \boldsymbol{\varphi}) \,,\end{aligned}$$

and the rule is proven. ■

Proposition 1.2.3 *Given two morphisms* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *and* $\boldsymbol{\phi} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{Q})$ *, the pull back of a scalar field* $f \in \mathrm{C}^0(\mathbb{Q}\,;\Re)$ *and of a covector field* $\mathbf{v}^* \in \mathrm{C}^0(\mathbb{Q}\,;\mathbb{T}^*\mathbb{Q})$ *fulfil the reverse chain rules:*

$$(\boldsymbol{\phi} \circ \boldsymbol{\varphi}){\downarrow}f = (\boldsymbol{\varphi}{\downarrow} \circ \boldsymbol{\phi}{\downarrow})f$$
$$(\boldsymbol{\phi} \circ \boldsymbol{\varphi}){\downarrow}\mathbf{v}^* = (\boldsymbol{\varphi}{\downarrow} \circ \boldsymbol{\phi}{\downarrow})\mathbf{v}^* \,.$$

Proof. Being $\boldsymbol{\varphi}{\downarrow}f = f \circ \boldsymbol{\varphi}$ and $\boldsymbol{\varphi}{\downarrow}\mathbf{v}^* = T^*\boldsymbol{\varphi} \circ \mathbf{v} \circ \boldsymbol{\varphi}$, we have that

$$\begin{aligned}(\boldsymbol{\phi} \circ \boldsymbol{\varphi}){\downarrow}f &= f \circ (\boldsymbol{\phi} \circ \boldsymbol{\varphi}) = (f \circ \boldsymbol{\phi}) \circ \boldsymbol{\varphi} \\ &= \boldsymbol{\varphi}{\downarrow}(\boldsymbol{\phi}{\downarrow}f) = (\boldsymbol{\varphi}{\downarrow} \circ \boldsymbol{\phi}{\downarrow})f \,,\end{aligned}$$

$$\begin{aligned}(\boldsymbol{\phi} \circ \boldsymbol{\varphi}){\downarrow}\mathbf{v}^* &= T^*(\boldsymbol{\phi} \circ \boldsymbol{\varphi}) \circ \mathbf{v}^* \circ (\boldsymbol{\phi} \circ \boldsymbol{\varphi}) = T^*\boldsymbol{\varphi} \circ T^*\boldsymbol{\phi} \circ \mathbf{v}^* \circ (\boldsymbol{\phi} \circ \boldsymbol{\varphi}) \\ &= \boldsymbol{\varphi}{\downarrow}(\boldsymbol{\phi}{\downarrow}\mathbf{v}^*) = (\boldsymbol{\varphi}{\downarrow} \circ \boldsymbol{\phi}{\downarrow})\mathbf{v}^* \,,\end{aligned}$$

and the rules are proven. ■

If $\varphi \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ and $\phi \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{Q})$ are diffeomorphisms, we have that:

$$(\phi \circ \varphi){\downarrow} = ((\phi \circ \varphi)^{-1}){\uparrow} = (\varphi^{-1} \circ \phi^{-1}){\uparrow} = \varphi{\downarrow} \circ \phi{\downarrow}\,.$$

The next proposition states that the directional derivative is *natural* with respect to the push. A more general result concerning the LIE derivative will be provided in Proposition 1.3.4.

Proposition 1.2.4 (Push of the directional derivative) *Let* $\varphi \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *be a diffeomorphism,* $f \in \mathrm{C}^1(\mathbb{M}\,;\Re)$ *,* $g \in \mathrm{C}^1(\mathbb{N}\,;\Re)$ *be scalar functions and* $\mathbf{v} : \mathbb{M} \mapsto \mathbb{TM}$ *,* $\mathbf{u} : \mathbb{N} \mapsto \mathbb{TN}$ *be vector fields. Then we have that*

$$\varphi{\uparrow}(\mathbf{v}\,f) = (\varphi{\uparrow}\mathbf{v})\,(\varphi{\uparrow}f)\,, \quad \forall\,\mathbf{v} : \mathbb{M} \mapsto \mathbb{TM}\,,$$

$$\varphi{\downarrow}(\mathbf{u}\,g) = (\varphi{\downarrow}\mathbf{u})\,(\varphi{\downarrow}g)\,, \quad \forall\,\mathbf{u} : \mathbb{N} \mapsto \mathbb{TN}\,.$$

Proof. The former equality is proven as follows

$$(\varphi{\uparrow}\mathbf{v})\,(\varphi{\uparrow}f) \circ \varphi = \mathbf{v}\,(\varphi{\uparrow}f \circ \varphi) = \mathbf{v}\,f = \varphi{\uparrow}(\mathbf{v}\,f) \circ \varphi\,.$$

The latter equality is obtained in an analogous way. ■

The definition of the push forward of a covector field and Proposition 1.2.4 imply that

$$(\varphi{\uparrow}Tf)\,(\varphi{\uparrow}\mathbf{v}) := \varphi{\uparrow}(Tf\,(\mathbf{v})) = \varphi{\uparrow}(\mathbf{v}\,f) = T(\varphi{\uparrow}f)\,(\varphi{\uparrow}\mathbf{v})\,, \quad \forall\,\mathbf{v} : \mathbb{M} \mapsto \mathbb{TM}\,,$$

that is $\varphi{\uparrow}(Tf) = T(\varphi{\uparrow}f)$. Analogously we get that $\varphi{\downarrow}(Tg) = T(\varphi{\downarrow}g)$.

Remark 1.2.2 *Another useful formula is*

$$\varphi{\uparrow}(f\,\mathbf{v}) = (\varphi{\uparrow}f)\,(\varphi{\uparrow}\mathbf{v})\,.$$

The proof follows from the relations

$$\begin{aligned}(\varphi{\uparrow}(f\,\mathbf{v}))k \circ \varphi &= (f\,\mathbf{v})(k \circ \varphi) = f\,(\mathbf{v}(k \circ \varphi))\\ &= (f \circ \varphi^{-1} \circ \varphi)\,((\varphi{\uparrow}\mathbf{v})k \circ \varphi)\\ &= (\varphi{\uparrow}f)\,(\varphi{\uparrow}\mathbf{v})k \circ \varphi\,.\end{aligned}$$

Despite of its resemblance to the formula in Proposition 1.2.4, it is to be stressed that here the field $f\,\mathbf{v}$ *is simply the product between the scalar field* f *and the* *vector field* $\mathbf{v}$ *. In greater generality, let* $\boldsymbol{\mu}\mathbf{a}$ *be the contraction of a tensor* $\boldsymbol{\mu} \in BL(\mathbb{TM}^2 ; \Re)$ *with a vector* $\mathbf{a} \in \mathbb{TM}$ *, defined by:*

$$\boldsymbol{\mu}\mathbf{a} \cdot \mathbf{b} := \boldsymbol{\mu}(\mathbf{a}, \mathbf{b}) , \quad \forall\, \mathbf{b} \in \mathbb{TM} .$$

Then the push of the contraction is equal to the contraction of the pushes:

$$\boldsymbol{\varphi}{\uparrow}(\boldsymbol{\mu}\mathbf{a}) = \boldsymbol{\varphi}{\uparrow}\boldsymbol{\mu}\, \boldsymbol{\varphi}{\uparrow}\mathbf{a} .$$

Indeed

$$\begin{aligned}(\boldsymbol{\varphi}{\uparrow}\boldsymbol{\mu}\mathbf{a}) \cdot \boldsymbol{\varphi}{\uparrow}\mathbf{b} &= \boldsymbol{\varphi}{\uparrow}(\boldsymbol{\mu}\mathbf{a} \cdot \mathbf{b}) = \boldsymbol{\varphi}{\uparrow}(\boldsymbol{\mu}(\mathbf{a}, \mathbf{b})) \\ &= (\boldsymbol{\varphi}{\uparrow}\boldsymbol{\mu})(\boldsymbol{\varphi}{\uparrow}\mathbf{a}, \boldsymbol{\varphi}{\uparrow}\mathbf{b}) = (\boldsymbol{\varphi}{\uparrow}\boldsymbol{\mu}\, \boldsymbol{\varphi}{\uparrow}\mathbf{a}) \cdot \boldsymbol{\varphi}{\uparrow}\mathbf{b} .\end{aligned}$$

1.2.6 Invariance under diffeomorphisms

A fundamental property of the duality pairing, which has basic implications in physical applications, is that it is invariant under diffeomorphic transformations. To understand this fact we must briefly recall how tensors change under diffeomorphisms.

Let us consider a morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M} ; \mathbb{N})$ between the manifolds $\mathbb{M}$ and $\mathbb{N}$ which are embedded in a larger riemannian manifold $\mathbb{S}$, i.e. a manifold endowed with a field $\mathbf{g} \in \mathrm{C}^1(\mathbb{S} ; BL(\mathbb{TS}^2 ; \Re))$ of metric tensors.

We put the question: which is the vector associated with the push-forward of a form associated with a given vector? A direct computation provides the answer. Indeed, being:

$$\begin{aligned}&T\boldsymbol{\varphi}(\mathbf{m}) \in \mathrm{C}^1(\mathbb{T}_{\mathbf{m}}\mathbb{M} ; \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{m})}\mathbb{N}) , \\ &T\boldsymbol{\varphi}^T(\boldsymbol{\varphi}(\mathbf{m})) \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{m})}\mathbb{N} ; \mathbb{T}_{\mathbf{m}}\mathbb{M}) , \\ &T\boldsymbol{\varphi}^{-1}(\boldsymbol{\varphi}(\mathbf{m})) \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{m})}\mathbb{N} ; \mathbb{T}_{\mathbf{m}}\mathbb{M}) , \\ &T\boldsymbol{\varphi}^{-T}(\mathbf{m}) \in \mathrm{C}^1(\mathbb{T}_{\mathbf{m}}\mathbb{M} ; \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{m})}\mathbb{N}) ,\end{aligned}$$

we have that:

$$\langle \boldsymbol{\varphi}{\uparrow}(\mathbf{g}_{\mathbf{m}}\mathbf{a}), \mathbf{w} \rangle = \langle \mathbf{g}_{\mathbf{m}}\mathbf{a}, T\boldsymbol{\varphi}^{-1}{\cdot}\mathbf{w} \rangle \circ \boldsymbol{\varphi}^{-1} = \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}(T\boldsymbol{\varphi}^{-T}{\cdot}\mathbf{a}, \mathbf{w}) , \quad \begin{aligned}&\forall\, \mathbf{a} \in \mathbb{T}_{\mathbf{m}}\mathbb{M} , \\ &\forall\, \mathbf{w} \in \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{m})}\mathbb{N} ,\end{aligned}$$

which can be written: $\boldsymbol{\varphi}{\uparrow}(\mathbf{g}_{\mathbf{m}}\mathbf{a}) = \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}(T\boldsymbol{\varphi}^{-T}\mathbf{a})$.

The pull-back of $\boldsymbol{\beta} \in BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{m})}\mathbb{N}^2 ; \Re)$ is computed as follows:

$$
\begin{aligned}
(\boldsymbol{\varphi}\downarrow\boldsymbol{\beta})(\mathbf{a},\mathbf{b}) = \boldsymbol{\beta}(T\boldsymbol{\varphi}\cdot\mathbf{a}, T\boldsymbol{\varphi}\cdot\mathbf{b}) &= \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}(\mathbf{g}_{\mathbf{m}}^{-1}\boldsymbol{\beta}\, T\boldsymbol{\varphi}\cdot\mathbf{a}, T\boldsymbol{\varphi}\cdot\mathbf{b}) \\
&= \mathbf{g}_{\mathbf{m}}(T\boldsymbol{\varphi}^T\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}^{-1}\boldsymbol{\beta}\, T\boldsymbol{\varphi}\cdot\mathbf{a}, \mathbf{b}) \,, \quad \forall\, \mathbf{a},\mathbf{b} \in \mathbb{T}_{\mathbf{m}}\mathbb{M} \,,
\end{aligned}
$$

so that $\mathbf{g}_{\mathbf{m}}^{-1}(\boldsymbol{\varphi}\downarrow\boldsymbol{\beta}) = T\boldsymbol{\varphi}^T(\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}^{-1}\boldsymbol{\beta})\, T\boldsymbol{\varphi}$.

The pull-back of $\boldsymbol{\alpha}^* \in BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{m})}\mathbb{N}^2 ; \Re)$ is evaluated as:

$$
\begin{aligned}
(\boldsymbol{\varphi}\downarrow\boldsymbol{\alpha}^*)(\mathbf{g}_{\mathbf{m}}\mathbf{a}, \mathbf{g}_{\mathbf{m}}\mathbf{b}) &= \boldsymbol{\alpha}^*(\boldsymbol{\varphi}\uparrow(\mathbf{g}_{\mathbf{m}}\mathbf{a}), \boldsymbol{\varphi}\uparrow(\mathbf{g}_{\mathbf{m}}\mathbf{b})) \\
&= \boldsymbol{\alpha}^*(\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}(T\boldsymbol{\varphi}^{-T}\mathbf{a}), \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}(T\boldsymbol{\varphi}^{-T}\mathbf{b})) \\
&= \langle (\boldsymbol{\alpha}^*\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}) \cdot T\boldsymbol{\varphi}^{-T}\mathbf{a}, \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}(T\boldsymbol{\varphi}^{-T}\mathbf{b}) \rangle \\
&= \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}((\boldsymbol{\alpha}^*\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}) \cdot T\boldsymbol{\varphi}^{-T}\mathbf{a}, T\boldsymbol{\varphi}^{-T}\mathbf{b}) \\
&= \mathbf{g}_{\mathbf{m}}(T\boldsymbol{\varphi}^{-1}(\boldsymbol{\alpha}^*\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})})\, T\boldsymbol{\varphi}^{-T}\mathbf{a}, \mathbf{b}) \,, \quad \forall\, \mathbf{a},\mathbf{b} \in \mathbb{T}_{\mathbf{m}}\mathbb{M} \,,
\end{aligned}
$$

and hence $(\boldsymbol{\varphi}\downarrow\boldsymbol{\alpha}^*)\, \mathbf{g}_{\mathbf{m}} = T\boldsymbol{\varphi}^{-1}(\boldsymbol{\alpha}^*\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})})\, T\boldsymbol{\varphi}^{-T}$ and the invariance property:

$$
\begin{aligned}
\langle \boldsymbol{\varphi}\downarrow\boldsymbol{\alpha}^*, \boldsymbol{\varphi}\downarrow\boldsymbol{\beta} \rangle &= \langle T\boldsymbol{\varphi}^{-1}(\boldsymbol{\alpha}^*\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})})\, T\boldsymbol{\varphi}^{-T}, T\boldsymbol{\varphi}^T(\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}^{-1}\boldsymbol{\beta})\, T\boldsymbol{\varphi} \rangle_{\mathbf{g}} \\
&= \langle \boldsymbol{\alpha}^*\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}, \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{m})}^{-1}\boldsymbol{\beta} \rangle_{\mathbf{g}} \circ \boldsymbol{\varphi} = \langle \boldsymbol{\alpha}^*, \boldsymbol{\beta} \rangle \circ \boldsymbol{\varphi} \,.
\end{aligned}
$$

1.2.7 Flows and vector fields

Let us first consider the case of time independent vector fields.

The integral curve of a vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M} ; \mathbb{TM})$ passing thru $\mathbf{x} \in \mathbb{M}$ for $\lambda = 0$ is the unique curve $\mathbf{c} \in \mathrm{C}^1(I ; \mathbb{M})$ solution of the differential equation

$$\partial_{\mu=\lambda}\, \mathbf{c}(\mu) = \mathbf{v}(\mathbf{c}(\lambda)) \,, \quad \lambda \in I \,,$$

under the initial condition $\mathbf{c}(0) = \mathbf{x} \in \mathbb{M}$.

- The *flow* associated with the vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M} ; \mathbb{TM})$ is the application

  $$\mathbf{Fl}^{\mathbf{v}} : \mathbb{M} \times I \mapsto \mathbb{M} \,,$$

  such that $\partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{v}$ equivalent to

  $$\partial_{\mu=\lambda}\, \mathbf{Fl}^{\mathbf{v}}_{\mu} = \mathbf{v} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \,, \quad \lambda \in I \,.$$

Then $\mathbf{c}(\lambda) = \mathbf{Fl}^{\mathbf{v}}_{\lambda}(\mathbf{x})$ is the integral curve of the vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ passing thru $\mathbf{x} \in \mathbb{M}$ for $\lambda = 0$.

By uniqueness of the integral curve, the following *group property* holds

$$\mathbf{Fl}^{\mathbf{v}}_{\mu} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{v}}_{\mu} = \mathbf{Fl}^{\mathbf{v}}_{\lambda+\mu}\,.$$

Since

$$\mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{v}}_{-\lambda} = \mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{Fl}^{\mathbf{v}}_{0} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})\,,$$

is the identity map, we infer that

$$\mathbf{Fl}^{\mathbf{v}}_{-\lambda} = (\mathbf{Fl}^{\mathbf{v}}_{\lambda})^{-1}\,.$$

Proposition 1.2.5 (Flows of morphism-related vector fields) *The vector fields* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *and* $\mathbf{X} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{TN})$ *are related by the morphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$, *according to the commutative diagram*

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\boldsymbol{\varphi}} & \mathbb{TN} \\ \mathbf{v}\uparrow & & \uparrow\mathbf{X} \\ \mathbb{M} & \xrightarrow{\boldsymbol{\varphi}} & \mathbb{N} \end{array} \iff \quad \mathbf{X} \circ \boldsymbol{\varphi} = T\boldsymbol{\varphi} \circ \mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TN})\,,$$

if and only if the flows $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *and* $\mathbf{Fl}^{\mathbf{X}}_{\lambda} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{N})$ *are related by*

$$\boldsymbol{\varphi} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{Fl}^{\mathbf{X}}_{\lambda} \circ \boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})\,.$$

Proof. Taking the derivatives:

$$\partial_{\lambda=0}\, \boldsymbol{\varphi} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} = T\boldsymbol{\varphi} \circ \mathbf{v}\,,$$

$$\partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{X}}_{\lambda} \circ \boldsymbol{\varphi} = \mathbf{X} \circ \boldsymbol{\varphi}\,,$$

we see that the speed of the flow $\mathbf{Fl}^{\mathbf{X}}_{\lambda}$ at the point $\boldsymbol{\varphi}(\mathbf{x}) \in \mathbb{N}$ is equal to the speed of the flow $\boldsymbol{\varphi} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}$ at $\mathbf{x} \in \mathbb{M}$. The converse result follows from the uniqueness of the solution of the differential equation defining the flow. ■

If $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ is a diffeomorphism, it is natural to give the following definition (see fig 1.11).

- The *push of the flow* $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ thru $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ is the flow $\boldsymbol{\varphi}{\uparrow}\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{N})$ defined by

$$\boldsymbol{\varphi}{\uparrow}\mathbf{Fl}^{\mathbf{v}}_{\lambda} = \boldsymbol{\varphi} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \boldsymbol{\varphi}^{-1}\,.$$

The result of Proposition 1.2.5 can then be stated as follows.

- The flow of the push is equal to the push of the flow:

$$\boldsymbol{\varphi}\!\uparrow\!\mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{Fl}^{\boldsymbol{\varphi}\uparrow\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{N})\,.$$

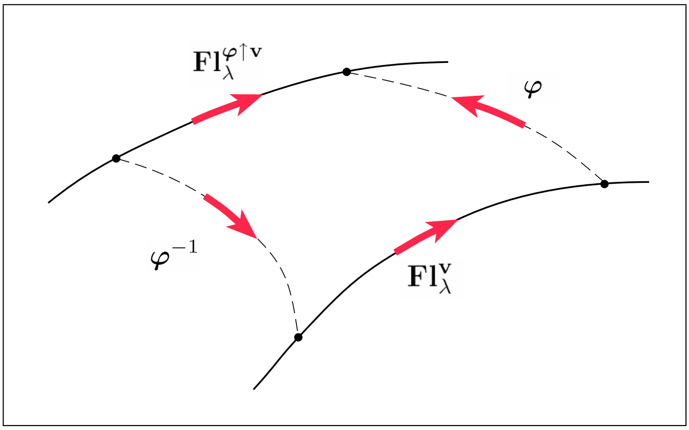

Figure 1.11: Push of a flow

In the special case $\mathbb{N} = \mathbb{M}$ we get the following.

Corollary 1.2.1 (Drag and commutation) *The push induced by an invertible morphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *drags a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *if and only if the morphism commutes with the flow of the field, that is*

$$\mathbf{v} = \boldsymbol{\varphi}\!\uparrow\!\mathbf{v} \iff \boldsymbol{\varphi} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \boldsymbol{\varphi}\,.$$

We underline that the equality $\mathbf{v} = \boldsymbol{\varphi}\!\uparrow\!\mathbf{v}$, expressing the property that the morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ drags the vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, means that the push forward $\boldsymbol{\varphi}\!\uparrow\!\mathbf{v}$ of a point value $\{\mathbf{x}\,,\mathbf{v}\}$ at $\mathbf{x} \in \mathbb{M}$ of the vector field $\mathbf{v}$ is equal to the point value $\{\boldsymbol{\varphi}(\mathbf{x})\,,\mathbf{v}\}$ at $\boldsymbol{\varphi}(\mathbf{x}) \in \mathbb{M}$ of the vector field $\mathbf{v}$.

In particular, setting $\boldsymbol{\varphi} = \mathbf{Fl}^{\mathbf{v}}_{\lambda}$, we obtain that

- A tangent vector field is dragged by its flow, i.e. $\mathbf{v} = \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\uparrow\!\mathbf{v}$.

1.2.8 Fibred manifolds and bundles

A comprehensive treatment of fibred manifolds can be found in [177]. Basic definitions and some results will be summarized hereafter.

Definition 1.2.8 *A fibred manifold is a triple* $\{\mathbb{E}, \mathbf{p}, \mathbb{M}\}$ *where* $\mathbb{E}$ *and* $\mathbb{M}$ *are manifolds and* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *is an surjective submersion called the projection. The manifold* $\mathbb{E}$ *is called the total space and* $\mathbb{M}$ *the base space. For each* $\mathbf{m} \in \mathbb{M}$ *the subset* $\mathbf{p}^{-1}(\mathbf{m})$ *is called the fibre over* $\mathbf{m}$ *and is denoted by* $\mathbb{E}_{\mathbf{m}}$ *.*

A fibred manifold may also be denoted by its projection $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$. In a fibred manifold the fibres over points of the base space may have quite different topological properties. In most applications it is however natural to require that fibres be related by diffeomorphic relations. This leads to the definition of a fibre bundle.

Definition 1.2.9 *A fibre bundle* $\{\mathbb{E}, \mathbf{p}, \mathbb{M}, \mathbb{F}\}$ *with typical fibre* $\mathbb{F}$ *is a fibred manifold, with total space* $\mathbb{E}$ *and projection* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *on the base space* $\mathbb{M}$ *, which is locally a cartesian product.*

This precisely means that the C^k-manifold $\mathbb{M}$ has an open atlas

$$\{\{U_i, \varphi_i\} \,|\, i \in I\}\,,$$

such that for each $i \in I$ there is a C^k-diffeomorphism $\phi_i : \mathbf{p}^{-1}(U_i) \mapsto U_i \times \mathbb{F}$ with $\boldsymbol{\pi}_i \circ \phi_i = \mathbf{p}$, being $\boldsymbol{\pi}_i : U_i \times \mathbb{F} \mapsto U_i$ the canonical projection on the first element.

- A manifold $\mathbb{E}$ which is a cartesian product $\mathbb{E} = \mathbb{M} \times \mathbb{F}$ is called a *trivial fibre bundle*.
- A *vector bundle* is a fibre bundle in which the fibre $\mathbb{F}$ is a vector space.

The *tangent bundle* $\mathbb{TM}$ to a manifold $\mathbb{M}$ is a vector bundle, with projection $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$, whose fibres are the tangent spaces to $\mathbb{M}$.

- A *fibre bundle morphism* $\boldsymbol{\chi} : \mathbb{E} \mapsto \mathbb{Z}$ between two fibre bundles $\mathbf{p}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ and $\mathbf{p}_{\mathbb{N}} \in \mathrm{C}^1(\mathbb{Z}\,;\mathbb{N})$ is a morphism which is *fibre preserving*:

$$\mathbf{p}_{\mathbb{M}}(\mathbf{a}) = \mathbf{p}_{\mathbb{M}}(\mathbf{b}) \implies (\mathbf{p}_{\mathbb{N}} \circ \boldsymbol{\chi})(\mathbf{a}) = (\mathbf{p}_{\mathbb{N}} \circ \boldsymbol{\chi})(\mathbf{b})\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{E}\,.$$

A fibre bundle morphism $\boldsymbol{\chi} : \mathbb{E} \mapsto \mathbb{Z}$ induces a *base morphism* $\boldsymbol{\varphi} : \mathbb{M} \mapsto \mathbb{N}$ according to the commutative diagram:

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{\boldsymbol{\chi}} & \mathbb{Z} \\ {\scriptstyle \mathbf{p}_{\mathbb{M}}}\downarrow & & \downarrow{\scriptstyle \mathbf{p}_{\mathbb{N}}} \\ \mathbb{M} & \xrightarrow{\boldsymbol{\varphi}} & \mathbb{N} \end{array} \qquad \text{with} \qquad \mathbf{p}_{\mathbb{N}} \circ \boldsymbol{\chi} = \boldsymbol{\varphi} \circ \mathbf{p}_{\mathbb{M}} \,.$$

It is also said that $\boldsymbol{\chi} : \mathbb{E} \mapsto \mathbb{Z}$ is a fibre bundle morphism over the base morphism $\boldsymbol{\varphi} : \mathbb{M} \mapsto \mathbb{N}$. More precisely it is the pair $(\boldsymbol{\chi}\,,\boldsymbol{\varphi})$, fulfilling the commutativity property above, which defines a fibre bundle morphism from $\mathbf{p}_{\mathbb{M}}$ to $\mathbf{p}_{\mathbb{N}}$ [177].

- A fibre bundle morphism from a fibre bundle $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ to itself is called an *endomorphism*.
- A *vector bundle homomorphism* $\boldsymbol{\chi} : \mathbb{E} \mapsto \mathbb{H}$ between two vector bundles $\mathbf{p}_{\mathbb{M}} \in C^1(\mathbb{E}\,;\mathbb{M})$ and $\mathbf{p}_{\mathbb{N}} \in C^1(\mathbb{H}\,;\mathbb{N})$ is a fibre bundle morphism which is fibre linear.
- A *vector bundle isomorphism* is an invertible homomorphism.
- A *vector bundle automorphism* is an invertible endomorphism.
- A *section* of the fibre bundle $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ is a right-inverse map $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$, that is:

  $$\mathbf{p} \circ \mathbf{s} = \mathbf{id}_{\mathbb{M}} \,,$$

  where $\mathbf{id}_{\mathbb{M}} \in C^1(\mathbb{M}\,;\mathbb{M})$ is the identity map. Tangent vector fields $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ are sections of the tangent vector bundle $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}\,;\mathbb{M})$ since they meet the property $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} = \mathbf{id}_{\mathbb{M}}$.
- A *section along a map* $\mathbf{f} \in C^1(\mathbb{N}\,;\mathbb{M})$ of the fibre bundle $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ is a map $\mathbf{s} \in C^1(\mathbb{N}\,;\mathbb{TE})$ such that:

  $$\mathbf{p} \circ \mathbf{s} = \mathbf{f} \,.$$

Let $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ be a fibre bundle and $T\mathbf{p} \in C^1(\mathbb{TE}\,;\mathbb{TM})$ the lifted fibre bundle by the tangent functor. The manifold $\mathbb{TE}$ has also the vector bundle structure of a tangent bundle denoted by $\boldsymbol{\tau}_{\mathbb{E}} \in C^1(\mathbb{TE}\,;\mathbb{E})$.

Definition 1.2.10 (Projectable vector fields) *A vector field* $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ *tangent to a fibre bundle* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ *is said to be projectable if there exists a vector field* $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ *which completes the commutative diagram:*

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{\ \mathbf{X}\ } & \mathbb{TE} \\ {\scriptstyle \mathbf{p}}\downarrow & & \downarrow{\scriptstyle T\mathbf{p}} \\ \mathbb{M} & \xrightarrow{\ \mathbf{v}\ } & \mathbb{TM} \end{array} \quad \Longleftrightarrow \quad T\mathbf{p} \circ \mathbf{X} = \mathbf{v} \circ \mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{TM})\,.$$

We underline that the map $\mathbf{v} \circ \mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{TM})$ is fibrewise constant in $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$. Then projectability means that the tangent vectors $\mathbf{X}(\mathbf{e}) \in \mathbb{E}_{\mathbf{x}}$ based at points of a fibre $\mathbb{E}_{\mathbf{x}}$ have all the same base velocity $T\mathbf{p} \circ \mathbf{X}(\mathbf{e}) = \mathbf{v}_{\mathbf{x}}$. An equivalent definition is that the map $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ is projectable if there exists a map $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ such that the pair $(\mathbf{X}\,,\mathbf{v})$ is a bundle morphism from $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ to $T\mathbf{p} \in C^1(\mathbb{TE}\,;\mathbb{TM})$. If the map $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ is a tangent vector field, then also the map $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ is a tangent vector field. Indeed being

$$\begin{array}{ccc} \mathbb{E} & \xleftarrow{\ \boldsymbol{\tau}_{\mathbb{E}}\ } & \mathbb{TE} \\ {\scriptstyle \mathbf{p}}\downarrow & & \downarrow{\scriptstyle T\mathbf{p}} \\ \mathbb{M} & \xleftarrow{\ \boldsymbol{\tau}_{\mathbb{M}}\ } & \mathbb{TM} \end{array} \quad \Longleftrightarrow \quad \boldsymbol{\tau}_{\mathbb{M}} \circ T\mathbf{p} = \mathbf{p} \circ \boldsymbol{\tau}_{\mathbb{E}}\,.$$

we have that $\boldsymbol{\tau}_{\mathbb{M}} \circ T\mathbf{p} \circ \mathbf{X} = \boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} \circ \mathbf{p} = \mathbf{p} \circ \boldsymbol{\tau}_{\mathbb{E}} \circ \mathbf{X}$. Then $\boldsymbol{\tau}_{\mathbb{E}} \circ \mathbf{X} = \mathbf{id}_{\mathbb{E}}$ implies that $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} = \mathbf{id}_{\mathbb{M}}$ by the surjectivity of the projection $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}\,;\mathbb{M})$.

Corollary 1.2.2 (Flows of projectable vector fields) *The pair of tangent vector fields* $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ *and* $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ *is a bundle morphism from* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ *to* $T\mathbf{p} \in C^1(\mathbb{TE}\,;\mathbb{TM})$ *if and only if for any* $\lambda \in \Re$ *the flow* $\mathbf{Fl}^{\mathbf{X}}_{\lambda} \in C^1(\mathbb{E}\,;\mathbb{E})$ *is a bundle morphism which projects to the flow* $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in C^1(\mathbb{M}\,;\mathbb{M})$ *according to the commutative diagram:*

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{\ \mathbf{Fl}^{\mathbf{X}}_{\lambda}\ } & \mathbb{E} \\ {\scriptstyle \mathbf{p}}\downarrow & & \downarrow{\scriptstyle \mathbf{p}} \\ \mathbb{M} & \xrightarrow{\ \mathbf{Fl}^{\mathbf{v}}_{\lambda}\ } & \mathbb{M} \end{array} \quad \Longleftrightarrow \quad \mathbf{p} \circ \mathbf{Fl}^{\mathbf{X}}_{\lambda} = \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})\,.$$

Proof. This is a special case of Proposition 1.2.5 page 38. ■

A product bundle $\mathbf{p} \times \boldsymbol{\pi} \in C^1(\mathbb{E} \times \mathbb{H}\,; \mathbb{M} \times \mathbb{N})$ is the cartesian product of two given ones $\mathbf{p} \in C^1(\mathbb{E}\,; \mathbb{M})$ and $\boldsymbol{\pi} \in C^1(\mathbb{H}\,; \mathbb{N})$. If there is some relationship between the given bundles, other constructions may be performed. So, if the base spaces are identical, we get the special important construction of a fibred product bundle over the common base. On the other hand, if the total space of the fibred product is considered but choosing a different base space, we get the definition of pull-back bundle.

Definition 1.2.11 (Fibred product bundle) *Given two fibre bundles* $\mathbf{p} \in C^1(\mathbb{E}\,; \mathbb{M})$ *and* $\boldsymbol{\pi} \in C^1(\mathbb{H}\,; \mathbb{M})$ *over the same base* $\mathbb{M}$*, the fibred product bundle* $\mathbf{p} \times_{\mathbb{M}} \boldsymbol{\pi} \in C^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{H}\,; \mathbb{M})$ *is the bundle whose total space* $\mathbb{E} \times_{\mathbb{M}} \mathbb{H}$ *is the subset of the cartesion product:*

$$\{(\mathbf{e}\,,\mathbf{h}) \in \mathbb{E} \times \mathbb{H} \mid \mathbf{p}(\mathbf{e}) = \boldsymbol{\pi}(\mathbf{h})\}\,,$$

with the projection $\mathbf{p} \times_{\mathbb{M}} \boldsymbol{\pi}$ *given by*

$$(\mathbf{p} \times_{\mathbb{M}} \boldsymbol{\pi})(\mathbf{e}\,,\mathbf{h}) = \mathbf{p}(\mathbf{e}) = \boldsymbol{\pi}(\mathbf{h})\,.$$

The restrictions of the cartesian-product projections to the total space of a fibred product bundle, yield two more bundle structures [177].

Definition 1.2.12 *The maps* $\mathbf{p}{\downarrow}\boldsymbol{\pi} \in C^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{H}\,; \mathbb{E})$ *and* $\boldsymbol{\pi}{\downarrow}\mathbf{p} \in C^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{H}\,; \mathbb{H})$ *are fibre bundles defined by*

$$\mathbf{p}{\downarrow}\boldsymbol{\pi}(\mathbf{e}\,,\mathbf{h}) := \mathbf{e}\,, \qquad \boldsymbol{\pi}{\downarrow}\mathbf{p}(\mathbf{e}\,,\mathbf{h}) := \mathbf{h}\,.$$

The fibred product of two *vector bundles* $\mathbf{p} \in C^1(\mathbb{E}\,; \mathbb{M})$ and $\boldsymbol{\pi} \in C^1(\mathbb{H}\,; \mathbb{M})$ over the same base $\mathbb{M}$ is called the Whitney sum of the two bundles.

Definition 1.2.13 (Pull-back bundle) *Given a fibre bundle* $(\mathbb{E}\,, \mathbf{p}\,, \mathbb{M})$ *and a map* $\mathbf{f} \in C^1(\mathbb{H}\,; \mathbb{M})$ *the pull-back bundle* $(\mathbf{f}{\downarrow}\mathbb{E}\,, \mathbf{f}{\downarrow}\mathbf{p}\,, \mathbb{H})$ *by* $\mathbf{f}$ *is the fibre bundle whose total space* $\mathbf{f}{\downarrow}\mathbb{E}$ *is the subset of the cartesian product:*

$$\{(\mathbf{e}\,,\mathbf{h}) \in \mathbb{E} \times \mathbb{H} \mid \mathbf{p}(\mathbf{e}) = \mathbf{f}(\mathbf{h})\}\,,$$

with the projection $\mathbf{f}{\downarrow}\mathbf{p}$ *given by*

$$\mathbf{f}{\downarrow}\mathbf{p}(\mathbf{e}\,,\mathbf{h}) = \mathbf{h}\,.$$

Lemma 1.2.2 *The space of sections of the pull-back bundle* $\mathbf{f}{\downarrow}\mathbf{p} \in C^1(\mathbf{f}{\downarrow}\mathbb{E}\,; \mathbb{H})$ *is isomorphic to the space of sections of the fibre bundle* $\mathbf{p} \in C^1(\mathbb{E}\,; \mathbb{M})$ *along* $\mathbf{f} \in C^1(\mathbb{H}\,; \mathbb{M})$*.*

Proof. Let $\mathbf{s} \in \mathrm{C}^1(\mathbb{H};\mathbb{E})$ be a section of $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$ along $\mathbf{f} \in \mathrm{C}^1(\mathbb{H};\mathbb{M})$ so that $\mathbf{p} \circ \mathbf{s} = \mathbf{f}$. Then the pair $(\mathbf{s}, \mathbf{id}_{\mathbb{H}}) \in \mathrm{C}^1(\mathbb{H};\mathbf{f}{\downarrow}\mathbb{E})$ is a section of the pull-back bundle $\mathbf{f}{\downarrow}\mathbf{p} \in \mathrm{C}^1(\mathbf{f}{\downarrow}\mathbb{E};\mathbb{H})$ since $(\mathbf{f}{\downarrow}\mathbf{p} \circ (\mathbf{s}, \mathbf{id}_{\mathbb{H}}))\mathbf{h} = \mathbf{f}{\downarrow}\mathbf{p}((\mathbf{s}(\mathbf{h}), \mathbf{h})) = \mathbf{h}$. Vice versa, given a section $\mathbf{s_f} \in \mathrm{C}^1(\mathbb{H};\mathbf{f}{\downarrow}\mathbb{E})$ of the pull-back bundle with $\mathbf{s_f}(\mathbf{h}) = (\mathbf{e}, \mathbf{h})$, the map $\mathbf{s} \in \mathrm{C}^1(\mathbb{H};\mathbb{E})$ defined by $\mathbf{s}(\mathbf{h}) := \mathbf{e}$ fulfils the relation $(\mathbf{p} \circ \mathbf{s})(\mathbf{h}) = \mathbf{p}(\mathbf{e}) = \mathbf{f}(\mathbf{h})$. ■

The pair $(\mathbf{p}{\downarrow}\mathbf{f}, \mathbf{f})$ is a fibre bundle morphism from $\mathbf{f}{\downarrow}\mathbf{p}$ to $\mathbf{p}$:

$$\begin{array}{ccc} \mathbf{f}{\downarrow}\mathbb{E} & \xrightarrow{\mathbf{p}{\downarrow}\mathbf{f}} & \mathbb{E} \\ {\scriptstyle \mathbf{f}{\downarrow}\mathbf{p}}\Big\downarrow & & \Big\downarrow{\scriptstyle \mathbf{p}} \\ \mathbb{H} & \xrightarrow{\mathbf{f}} & \mathbb{M} \end{array} \quad \Longleftrightarrow \quad \mathbf{p} \circ \mathbf{p}{\downarrow}\mathbf{f} = \mathbf{f} \circ \mathbf{f}{\downarrow}\mathbf{p} \in \mathrm{C}^0(\mathbf{f}{\downarrow}\mathbb{E};\mathbb{M}).$$

The tipical fibres of $\mathbf{p}$ and $\mathbf{f}{\downarrow}\mathbf{p}$ are the same. The total space $\mathbf{f}{\downarrow}\mathbb{E}$ of the pull-back bundle $\mathbf{f}{\downarrow}\mathbf{p}$ may be thought of as formed by copies of the fibres of the fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$ with base points transplanted from $\mathbb{M}$ to $\mathbb{H}$ by the map $\mathbf{f} \in \mathrm{C}^1(\mathbb{H};\mathbb{M})$.

The notion of pull-back bundle permits to define the cotangent map of any morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$.

Definition 1.2.14 (Cotangent map of a morphism) *The cotangent map of a morphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$ *is the map* $T^*\boldsymbol{\varphi} \in \mathrm{C}^1(\boldsymbol{\varphi}{\downarrow}\mathbb{T}^*\mathbb{N};\mathbb{T}^*\mathbb{M})$ *defined by:*

$$T^*\boldsymbol{\varphi}(\mathbf{x}, \boldsymbol{\omega}) := T^*_{\mathbf{x}}\boldsymbol{\varphi} \cdot \boldsymbol{\omega}, \quad \forall\, \mathbf{x} \in \mathbb{M}, \quad \forall\, \boldsymbol{\omega} \in \mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N},$$

with $T^*_{\mathbf{x}}\boldsymbol{\varphi} \in BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N};\mathbb{T}^*_{\mathbf{x}}\mathbb{M})$ *bounded linear map dual to the tangent map* $T_{\mathbf{x}}\boldsymbol{\varphi} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M};\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N})$. *If the morphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$ *is invertible, we may set:*

$$T^*\boldsymbol{\varphi} \cdot \boldsymbol{\omega} := T^*\boldsymbol{\varphi}(\boldsymbol{\varphi}^{-1}(\boldsymbol{\tau}^*_{\mathbb{M}}(\boldsymbol{\omega})), \boldsymbol{\omega}),$$

thus recovering the special definition of the cotangent map $T^*\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{N};\mathbb{T}^*\mathbb{M})$ *of an invertible morphism.*

Let us define the pull-back by a morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$ of a one-form $\boldsymbol{\omega} \in \mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N}$ at the point $\mathbf{x} \in \mathbb{M}$ as $\boldsymbol{\varphi}{\downarrow}(\mathbf{x}, \boldsymbol{\omega}) := T^*\boldsymbol{\varphi}(\mathbf{x}, \boldsymbol{\omega}) \in \mathbb{T}^*_{\mathbf{x}}\mathbb{M}$. Then

$$\boldsymbol{\varphi}{\downarrow} = T^*\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M} \times_{\mathbb{N}} \mathbb{T}^*\mathbb{N};\mathbb{T}^*\mathbb{M}).$$

If the morphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$ is invertible, this definition reduces to the formula of Proposition 1.2.1.

In particular, if the morphism is the projection $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$, the pull-back bundle $\boldsymbol{\tau}^*_{\mathbb{M}}{\downarrow}\mathbb{T}^*\mathbb{M}$ is equal to the WHITNEY sum $\mathbb{T}^*\mathbb{M} \times_{\mathbb{M}} \mathbb{T}^*\mathbb{M}$ and the cotangent map $T^*\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\boldsymbol{\tau}^*_{\mathbb{M}}{\downarrow}\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})$ is defined by the relation:

$$T^*\boldsymbol{\tau}^*_{\mathbb{M}}(\mathbf{u}^*\,,\mathbf{v}^*) := T^*_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{v}^*\,, \quad \forall\, \mathbf{u}^*, \mathbf{v}^* \in \mathbb{T}^*_{\mathbf{x}}\mathbb{M}\,,$$

with $T^*_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \in BL(\mathbb{T}^*_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}^*_{\mathbf{u}^*}\mathbb{T}^*\mathbb{M})$ dual to $T_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \in BL(\mathbb{T}_{\mathbf{u}^*}\mathbb{T}^*\mathbb{M}\,;\mathbb{T}_{\mathbf{x}}\mathbb{M})$.

Lemma 1.2.3 (Cotangent map of the cotangent bundle projection) *The cotangent map* $T^*\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M} \times_{\mathbb{M}} \mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})$ *is a linear homomorphism between the bundles* $\mathbb{T}^*\mathbb{M} \times_{\mathbb{M}} \mathbb{T}^*\mathbb{M}$ *and* $(\mathbb{T}^*\mathbb{T}^*\mathbb{M}\,, \boldsymbol{\tau}^*_{\mathbb{TM}}\,, \mathbb{T}^*\mathbb{M})$ *over the identity in* $\mathbb{T}^*\mathbb{M}$ *which is fibrewise injective and horizontal valued.*

Proof. Fiberwise injectivity follows from the polarity relation

$$\mathbf{ker}(T^*_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}}) = (\mathbf{im}(T_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}}))^{\circ} = \{0\}\,,$$

a direct consequence of the assumption that the projection is a submersion. Horizontal valuedness means that the form $T^*_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{v}^* \in \mathbb{T}^*_{\mathbf{u}^*}\mathbb{T}^*\mathbb{M}$ vanishes on vertical vectors $\mathbf{X}_{\mathbf{u}^*} \in \mathbf{ker}(T_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}})$ and this follows from the duality relation:

$$\langle T^*_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{v}^*, \mathbf{X}_{\mathbf{u}^*} \rangle = \langle \mathbf{v}^*, T_{\mathbf{u}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{X}_{\mathbf{u}^*} \rangle\,.$$

Linearity in $\mathbf{v}^* \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ for any fixed $\mathbf{u}^* \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ is clear. ∎

Definition 1.2.15 (Liouville one-form) *The canonical or* LIOUVILLE *one-form* $\boldsymbol{\theta}\mathbb{M} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})$ *is the horizontal valued form defined by*

$$\boldsymbol{\theta}\mathbb{M} := T^*\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathrm{DIAG}\,,$$

with the diagonal map $\mathrm{DIAG} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{M} \times_{\mathbb{M}} \mathbb{T}^*\mathbb{M})$ *given by*

$$\mathrm{DIAG}(\mathbf{v}^*) = (\mathbf{v}^*\,,\mathbf{v}^*)\,, \quad \forall\, \mathbf{v}^* \in \mathbb{T}^*\mathbb{M}\,.$$

Then $\boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*) := T^*_{\mathbf{v}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{v}^* \in \mathbb{T}^*_{\mathbf{v}^*}\mathbb{T}^*\mathbb{M}$ *and* $\boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*) = 0 \iff \mathbf{v}^* = 0\,.$

In a similar way, if the morphism is the projection $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$, the pull-back bundle $\boldsymbol{\tau}_{\mathbb{M}}{\downarrow}\mathbb{T}^*\mathbb{M}$ is equal to the WHITNEY sum $\mathbb{TM} \times_{\mathbb{M}} \mathbb{T}^*\mathbb{M}$ and the cotangent map $T^*\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\boldsymbol{\tau}_{\mathbb{M}}{\downarrow}\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{TM})$ is defined by the relation:

$$T^*\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{u}\,,\mathbf{v}^*) := T^*_{\mathbf{u}}\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{v}^*\,, \quad \forall\, \mathbf{u} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,, \quad \forall\, \mathbf{v}^* \in \mathbb{T}^*_{\mathbf{x}}\mathbb{M}\,,$$

with $T^*_{\mathbf{u}}\boldsymbol{\tau}_{\mathbb{M}} \in BL(\mathbb{T}^*_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}^*_{\mathbf{u}}\mathbb{T}\mathbb{M})$ dual to $T_{\mathbf{u}}\boldsymbol{\tau}_{\mathbb{M}} \in BL(\mathbb{T}_{\mathbf{u}}\mathbb{T}\mathbb{M}\,;\mathbb{T}_{\mathbf{x}}\mathbb{M})$. Given a bundle morphism $\mathbf{A} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ we may define the map

$$T^*\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{u}\,,\mathbf{A}(\mathbf{v})) := T^*_{\mathbf{u}}\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{A}(\mathbf{v})\,, \quad \forall\,\mathbf{u},\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,, \quad \mathbf{A}(\mathbf{v}) \in \mathbb{T}^*_{\mathbf{x}}\mathbb{M}\,,$$

Definition 1.2.16 (Poincaré-Cartan one-form) *The* Poincaré-Cartan *one-form* $\boldsymbol{\theta}\mathbf{A} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{T}^*\mathbb{T}\mathbb{M})$ *is the horizontal valued form defined by*

$$\boldsymbol{\theta}\mathbf{A} := T^*\boldsymbol{\tau}_{\mathbb{M}} \circ (\mathbf{id}_{\,\mathbb{T}\mathbb{M}}\,,\mathbf{A})\,,$$

Then $\boldsymbol{\theta}\mathbf{A}(\mathbf{v}) := T^*_{\mathbf{v}}\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{A}(\mathbf{v}) \in \mathbb{T}^*_{\mathbf{v}}\mathbb{T}\mathbb{M}$ *and* $\boldsymbol{\theta}\mathbf{A}(\mathbf{v}) = 0 \iff \mathbf{A}(\mathbf{v}) = 0$ *.*

Lemma 1.2.4 *The* Poincaré-Cartan *one-form* $\boldsymbol{\theta}\mathbf{A} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{T}^*\mathbb{T}\mathbb{M})$ *is the pull-back of the* Liouville *one-form by means of the bundle morphism* $\mathbf{A} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ *, i.e.*

$$\boldsymbol{\theta}\mathbf{A}(\mathbf{v}) := \mathbf{A}{\downarrow}(\mathbf{v}\,,\boldsymbol{\theta}\mathbb{M}) \in \mathbb{T}^*_{\mathbf{v}}\mathbb{T}\mathbb{M}\,.$$

Proof. By definition $\mathbf{A}{\downarrow}(\mathbf{v}\,,\boldsymbol{\theta}\mathbb{M}) = T^*\mathbf{A}(\mathbf{v}\,,\boldsymbol{\theta}\mathbb{M}) = T^*_{\mathbf{v}}\mathbf{A} \cdot \boldsymbol{\theta}\mathbb{M}(\mathbf{A}(\mathbf{v}))$. Being $\boldsymbol{\theta}\mathbb{M}(\mathbf{A}(\mathbf{v})) = T^*\boldsymbol{\tau}^*_{\mathbb{M}}(\mathbf{A}(\mathbf{v})\,,\mathbf{A}(\mathbf{v}))$, we infer that

$$\begin{aligned}
\mathbf{A}{\downarrow}(\mathbf{v}\,,\boldsymbol{\theta}\mathbb{M}) = T^*_{\mathbf{v}}\mathbf{A} \cdot \boldsymbol{\theta}\mathbb{M}(\mathbf{A}(\mathbf{v})) &= T^*_{\mathbf{v}}\mathbf{A} \cdot T^*\boldsymbol{\tau}^*_{\mathbb{M}}(\mathbf{A}(\mathbf{v})\,,\mathbf{A}(\mathbf{v})) \\
&= T^*_{\mathbf{v}}\mathbf{A} \cdot T^*_{\mathbf{A}(\mathbf{v})}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{A}(\mathbf{v}) \\
&= T^*_{\mathbf{v}}(\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathbf{A}) \cdot \mathbf{A}(\mathbf{v}) \\
&= T^*_{\mathbf{v}}\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{A}(\mathbf{v}) \\
&= (T^*\boldsymbol{\tau}_{\mathbb{M}} \circ (\mathbf{id}_{\,\mathbb{T}\mathbb{M}}\,,\mathbf{A}))(\mathbf{v}) = \boldsymbol{\theta}\mathbf{A}(\mathbf{v})\,,
\end{aligned}$$

and the result is proved. ■

1.2.9 Linear operations in vector bundles

In a vector bundle $(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})$ the fibrewise addition $+_{\mathbf{p}}$ defines a bilinear homomorphism $\mathbf{add}_{\,(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})} \in \mathrm{C}^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{E}\,;\mathbb{E})$ over the identity $\mathbf{id}_{\,\mathbb{M}}$ by

$$\begin{array}{ccc}
\mathbb{E} \times_{\mathbb{M}} \mathbb{E} & \xrightarrow{\mathbf{add}_{\,(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}} & \mathbb{E} \\
{\scriptstyle (\mathbf{p}\,,\mathbf{p})}\big\downarrow & & \big\downarrow{\scriptstyle \mathbf{p}} \\
\mathbb{M} & \xrightarrow{\mathbf{id}_{\,\mathbb{M}}} & \mathbb{M}
\end{array}
\iff
\begin{cases}
\mathbf{add}_{\,(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})} \cdot (\mathbf{u}\,,\mathbf{v})_{\mathbf{x}} := \mathbf{u}_{\mathbf{x}} +_{\mathbf{p}} \mathbf{v}_{\mathbf{x}}\,, \\
\mathbf{u}_{\mathbf{x}}, \mathbf{v}_{\mathbf{x}} \in \mathbb{E}_{\mathbf{x}}\,.
\end{cases}$$

The fibrewise addition on the vector bundle $(\mathbb{TE}\,, T\mathbf{p}\,, \mathbb{TM})$, denoted by $+_{T\mathbf{p}}$, is induced by acting with the tangent functor.

The bilinear homomorphism $\mathbf{add}_{(\mathbb{TE}\,,T\mathbf{p}\,,\mathbb{TM})} \in \mathrm{C}^1(\mathbb{TE} \times_{\mathbb{TM}} \mathbb{TE}\,; \mathbb{TE})$ over the identity $\mathbf{id}_{\mathbb{TM}}$ is thus defined by $\mathbf{add}_{(\mathbb{TE}\,,T\mathbf{p}\,,\mathbb{TM})} := T\mathbf{add}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}$ and explicitly

$$\mathbf{add}_{(\mathbb{TE}\,,T\mathbf{p}\,,\mathbb{TM})} \cdot (\mathbf{X}\,, \mathbf{Y})_{\mathbf{w}_\mathbf{x}} = T\mathbf{add}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})\,, \boldsymbol{\tau}_{\mathbb{E}}(\mathbf{Y}))_{\mathbf{x}} \cdot (\mathbf{X}\,, \mathbf{Y})_{\mathbf{w}_\mathbf{x}}\,,$$

so that

$$\begin{array}{ccc} \mathbb{TE} \times_{\mathbb{TM}} \mathbb{TE} & \xrightarrow{\mathbf{add}_{(\mathbb{TE}\,,T\mathbf{p}\,,\mathbb{TM})}} & \mathbb{TE} \\ {\scriptstyle (T\mathbf{p}\,,T\mathbf{p})}\downarrow & & \downarrow{\scriptstyle T\mathbf{p}} \\ \mathbb{TM} & \xrightarrow{\mathbf{id}_{\mathbb{TM}}} & \mathbb{TM} \end{array} \iff \begin{cases} \mathbf{add}_{(\mathbb{TE}\,,T\mathbf{p}\,,\mathbb{TM})} \cdot (\mathbf{X}\,, \mathbf{Y})_{\mathbf{w}_\mathbf{x}} := \\ \mathbf{X}_{\mathbf{w}_\mathbf{x}} +_{T\mathbf{p}} \mathbf{Y}_{\mathbf{w}_\mathbf{x}}\,, \end{cases}$$

where the pair $(\mathbf{X}\,, \mathbf{Y})_{\mathbf{w}_\mathbf{x}}$ is such that $T\mathbf{p} \cdot \mathbf{X} = T\mathbf{p} \cdot \mathbf{Y} = \mathbf{w}_\mathbf{x} \in \mathbb{T}_\mathbf{x}\mathbb{M}$ and $\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X}) = \mathbf{u}_\mathbf{x}$, $\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{Y}) = \mathbf{v}_\mathbf{x}$ so that $\mathbf{p}(\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})) = \mathbf{p}(\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{Y})) = \mathbf{x} \in \mathbb{M}$.

Analogously for the scalar multiplications $\mathbf{mult}^{\alpha}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}) \in BL(\mathbb{E}_{\mathbf{p}(\mathbf{e})}\,; \mathbb{E}_{\mathbf{p}(\mathbf{e})})$ and $T\mathbf{mult}^{\alpha}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}) \cdot \mathbf{X} \in BL(\mathbb{T}_\mathbf{e}\mathbb{E}\,; \mathbb{T}_\mathbf{e}\mathbb{E})$.

1.2.10 Exact sequences

Let $\mathbb{Q}, \mathbb{M}, \mathbb{N}$ be manifolds and $\mathbf{f} \in \mathrm{C}^1(\mathbb{Q}\,; \mathbb{M})$ $\mathbf{g} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{N})$ be manifolds morphisms. A sequence:

$$\mathbb{Q} \xrightarrow{\ \mathbf{f}\ } \mathbb{M} \xrightarrow{\ \mathbf{g}\ } \mathbb{N}$$

is called exact if $\mathbf{im}(\mathbf{f}) = \mathbf{ker}(\mathbf{g})$. Let us now consider a sequence of vector bundles $(\mathbb{E}_1\,, \mathbf{p}_1\,, \mathbb{M})$, $(\mathbb{E}_2\,, \mathbf{p}_2\,, \mathbb{M})$, $(\mathbb{E}_3\,, \mathbf{p}_3\,, \mathbb{M})$ over the same base manifold $\mathbb{M}$. Denoting by 0 the null vector bundle, the exact sequence

$$0 \longrightarrow \mathbb{E}_1 \xrightarrow{\ \mathbf{f}\ } \mathbb{E}_2 \xrightarrow{\ \mathbf{g}\ } \mathbb{E}_3 \longrightarrow 0$$

implies that $\mathbf{f} \in \mathrm{C}^1(\mathbb{E}_1\,; \mathbb{E}_2)$ is injective and $\mathbf{g} \in \mathrm{C}^1(\mathbb{E}_2\,; \mathbb{E}_3)$ is surjective.

Definition 1.2.17 (Splitting) *The exact sequence above is said to admit a splitting if there exists an injective vector bundle morphism* $\mathbf{h} \in \mathrm{C}^1(\mathbb{E}_3\,; \mathbb{E}_2)$ *such that* $\mathbf{g} \circ \mathbf{h} = \mathbf{id}_{\mathbb{E}_3}$. *Then* $\mathbb{E}_2 = \boldsymbol{im}(\mathbf{f}) \oplus \boldsymbol{im}(\mathbf{h})$.

1.2.11 Second tangent bundle

Higher order tangent bundles play an important role in the geometric description of many fundamental issues in physics. The *second tangent bundle* is of special importance in dynamics on manifolds.

Let us consider the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ of a manifold $\mathbb{M}$ and its second tangent bundle $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$.

The tangent map $T\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^0(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ of the projection $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ provides another vector bundle structure on the base manifold $\mathbb{TM}$ with the commutative diagram:

$$\begin{array}{ccc} \mathbb{T}^2\mathbb{M} & \xrightarrow{T\boldsymbol{\tau}_{\mathbb{M}}} & \mathbb{TM} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{TM}}}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\tau}_{\mathbb{M}}} \\ \mathbb{TM} & \xrightarrow{\boldsymbol{\tau}_{\mathbb{M}}} & \mathbb{M} \end{array} \iff \quad \boldsymbol{\tau}_{\mathbb{M}} \circ T\boldsymbol{\tau}_{\mathbb{M}} = \boldsymbol{\tau}_{\mathbb{M}} \circ \boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^0(\mathbb{T}^2\mathbb{M}\,;\mathbb{M})\,.$$

The relation between the two bundle structures is conveniently described in terms of the canonical involution, as described in [80] and in the next paragraph.

1.2.12 Canonical involution

Definition 1.2.18 (Flip) *The canonical involution* $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{T}^2\mathbb{M})$ *is defined by*

$$\mathbf{k}_{\mathbb{T}^2\mathbb{M}}\,(\partial_{\mu=0}\,\partial_{\lambda=0}\,\mathbf{c}(\lambda,\mu)) := \partial_{\lambda=0}\,\partial_{\mu=0}\,\mathbf{c}(\lambda,\mu)\,, \quad \forall\,\mathbf{c} \in \mathrm{C}^2(\Re \times \Re\,;\mathbb{M})\,.$$

In a local chart $(U\,,\varphi)$, setting $\varphi \circ \mathbf{c} = c$, we have that

$$\begin{aligned} &(T^2\varphi \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T^2\varphi)\,(c(0,0)\,,\partial_{\lambda=0}\,c(\lambda,0)\,,\partial_{\mu=0}\,c(0,\mu)\,,\partial_{\mu=0}\,\partial_{\lambda=0}\,c(\lambda,\mu)) \\ &\qquad := (c(0,0)\,,\partial_{\mu=0}\,c(0,\mu)\,,\partial_{\lambda=0}\,c(\lambda,0)\,,\partial_{\lambda=0}\,\partial_{\mu=0}\,c(\lambda,\mu))\,. \end{aligned}$$

or in terms of components:

$$(T^2\varphi \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T^2\varphi)(x\,,u\,,v\,,\xi) := (x\,,v\,,u\,,\xi)\,.$$

The *flip* nickname underlines that the map performs an exchange in the order of the iterated derivation.

We denote by $\pm_{\boldsymbol{\tau}_{\mathbb{TM}}}$ and $\pm_{T\boldsymbol{\tau}_{\mathbb{M}}} := T\pm_{\boldsymbol{\tau}_{\mathbb{M}}}$ respectively the fibrewise addition (subtraction) in the vector bundles $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ and $T\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^0(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$. Often $\pm_{\boldsymbol{\tau}_{\mathbb{TM}}}$ is simply denoted by $\pm$. Likewise $\cdot_{\boldsymbol{\tau}_{\mathbb{TM}}}$ and $\cdot_{T\boldsymbol{\tau}_{\mathbb{M}}}$ are the fibrewise multiplications, with $\cdot$ denoting $\cdot_{\boldsymbol{\tau}_{\mathbb{TM}}}$ by default.

Lemma 1.2.5 *The flip is involutive:* $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} = \mathbf{id}_{\mathbb{T}^2\mathbb{M}}$ *, such that:*

$$\boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} = T\boldsymbol{\tau}_{\mathbb{M}} \,, \qquad T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} = \boldsymbol{\tau}_{\mathbb{TM}} \,,$$

and provides a linear isomorphism between the bundles $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ *and* $T\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ *defined by*

$$\begin{cases} \mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X} +_{\boldsymbol{\tau}_{\mathbb{TM}}} \mathbf{Y}) = \mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}) +_{T\boldsymbol{\tau}_{\mathbb{M}}} \mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{Y}) \,, & \boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}) = \boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{Y}) \,, \\ \mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\alpha \cdot_{\boldsymbol{\tau}_{\mathbb{TM}}} \mathbf{X}) = \alpha \cdot_{T\boldsymbol{\tau}_{\mathbb{M}}} \mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}) \,, & \alpha \in \Re \,. \end{cases}$$

Moreover, for any $\mathbf{f} \in \mathrm{C}^2(\mathbb{M}\,;\mathbb{N})$ *:*

$$\mathbf{k}_{\mathbb{T}^2\mathbb{N}} \circ T^2\mathbf{f} = T^2\mathbf{f} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \,,$$

where $T^2\mathbf{f} \in \mathrm{C}^0(\mathbb{T}^2\mathbb{M}\,;\mathbb{T}^2\mathbb{N})$ *.*

Proof. Involutivity is clear from the definition. The base vector of the vector $\mathbf{X}(\mathbf{v}) := \partial_{\mu=0}\,\partial_{\lambda=0}\,\mathbf{c}(\lambda,\mu) \in \mathbb{T}_{\mathbf{v}}\mathbb{TM}$ in the bundle $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ is $\mathbf{v} = \boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}(\mathbf{v})) = \partial_{\lambda=0}\,\mathbf{c}(\lambda,0) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}$ and the base-point velocity is $T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}(\mathbf{v}) = \partial_{\mu=0}\,\mathbf{c}(0,\mu) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}$ where $\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v}) = \mathbf{c}(0,0) \in \mathbb{M}$.

The flip involution transforms the vector $\mathbf{X}(\mathbf{v})$ into the vector $\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}(\mathbf{v})) = \partial_{\lambda=0}\,\partial_{\mu=0}\,\mathbf{c}(\lambda,\mu) \in \mathbb{T}_{T\boldsymbol{\tau}_{\mathbb{M}}\cdot\mathbf{X}(\mathbf{v})}\mathbb{TM}$ whose base vector is $\boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}(\mathbf{v}))) = \partial_{\mu=0}\,\mathbf{c}(0,\mu) = T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}(\mathbf{v}) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}$ and whose base-point velocity is $\mathbf{v} = \boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}(\mathbf{v})) = \partial_{\lambda=0}\,\mathbf{c}(\lambda,0) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}$. The flip is a bundle morphism between the bundles $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ and $T\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^0(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ since two vectors with the same base point in $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ are transformed in vectors with the same base velocity and hence with the same base point in $T\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^0(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ and vice versa. Fiberwise linearity of the flip follows from the rules:

$$\begin{aligned} (x\,,u\,,v\,,\xi) +_{\boldsymbol{\tau}_{\mathbb{TM}}} (x\,,u\,,w\,,\zeta) &= (x\,,u\,,v+w\,,\xi+\zeta) \\ (x\,,v\,,u\,,\xi) +_{T\boldsymbol{\tau}_{\mathbb{M}}} (x\,,w\,,u\,,\zeta) &= (x\,,v+w\,,u\,,\xi+\zeta)\,, \end{aligned}$$

and

$$\begin{aligned} \alpha \cdot_{\boldsymbol{\tau}_{\mathbb{TM}}} (x\,,u\,,v\,,\xi) &= (x\,,u\,,\alpha v\,,\alpha\xi) \\ \alpha \cdot_{T\boldsymbol{\tau}_{\mathbb{M}}} (x\,,v\,,u\,,\xi) &= (x\,,\alpha u\,,v\,,\alpha\xi)\,, \end{aligned}$$

For any $\mathbf{f} \in \mathrm{C}^2(\mathbb{M}\,;\mathbb{N})$ and $\mathbf{X} \in \mathbb{T}_{\mathbf{u}}\mathbb{TM}$ by definition we have: $T^2\mathbf{f}(\mathbf{u}) \cdot \mathbf{X} = \partial_{\mu=0}\,\partial_{\lambda=0}\,(\mathbf{f} \circ \mathbf{c})(\lambda,\mu)$, with $\mathbf{u}_\mu = \partial_{\lambda=0}\,\mathbf{c}(\lambda,\mu)$, $\mathbf{u} = \mathbf{u}_0$ and $\mathbf{X} = \partial_{\mu=0}\,\mathbf{u}_\mu$.

Then:

$$(T^2\mathbf{f}(\mathbf{u}) \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}}) \cdot \mathbf{X} = \partial_{\lambda=0}\, \partial_{\mu=0}\, (\mathbf{f} \circ \mathbf{c})(\lambda, \mu)$$
$$= \mathbf{k}_{\mathbb{T}^2\mathbb{N}}(\partial_{\mu=0}\, \partial_{\lambda=0}\, (\mathbf{f} \circ \mathbf{c})(\lambda, \mu)) = (\mathbf{k}_{\mathbb{T}^2\mathbb{N}} \circ T^2\mathbf{f}(\mathbf{u})) \cdot \mathbf{X}\,,$$

and the second assertion follows. ■

Lemma 1.2.6 *Acting with the tangent functor on a section* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *of the tangent bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ *, we get a section* $T\mathbf{v} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *of the bundle* $T\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ *. The map* $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *is a section of the bundle* $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ *.*

Proof. Since $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ is a section, we have that $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} = \mathbf{id}_{\mathbb{M}}$. Then: $T\boldsymbol{\tau}_{\mathbb{M}} \circ T\mathbf{v} = T(\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v}) = T\mathbf{id}_{\mathbb{M}} = \mathbf{id}_{\mathbb{TM}}$ and the second statement follows since $\boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} = T\boldsymbol{\tau}_{\mathbb{M}} \circ T\mathbf{v}$. ■

Note that for any section $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ the diagram

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\mathbf{v}} & \mathbb{T}^2\mathbb{M} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{M}}}\big\downarrow & & \big\downarrow{\scriptstyle T\boldsymbol{\tau}_{\mathbb{M}}} \\ \mathbb{M} & \xrightarrow{\mathbf{v}} & \mathbb{TM} \end{array}$$

is **not** commutative. Indeed the map $\mathbf{v} \circ \boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ is fibrewise constant in $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ and hence cannot be equal to the identity $\mathbf{id}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$. The next result provides the relation between the vector field associated with a flow and the one associated with its tangent map.

Lemma 1.2.7 (Velocity of the tangent flow) *Let* $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *be the flow of the vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *and* $T\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ *the relevant tangent map. Then the following formula holds*

$$\mathbf{Fl}^{\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v}}_{\lambda} = T\mathbf{Fl}^{\mathbf{v}}_{\lambda}\,,$$

where $T\mathbf{v} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *is the map tangent to* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *and* $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{T}^2\mathbb{M})$ *is the canonical flip.*

Proof. Let $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ be vector fields and $\mathbf{Fl}^{\mathbf{u}}_{\mu}, \mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ their flows. Then the velocity of the curve:

$$T\mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{u} = \partial_{\mu=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu}\,,$$

is given by

$$\begin{aligned} \partial_{\lambda=0}\; T\mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{u} = \partial_{\lambda=0}\, \partial_{\mu=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} &= \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \partial_{\mu=0}\, \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \\ &= \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \partial_{\mu=0}\, (\mathbf{v} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu}) = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{u}\,. \end{aligned}$$

The arbitrarity of $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ implies that $\partial_{\lambda=0}\; T\mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v}\,.$ ■

The following result translates the classical result by Euler and H.A. Schwarz on the symmetry of the iterated derivative, into a property of the second tangent map of a vector valued functional on a manifold $\mathbb{M}$.

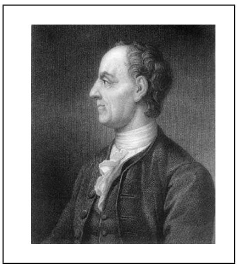

Figure 1.12: Leonhard Euler (1707 - 1783)

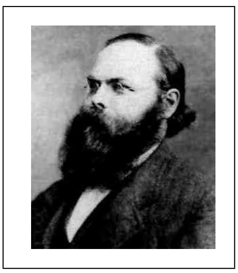

Figure 1.13: Hermann Amandus Schwarz (1843 - 1921)

Lemma 1.2.8 (Euler-Schwarz) *For any functional* $f \in C^2(\mathbb{M};\Re)$ *and any pair of vector fields* $\mathbf{u}, \mathbf{v} \in C^1(\mathbb{M};\mathbb{TM})$ *we have that*

$$T^2 f \circ T\mathbf{v} \circ \mathbf{u} = \mathbf{k}_{\mathbb{TT}\Re} \circ T^2 f \circ T\mathbf{v} \circ \mathbf{u}\,.$$

Proof. From Lemma 1.2.5 we have: $T^2 f \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} = \mathbf{k}_{\mathbb{TT}\Re} \circ T^2 f$. Then, setting $f_{\mathbf{vu}}(\lambda, \mu) := f \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu}$, we get

$$\begin{aligned} \partial_{\mu=0}\, \partial_{\lambda=0}\, f_{\mathbf{vu}}(\lambda,\mu) &:= \partial_{\mu=0}\, \partial_{\lambda=0}\, f \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \\ &= \partial_{\mu=0}\, Tf \circ \mathbf{v} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \\ &= T^2 f \circ T\mathbf{v} \circ \mathbf{u}\,, \end{aligned}$$

and

$$\begin{aligned} \partial_{\lambda=0}\, \partial_{\mu=0}\, f_{\mathbf{vu}}(\lambda,\mu) &:= \partial_{\lambda=0}\, \partial_{\mu=0}\, f \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \\ &= \partial_{\lambda=0}\, Tf \circ T\mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{u} \\ &= T^2 f \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{u} \\ &= \mathbf{k}_{\mathbb{TT}\Re} \circ T^2 f \circ T\mathbf{v} \circ \mathbf{u}\,, \end{aligned}$$

and then the equivalence of the statement with the standard form of SCHWARZ theorem is apparent. ■

1.2.13 Sprays

Let $\mathbf{X} \in C^1(\mathbb{TM};\mathbb{T}^2\mathbb{M})$ be a section of the tangent bundle $\boldsymbol{\tau}_{\mathbb{TM}} \in C^1(\mathbb{T}^2\mathbb{M};\mathbb{TM})$, so that $\boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{X} = \mathbf{id}_{\mathbb{TM}}$, i.e. $\mathbf{X}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{TM}$. The associated flow $\mathbf{Fl}^{\mathbf{X}}_{\lambda} \in C^1(\mathbb{TM};\mathbb{TM})$ is defined by the differential equation

$$\partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{X}}_{\lambda} = \mathbf{X}\,, \qquad \mathbf{Fl}^{\mathbf{X}}_{0} = \mathbf{id}_{\mathbb{TM}}\,.$$

We give the following definition.

Definition 1.2.19 (Spray) *A section* $\mathbf{X} \in C^1(\mathbb{TM};\mathbb{T}^2\mathbb{M})$ *of the tangent bundle* $\boldsymbol{\tau}_{\mathbb{TM}} \in C^1(\mathbb{T}^2\mathbb{M};\mathbb{TM})$ *is called a spray if it is also a section of the bundle* $T\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{T}^2\mathbb{M};\mathbb{TM})$ *, that is if* $T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{X} = \boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{X} = \mathbf{id}_{\mathbb{TM}}$ *.*

Lemma 1.2.9 *A section* $\mathbf{X} \in C^1(\mathbb{TM};\mathbb{T}^2\mathbb{M})$ *of the tangent bundle* $\boldsymbol{\tau}_{\mathbb{TM}} \in C^1(\mathbb{T}^2\mathbb{M};\mathbb{TM})$ *is a spray if and only if* $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{X} = \mathbf{X}$ *.*

Proof. The *if* part follows from Lemma 1.2.5 since

$$\mathbf{X} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{X} \implies T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{X} = T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{X} = \boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{X} = \mathbf{id}_{\mathbb{TM}} \,.$$

The *only if* part amounts to prove that

$$T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{X} = \boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{X} = \mathbf{id}_{\mathbb{TM}} \implies \mathbf{X} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{X} \,.$$

Let $\mathbf{c} \in \mathrm{C}^2(\Re \times \Re\,; \mathbb{M})$ and set $\mathbf{c}(\lambda, \mu) = \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu}$ with $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TM})$ vector fields, and $\mathbf{X} = \partial_{\mu=0}\, \partial_{\lambda=0}\, \mathbf{c}(\lambda, \mu)$. Then

$$\begin{aligned}
\mathbf{X} &= \partial_{\mu=0}\, \partial_{\lambda=0}\, \mathbf{c}(\lambda, \mu) = \partial_{\mu=0}\, \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \\
&= \partial_{\mu=0}\, \mathbf{v} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} = T\mathbf{v} \circ \mathbf{u} \,, \\
\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{X} &= \partial_{\lambda=0}\, \partial_{\mu=0}\, \mathbf{c}(\lambda, \mu) = \partial_{\lambda=0}\, \partial_{\mu=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \\
&= \partial_{\lambda=0}\, T\mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{u} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{u} \,,
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{X} &= \partial_{\lambda=0}\, \mathbf{c}(\lambda, 0) = \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{0} = \mathbf{v} \,, \\
T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{X} &= \partial_{\mu=0}\, \mathbf{c}(0, \mu) = \partial_{\mu=0}\, \mathbf{Fl}^{\mathbf{v}}_{0} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} = \mathbf{u} \,.
\end{aligned}$$

Note that, assuming commutation of the flows, i.e. $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} = \mathbf{Fl}^{\mathbf{u}}_{\mu} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}$, we have $\mathbf{X} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{u} \circ \mathbf{v}$. Now the assumption $T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{X} = \boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{X}$ implies that $\mathbf{u} = \mathbf{v}$ and hence that

$$\mathbf{X} = T\mathbf{u} \circ \mathbf{u} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{u} \circ \mathbf{u} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{X} \,,$$

which was to be proved. ■

The tangent map $T\mathbf{v} \in \mathrm{C}^0(\mathbb{TM}\,; \mathbb{T}^2\mathbb{M})$ of any map $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TM})$ cannot be a spray. Indeed, by remark 1.2.1 on page 29, we know that $T\mathbf{v} \in \mathrm{C}^0(\mathbb{TM}\,; \mathbb{T}^2\mathbb{M})$ is not a section of the bundle $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,; \mathbb{TM})$. On the contrary, the compostion $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \in \mathrm{C}^0(\mathbb{TM}\,; \mathbb{T}^2\mathbb{M})$ is a section of the bundle $\boldsymbol{\tau}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,; \mathbb{TM})$, as shown by lemma 1.2.6. By definition a spray is not a projectable vector field.

1.2.14 Second order vectors

Definition 1.2.20 (Second order vectors) *A vector $\mathbf{X} \in \mathbb{T}_{\mathbf{v}}\mathbb{TM}$ is said to be second order if the velocity of the base point $\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v}) \in \mathbb{M}$ is equal to the the tangent vector $\mathbf{v} \in \mathbb{TM}$, i.e. if $T_{\mathbf{v}}\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}(\mathbf{v}) = \boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}(\mathbf{v})) = \mathbf{v}$.*

The motivation is the following. A second order vector $\mathbf{X} \in \mathbb{T}_{\mathbf{v}}\mathbb{TM}$ is the velocity $\mathbf{X}(\mathbf{v}) = \partial_{\lambda=0}\,\mathbf{c}(\lambda)$ of a curve $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{TM})$ at $\mathbf{c}(0) = \mathbf{v} \in \mathbb{TM}$, such that the projected curve $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{M})$ has velocity given by

$$\partial_{\lambda=0}\,(\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{c}) = T\boldsymbol{\tau}_{\mathbb{M}} \circ \partial_{\lambda=0}\,\mathbf{c} = T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{X}(\mathbf{v}) = \mathbf{v}\,.$$

In a local chart the components of $\mathbf{v}(\mathbf{x}) \in \mathbb{TM}$ are $(x\,,v)$ and the components of $\mathbf{X}(\mathbf{v})$ are $((\mathrm{pr}_1 \circ X)(x\,,v)\,,(\mathrm{pr}_2 \circ X)(x\,,v))$. The defining property ensures that

$$(\mathrm{pr}_1 \circ X)(x\,,v) = v\,.$$

Then the system of first order differential equations

$$\begin{cases} \dot{x} = (\mathrm{pr}_1 \circ X)(x\,,v)\,,\\ \dot{v} = (\mathrm{pr}_2 \circ X)(x\,,v)\,, \end{cases}$$

is equivalent to the second order differential equation

$$\ddot{x} = (\mathrm{pr}_2 \circ X)(x\,,\dot{x})\,.$$

1.2.15 Vertical bundle

With any fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ there are associated two vector bundle structures: $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{E})$ and $T\mathbf{p} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TM})$.

The former is the vector bundle tangent to the manifold $\mathbb{E}$ while the latter is the result of acting on the fibre bundle with the tangent functor.

These bundle structures are related to the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ by the commutative diagram:

$$\begin{array}{ccc} \mathbb{TE} & \xrightarrow{T\mathbf{p}} & \mathbb{TM} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{E}}}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\tau}_{\mathbb{M}}} \\ \mathbb{E} & \xrightarrow{\mathbf{p}} & \mathbb{M} \end{array} \quad \Longleftrightarrow \quad \mathbf{p} \circ \boldsymbol{\tau}_{\mathbb{E}} = \boldsymbol{\tau}_{\mathbb{M}} \circ T\mathbf{p} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{M})\,.$$

Definition 1.2.21 *The vertical bundle is the subbundle* $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{VE}\,;\mathbb{E})$ *whose fibres are the point kernels of the tangent map* $T\mathbf{p} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TM})$*, that is*

$$\mathbb{V}_{\mathbf{b}}\mathbb{E} := \boldsymbol{ker}(T_{\mathbf{b}}\mathbf{p})\,, \quad \mathbf{b} \in \mathbb{E}\,.$$

Vertical vectors fields $\mathbf{V} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{VE})$ *on a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *are characterized by being projectable to the zero section of* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$*.*

In each tangent space $\mathbb{T}_{\mathbf{b}}\mathbb{E}$ the subspace of vertical vectors is made of tangents at $\mathbf{b} \in \mathbb{E}$ to the curves $\mathbf{c}_{\mathbb{E}} \in \mathrm{C}^1(I\,;\mathbb{E})$ such that the velocity of the projected curve $\mathbf{p} \circ \mathbf{c}_{\mathbb{E}} \in \mathrm{C}^1(I\,;\mathbb{M})$ vanishes at $\mathbf{p}(\mathbf{b})$.

It is clear that these tangents vectors belong to the tangent space $\mathbb{T}_{\mathbf{b}}\mathbb{E}_{\mathbf{x}}$ to the fibre $\mathbb{E}_{\mathbf{x}}$ over $\mathbf{x} = \mathbf{p}(\mathbf{b}) \in \mathbb{M}$. Then $\dim \mathbb{V}_{\mathbf{b}}\mathbb{E} = \dim \mathbb{T}_{\mathbf{b}}\mathbb{E}_{\mathbf{x}}$.

Given a section $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ of the fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, and a vector $\mathbf{X} \in \mathbb{TE}$, the difference: $\mathbf{X} - (T\mathbf{s} \circ T\mathbf{p}) \cdot \mathbf{X}$ is a vertical vector. Indeed

$$T\mathbf{p} \cdot (T\mathbf{s} \circ T\mathbf{p}) \cdot \mathbf{X} = (T\mathbf{p} \circ T\mathbf{s} \circ T\mathbf{p}) \cdot \mathbf{X} = T\mathbf{p} \cdot \mathbf{X}\,.$$

Moreover, given two sections $\mathbf{s}, \overline{\mathbf{s}} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ such that $\mathbf{s}(\mathbf{x}) = \overline{\mathbf{s}}(\mathbf{x})$ and any vector $\mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$, we have that $T\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}}, T\overline{\mathbf{s}} \cdot \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E}$. Their difference is then meanginful and is a vertical vector:

$$T\mathbf{p} \circ (T\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}} - T\overline{\mathbf{s}} \cdot \mathbf{v}_{\mathbf{x}}) = \mathbf{v}_{\mathbf{x}} - \mathbf{v}_{\mathbf{x}} = 0\,.$$

This simple property has far reaching consequences being at the base of the concept of horizontal lifting and hence of connection on a fibre bundle (see section 1.4.1, page 87).

1.2.16 Vertical bundle of a vector bundle

In a vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, any fibre $\mathbb{E}_{\mathbf{p}(\mathbf{e})} := \mathbf{p}^{-1}\{\mathbf{p}(\mathbf{e})\}$ is a linear space and the tangent space $\mathbb{T}_{\mathbf{e}}\mathbb{E}_{\mathbf{p}(\mathbf{e})}$ at $\mathbf{e} \in \mathbb{E}$ to a fibre $\mathbb{E}_{\mathbf{p}(\mathbf{e})}$ over $\mathbf{p}(\mathbf{e}) \in \mathbb{M}$ may be identified with the fibre itself, i.e.

$$\mathbb{V}_{\mathbf{e}}\mathbb{E} = \mathbb{T}_{\mathbf{e}}\mathbb{E}_{\mathbf{p}(\mathbf{e})} \simeq \mathbb{E}_{\mathbf{p}(\mathbf{e})}\,, \quad \forall\, \mathbf{e} \in \mathbb{E}\,.$$

Accordingly, a vertical vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{VE})$ may be identified with a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ of the vector bundle. Identifications may however be a source of misinterpretation of the results. It is then preferable to consider in each fibre $\mathbb{E}_{\mathbf{p}(\mathbf{e})}$ a straight line through a point, say $\mathbf{e} \in \mathbb{E}_{\mathbf{p}(\mathbf{e})}$ which is represented in parametric form by the affine map:

$$\mathbf{aff}^{t}_{\mathbb{E}\times_{\mathbb{M}}\mathbb{E}}(\mathbf{e}, \boldsymbol{\eta}) := \mathbf{e} + t\boldsymbol{\eta}\,, \quad \boldsymbol{\eta} \in \mathbb{E}_{\mathbf{p}(\mathbf{e})}\,, \quad t \in \Re\,.$$

The parallel line through the origin is represented by the linear map

$$\mathbf{mult}^{\,t}_{\mathbb{E}\times_{\mathbb{M}}\mathbb{E}}(\boldsymbol{\eta}) := \mathbf{aff}^{t}_{\mathbb{E}\times_{\mathbb{M}}\mathbb{E}}(\mathbf{0}, \boldsymbol{\eta}) = t\boldsymbol{\eta}\,.$$

Definition 1.2.22 (Vertical lift) *The vector bundle isomorphism between the product vector bundle* $\mathbb{E} \times_{\mathbb{M}} \mathbb{E}$ *and the vertical vector bundle* $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{V}\mathbb{E}\,;\mathbb{E})$, *defined by:*

$$\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})} := \partial_{t=0}\, \boldsymbol{aff}^{\,t}_{\mathbb{E}\times_{\mathbb{M}}\mathbb{E}}\,,$$

and explicitly: $\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}\,,\boldsymbol{\eta}) := \partial_{t=0}\,(\mathbf{e}+t\boldsymbol{\eta}) \in \mathbb{V}_{\mathbf{e}}\mathbb{E}$ *for all* $(\mathbf{e}\,,\boldsymbol{\eta}) \in \mathbb{E}\times_{\mathbb{M}}\mathbb{E}$, *is called the (full) vertical lift* $\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \in \mathrm{C}^1(\mathbb{E}\times_{\mathbb{M}}\mathbb{E}\,;\mathbb{T}\mathbb{E})$.

For any given $\mathbf{e} \in \mathbb{E}$, the linear map $\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \in BL(\mathbb{E}_{\mathbf{p}(\mathbf{e})}\,;\mathbb{T}_{\mathbf{e}}\mathbb{E})$, called the vertical lift at $\mathbf{e} \in \mathbb{E}$, associates with any vector $\boldsymbol{\eta} \in \mathbb{E}_{\mathbf{p}(\mathbf{e})}$ the vertical vector

$$\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e})\cdot\boldsymbol{\eta} := \mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}\,,\boldsymbol{\eta}) \in \mathbb{V}_{\mathbf{e}}\mathbb{E} = \mathbb{T}_{\mathbf{e}}\mathbb{E}_{\mathbf{p}(\mathbf{e})}\,.$$

The map $\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \in BL(\mathbb{E}_{\mathbf{p}(\mathbf{e})}\,;\mathbb{V}_{\mathbf{e}}\mathbb{E})$ is a linear isomorphism.

Definition 1.2.23 (Liouville vector field) *In a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, *the* LIOUVILLE *vector field* $\mathbf{C}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{T}\mathbb{E})$ *is the vertical-valued vector field defined by*

$$\mathbf{C}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) := \mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}\,,\mathbf{e}) \in \mathbb{T}_{\mathbf{e}}\mathbb{E}_{\mathbf{p}(\mathbf{e})}\,,$$

so that $\boldsymbol{\tau}_{\mathbb{E}} \circ \mathbf{C}_{(\mathbb{E},\mathbf{p},\mathbb{M})} = \mathbf{id}_{\mathbb{E}}$. *Defining* $\mathrm{DIAG} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{E}\times_{\mathbb{M}}\mathbb{E})$ *by* $\mathrm{DIAG}(\mathbf{e}) = (\mathbf{e}\,,\mathbf{e})$ *we may set* $\mathbf{C}_{(\mathbb{E},\mathbf{p},\mathbb{M})} := \mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \circ \mathrm{DIAG}$.

Definition 1.2.24 (Vertical drill) *In a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, *the vertical drill at* $\mathbf{e} \in \mathbb{E}$ *is the linear map* $\mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \in BL(\mathbb{V}_{\mathbf{e}}\mathbb{E}\,;\mathbb{E}_{\mathbf{p}(\mathbf{e})})$ *which is the inverse of the vertical lift* $\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \in BL(\mathbb{E}_{\mathbf{p}(\mathbf{e})}\,;\mathbb{V}_{\mathbf{e}}\mathbb{E})$:

$$\begin{cases}\mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \circ \mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) = \mathbf{id}_{\mathbb{E}_{\mathbf{p}(\mathbf{e})}}\,,\\ \mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \circ \mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) = \mathbf{id}_{\mathbb{V}_{\mathbf{e}}\mathbb{E}}\,.\end{cases}$$

This fibrewise correspondence induces a surjective homomorphism from the vector bundle $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{V}\mathbb{E}\,;\mathbb{E})$ onto the vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, which we call the *vertical drill* $\mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \in \mathrm{C}^1(\mathbb{V}\mathbb{E}\,;\mathbb{E})$, defined by:

$$(\mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})}\cdot\mathbf{V}) := \mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{V}))\cdot\mathbf{V}\,, \quad \forall\,\mathbf{V} \in \mathbb{V}\mathbb{E}\,.$$

Then the vertical drill associates with a vertical vector $\mathbf{V} \in \mathbb{V}_{\mathbf{e}}\mathbb{E}$ the vector $\boldsymbol{\eta} \in \mathbb{E}_{\mathbf{p}(\mathbf{e})}$ whose vertical lift at $\mathbf{e} \in \mathbb{E}$ is equal to $\mathbf{V}$. Following [80] we define the *small vertical lift* as:

$$\mathbf{vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}\cdot\boldsymbol{\eta} := \mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{0})\cdot\boldsymbol{\eta} = \partial_{t=0}\,\mathbf{mult}^{\,t}_{\mathbb{E}\times_{\mathbb{M}}\mathbb{E}} \in \mathbb{V}_{\mathbf{0}}\mathbb{E}\,, \quad \forall\,\boldsymbol{\eta} \in \mathbb{E}\,,$$

so that $\mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \circ \mathbf{vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})} = \mathbf{id}_{\mathbb{E}}$.

Lemma 1.2.10 *In a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *the following diagrams commute:*

$$\begin{array}{ccc} \mathbb{TM} \times_{\mathbb{M}} \mathbb{TM} & \xrightarrow{\boldsymbol{aff}^t_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}}} & \mathbb{TM} \\ {\scriptstyle T\mathbf{s}}\big\downarrow & & \big\downarrow{\scriptstyle T\mathbf{s}} \\ \mathbb{TE} \times_{\mathbb{E}} \mathbb{TE} & \xrightarrow{\boldsymbol{aff}^t_{\mathbb{TE}\times_{\mathbb{E}}\mathbb{TE}}} & \mathbb{TE} \end{array} \quad \iff \quad \boldsymbol{aff}^t_{\mathbb{TE}\times_{\mathbb{E}}\mathbb{TE}} \circ T\mathbf{s} = T\mathbf{s} \circ \boldsymbol{aff}^t_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}}\,,$$

$$\begin{array}{ccc} \mathbb{TM} \times_{\mathbb{M}} \mathbb{TM} & \xrightarrow{\mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})}} & \mathbb{VTM} \\ {\scriptstyle T\mathbf{s}}\big\downarrow & & \big\downarrow{\scriptstyle T^2\mathbf{s}} \\ \mathbb{TE} \times_{\mathbb{E}} \mathbb{TE} & \xrightarrow{\mathbf{Vl}_{(\mathbb{TE},\boldsymbol{\tau}_{\mathbb{E}},\mathbb{E})}} & \mathbb{VTE} \end{array} \quad \iff \quad \mathbf{Vl}_{(\mathbb{TE},\boldsymbol{\tau}_{\mathbb{E}},\mathbb{E})} \circ T\mathbf{s} = T^2\mathbf{s} \circ \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})}\,,$$

where $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *is a section of the bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *and we have introduced the notation:*

$$T\mathbf{s} \cdot (\mathbf{a}\,,\mathbf{b}) := (T\mathbf{s} \cdot \mathbf{a}\,, T\mathbf{s} \cdot \mathbf{b}) \in \mathbb{TE} \times_{\mathbb{E}} \mathbb{TE}\,, \quad \forall\,(\mathbf{a}\,,\mathbf{b}) \in \mathbb{TM} \times_{\mathbb{M}} \mathbb{TM}\,.$$

We have likewise the relations:

$$\begin{aligned} \boldsymbol{mult}^t_{\mathbb{TE}\times_{\mathbb{E}}\mathbb{TE}} \circ T\mathbf{s} &= T\mathbf{s} \circ \boldsymbol{mult}^t_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}} \\ \mathbf{vl}_{\mathbb{TE}\times_{\mathbb{E}}\mathbb{TE}} \circ T\mathbf{s} &= T^2\mathbf{s} \circ \mathbf{vl}_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}}\,. \end{aligned}$$

Proof. It is enough to prove the former equality.

$$\begin{aligned} \mathbf{aff}^t_{\mathbb{TE}\times_{\mathbb{E}}\mathbb{TE}} \circ T\mathbf{s} \circ (\mathbf{a}\,,\mathbf{b}) &= \mathbf{aff}^t_{\mathbb{TE}\times_{\mathbb{E}}\mathbb{TE}} \circ (T\mathbf{s} \cdot \mathbf{a}\,, T\mathbf{s} \cdot \mathbf{b}) \\ &= T\mathbf{s} \cdot \mathbf{a} + t\,T\mathbf{s} \cdot \mathbf{b} = T\mathbf{s} \cdot (\mathbf{a} + t\,\mathbf{b}) \\ &= T\mathbf{s} \circ \mathbf{aff}^t_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}} \circ (\mathbf{a}\,,\mathbf{b})\,, \quad \forall\,(\mathbf{a}\,,\mathbf{b}) \in \mathbb{TM} \times_{\mathbb{M}} \mathbb{TM}\,. \end{aligned}$$

The latter equality is the $\partial_{t=0}$ derivative of the former one. ■

It is usual to identify the vertical drill with the identity map to simplify the exposition. We will adopt this point of view with some significant exceptions where the distinction between the vertical space $\mathbb{V}_{\mathbf{e}}\mathbb{E}$ and the linear space $\mathbb{E}_{\mathbf{e}}$ is essential for a correct presentation.

Lemma 1.2.11 (Correction flow) *A flow* $\mathbf{Fl}^{\mathbf{X}}_\lambda \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ *is a bundle automorphism, i.e. fiber-preserving and invertible for each* $\lambda \in I$ *, if and only if the velocity vector field* $\mathbf{X} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *is expressed by the sum*

$$\mathbf{X} = \mathbf{V} + \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v}\,,$$

where $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *is the projection of the vector field* $\mathbf{X} \in \mathrm{C}^0(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *to the base manifold and* $\mathbf{V} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *is a vertical vector field.*

Proof. Let us assume that the flow $\mathbf{Fl}^{\mathbf{X}}_\lambda \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ is a tangent-bundle morphism. Then the base flow $\mathbf{Fl}^{\mathbf{v}}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ is well defined. By acting with the tangent functor, we get the lifted flow $T\mathbf{Fl}^{\mathbf{v}}_\lambda \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ which for each $\lambda \in I$ is a fibre-bundle endomorphism over the same base flow as $\mathbf{Fl}^{\mathbf{X}}_\lambda \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$. It follows that the *correction flow*:

$$\mathbf{Fl}^{\mathbf{V}}_\lambda := \mathbf{Fl}^{\mathbf{X}}_\lambda \circ (T\mathbf{Fl}^{\mathbf{v}}_\lambda)^{-1} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})\,,$$

projects to the identity: $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{Fl}^{\mathbf{V}}_\lambda = \mathbf{id}_{\mathbb{M}} \circ \boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$. Taking the derivative $\partial_{\lambda=0}$ of the correction flow and invoking Lemma 1.2.7, we get: $\mathbf{V} = \mathbf{X} - \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ while, taking the derivative $\partial_{\lambda=0}$ of the projected flow, we get the verticality property:

$$T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{V} = \partial_{\lambda=0}\, \boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{Fl}^{\mathbf{V}}_\lambda = \partial_{\lambda=0}\, \mathbf{id}_{\mathbb{M}} \circ \boldsymbol{\tau}_{\mathbb{M}} = 0\,.$$

The converse implication is proved by reversing the arguments' order. From the decomposition formula we infer that

$$T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{X} = T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{V} + T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} = \boldsymbol{\tau}_{\mathbb{TM}} \circ T\mathbf{v} = \mathbf{v} \circ \boldsymbol{\tau}_{\mathbb{M}}\,,$$

and the result follows by corollary 1.2.2, page 42. ■

1.2.17 Automorphic flows

Definition 1.2.25 *A flow* $\mathbf{Fl}^{\mathbf{X}}_\lambda \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{E})$ *is said to be automorphic if for each* $\lambda \in \Re$ *it is a linear vector bundle automorphism, that is a fibre-preserving, fibre-linear and invertible map from the vector bundle* $(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})$ *onto itself.*

Then the base flow $\mathbf{Fl}^{\mathbf{v}}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ and its velocity vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ are well defined by the commutative diagrams

$$\begin{array}{ccccc} \mathbb{E} & \xrightarrow{\mathbf{Fl}^{\mathbf{X}}_\lambda} & \mathbb{E} & \xrightarrow{\mathbf{X}} & \mathbb{TE} \\ {\scriptstyle \mathbf{p}}\downarrow & & \downarrow{\scriptstyle \mathbf{p}} & & \downarrow{\scriptstyle T\mathbf{p}} \\ \mathbb{M} & \xrightarrow{\mathbf{Fl}^{\mathbf{v}}_\lambda} & \mathbb{M} & \xrightarrow{\mathbf{v}} & \mathbb{TM} \end{array} \quad \iff \quad \begin{cases} \mathbf{p} \circ \mathbf{Fl}^{\mathbf{X}}_\lambda = \mathbf{Fl}^{\mathbf{v}}_\lambda \circ \mathbf{p}\,, \\ T\mathbf{p} \circ \mathbf{X} = \mathbf{v} \circ \mathbf{p}\,. \end{cases}$$

and the vector field $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ projects over the vector field $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$.

Lemma 1.2.12 (Automorphic flows) *Let* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ *be a vector bundle and* $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ *a vector field projecting over* $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$. *Then the pair* $(\mathbf{X}\,,\mathbf{v})$ *is a homomorphism from the bundle* $(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})$ *to the bundle* $(\mathbb{TE}\,,T\mathbf{p}\,,\mathbb{TM})$ *iff the associated flow* $\mathbf{Fl}^{\mathbf{X}}_{\lambda} \in C^1(\mathbb{E}\,;\mathbb{E})$ *is automorphic.*

Proof. The fibrewise linearity of $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ is expressed by

$$\begin{aligned} \mathbf{X}(\alpha \cdot_{\mathbf{p}} \mathbf{e}_{\mathbf{x}}) &= \alpha \cdot_{T\mathbf{p}} \mathbf{X}(\mathbf{e}_{\mathbf{x}}) \in \mathbb{T}_{\mathbf{e}_{\mathbf{x}}}\mathbb{E}\,, && \forall\, \mathbf{e}_{\mathbf{x}} \in \mathbb{E}_{\mathbf{x}}\,, \alpha \in \Re\,, \\ \mathbf{X}(\mathbf{e}_{1\mathbf{x}} +_{\mathbf{p}} \mathbf{e}_{2\mathbf{x}}) &= \mathbf{X}(\mathbf{e}_{1\mathbf{x}}) +_{T\mathbf{p}} \mathbf{X}(\mathbf{e}_{2\mathbf{x}})\,, && \forall\, \mathbf{e}_{1\mathbf{x}}, \mathbf{e}_{2\mathbf{x}} \in \mathbb{E}_{\mathbf{x}}\,, \end{aligned}$$

where $\mathbf{X}(\mathbf{e}_{1\mathbf{x}}), \mathbf{X}(\mathbf{e}_{2\mathbf{x}}) \in (T\mathbf{p})^{-1}\{\mathbf{v}_{\mathbf{x}}\}$. By corollary 1.2.2, page 42, the flow $\mathbf{Fl}^{\mathbf{X}}_{\lambda} \in C^1(\mathbb{E}\,;\mathbb{E})$ projects on the flow $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in C^1(\mathbb{M}\,;\mathbb{M})$. Since the map $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ is fibre-respecting over $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$:

$$\mathbf{e}_{1\mathbf{x}}, \mathbf{e}_{2\mathbf{x}} \in \mathbb{E}_{\mathbf{x}} \implies \mathbf{X}(\mathbf{e}_{1\mathbf{x}}), \mathbf{X}(\mathbf{e}_{2\mathbf{x}}) \in (T\mathbf{p})^{-1}\{\mathbf{v}_{\mathbf{x}}\}\,,$$

and the sum $+_{T\mathbf{p}}$ is well-defined. We must prove that:

$$\begin{aligned} &\mathbf{Fl}^{\mathbf{X}}_{\lambda}(\alpha\, \mathbf{e}_{\mathbf{x}}) = \alpha\, \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{\mathbf{x}}) \in \mathbb{E}_{\mathbf{Fl}^{\mathbf{v}}_{\lambda}(\mathbf{x})}\,, \quad \forall\, \mathbf{e} \in \mathbb{E}\,, \alpha \in \Re\,, \\ &\mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{1\mathbf{x}} +_{\mathbf{p}} \mathbf{e}_{2\mathbf{x}}) = \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{1\mathbf{x}}) +_{\mathbf{p}} \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{2\mathbf{x}}) \in \mathbb{E}_{\mathbf{Fl}^{\mathbf{v}}_{\lambda}(\mathbf{x})} \quad \forall\, \mathbf{e}_{1\mathbf{x}}, \mathbf{e}_{2\mathbf{x}} \in \mathbb{E}_{\mathbf{x}}\,. \end{aligned}$$

The result follows from the uniqueness of the solution of the differential equation defining the flow. Fiberwise homogeneity is inferred from the equalities:

$$\begin{aligned} \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\alpha \cdot_{\mathbf{p}} \mathbf{e}_{\mathbf{x}}) &= \partial_{\lambda=0}\, (\mathbf{Fl}^{\mathbf{X}}_{\lambda} \circ \mathbf{mult}^{\,\alpha}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})})(\mathbf{e}_{\mathbf{x}}) \\ &= \mathbf{X}(\mathbf{mult}^{\,\alpha}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}_{\mathbf{x}})) \\ &= T\mathbf{mult}^{\,\alpha}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}_{\mathbf{x}}) \cdot \mathbf{X}(\mathbf{e}_{\mathbf{x}}) \\ &= T\mathbf{mult}^{\,\alpha}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}_{\mathbf{x}}) \cdot \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{\mathbf{x}}) \\ &= \partial_{\lambda=0}\, (\mathbf{mult}^{\,\alpha}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})} \circ \mathbf{Fl}^{\mathbf{X}}_{\lambda})(\mathbf{e}_{\mathbf{x}}) \\ &= \partial_{\lambda=0}\, \alpha \cdot_{\mathbf{p}} \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{\mathbf{x}})\,. \end{aligned}$$

Fiberwise additivity is inferred from the relation:

$$\begin{aligned}
\partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{1\mathbf{x}} +_{\mathbf{p}} \mathbf{e}_{2\mathbf{x}}) &= \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{add}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}_{1\mathbf{x}}\,, \mathbf{e}_{2\mathbf{x}})) \\
&= \mathbf{X}(\mathbf{add}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}_{1\mathbf{x}}\,, \mathbf{e}_{2\mathbf{x}})) \\
&= T\mathbf{add}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}_{1\mathbf{x}}\,, \mathbf{e}_{2\mathbf{x}}) \cdot (\mathbf{X}(\mathbf{e}_{1\mathbf{x}})\,, \mathbf{X}(\mathbf{e}_{2\mathbf{x}})) \\
&= T\mathbf{add}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{e}_{1\mathbf{x}}\,, \mathbf{e}_{2\mathbf{x}}) \cdot \partial_{\lambda=0}\, (\mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{1\mathbf{x}})\,, \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{2\mathbf{x}})) \\
&= \partial_{\lambda=0}\, \mathbf{add}_{(\mathbb{E}\,,\mathbf{p}\,,\mathbb{M})}(\mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{1\mathbf{x}})\,, \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{2\mathbf{x}})) \\
&= \partial_{\lambda=0}\, (\mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{1\mathbf{x}}) +_{\mathbf{p}} \mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{2\mathbf{x}}))\,.
\end{aligned}$$

We underline that $\mathbf{p}(\mathbf{e}_{1\mathbf{x}}) = \mathbf{p}(\mathbf{e}_{2\mathbf{x}}) = \mathbf{x} \in \mathbb{M}$ and then, by fibre-preservation, also $\mathbf{p}(\mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{1\mathbf{x}})) = \mathbf{p}(\mathbf{Fl}^{\mathbf{X}}_{\lambda}(\mathbf{e}_{2\mathbf{x}})) = \mathbf{Fl}^{\mathbf{v}}_{\lambda}(\mathbf{x})$ and, taking the derivative $\partial_{\lambda=0}$, also $T\mathbf{p} \cdot \mathbf{X}(\mathbf{e}_{1\mathbf{x}}) = T\mathbf{p} \cdot \mathbf{X}(\mathbf{e}_{2\mathbf{x}}) = \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$. ■

1.3 Lie derivative

The LIE *derivative* or *convective derivative* at $\mathbf{x} \in \mathbb{M}$ of a scalar, vector or tensor field on $\mathbb{M}$, along a vector field $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$, is the rate of variation of their pull back at $\mathbf{x} \in \mathbb{M}$ along the flow $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in C^1(\mathbb{M}\,;\mathbb{M})$ of $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$.

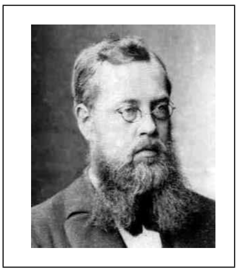

Figure 1.14: Marius Sophus Lie (1842 - 1899)

Then we have that:

- The LIE derivative $\mathcal{L}_{\mathbf{v}} f \in C^0(\mathbb{M}\,;\mathbb{T}\Re)$ of a scalar field $f \in C^1(\mathbb{M}\,;\Re)$ along the flow $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in C^1(\mathbb{M}\,;\mathbb{M})$ of a vector field $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ coincides with its directional derivative.

Indeed the chain rule of differentiation shows that

$$\mathcal{L}f \circ \mathbf{v} = \mathcal{L}_{\mathbf{v}} f := \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\! f = \partial_{\lambda=0}\, (f \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}) = Tf \circ \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\Re)\,.$$

Hence $\mathcal{L}f = Tf \in \mathrm{C}^0(\mathbb{TM}\,;\mathbb{T}\Re)$ with the commutative diagram:

$$\begin{array}{ccc} \Re & \xleftarrow{\;\boldsymbol{\tau}_{\Re}\;} & \mathbb{T}\Re \\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle Tf} \\ \mathbb{M} & \xrightarrow{\;\mathbf{v}\;} & \mathbb{TM} \end{array} \quad \iff \quad f = \boldsymbol{\tau}_{\Re} \circ \mathcal{L}f \circ \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\Re)\,.$$

By the identification $\mathbb{T}_t\Re \simeq \Re$ for all $t \in \Re$ we may write $\mathcal{L}_{\mathbf{v}} f \in \mathrm{C}^0(\mathbb{TM}\,;\Re)$.

Definition 1.3.1 (Lie derivative) *The* LIE *derivative of a vector field* $\mathbf{u} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *along the flow* $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *of a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is the vertical-valued vector field defined by*

$$\partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u} := \partial_{\lambda=0}\, T\mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathbb{V}_{\mathbf{u}}\mathbb{TM}\,.$$

The verticality of the LIE derivative is a direct consequence of the fact that the curve of tangent vector fields $(T\mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda})(\mathbf{x}) \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ passes through $\mathbf{u}(\mathbf{x})$ at $\lambda = 0$ and evolves in the linear fibre $\mathbb{T}_{\mathbf{x}}\mathbb{M}$. Accordingly, the base curve $(\boldsymbol{\tau}_{\mathbb{M}} \circ T\mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda})(\mathbf{x}) = (\mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda})(\mathbf{x}) = (\mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda})(\mathbf{x}) = \mathbf{x}$ degenerates to the point $\mathbf{x} \in \mathbb{M}$.

Lemma 1.3.1 (Lie derivative) *The* LIE *derivative of a tangent vector field* $\mathbf{u} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *along the flow* $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *of a tangent vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *may be evaluated by the formula:*

$$\partial_{\lambda=0}\, (\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u}) = T\mathbf{u} \circ \mathbf{v} - \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{u}\,.$$

Proof. A direct computation, based on LEIBNIZ and chain rules and on Lemma 1.2.7, gives the result:

$$\begin{aligned} \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u} &= \partial_{\lambda=0}\, T\mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \\ &= \partial_{\lambda=0}\, \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} - \partial_{\lambda=0}\, T\mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{u} \\ &= \partial_{\lambda=0}\, \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} - \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(T\mathbf{v})}_{\lambda} \circ \mathbf{u} \\ &= T\mathbf{u} \circ \mathbf{v} - \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{u}\,. \end{aligned}$$

Recalling that the flip commutes between the two vector bundle structures on $\mathbb{TM}$, the result can also be proved by the following computation:

$$\begin{aligned}\partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u} &= \partial_{\lambda=0}\,T\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\circ\mathbf{u}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\\ &= \partial_{\lambda=0}\,\partial_{\mu=0}\,\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\circ\mathbf{Fl}^{\mathbf{u}}_{\mu}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\\ &= \mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ\partial_{\mu=0}\,\partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\circ\mathbf{Fl}^{\mathbf{u}}_{\mu}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\\ &= \mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ(\partial_{\mu=0}\,T\mathbf{Fl}^{\mathbf{u}}_{\mu}\circ\mathbf{v}-_{T\boldsymbol{\tau}_{\mathbb{M}}}\partial_{\mu=0}\,\mathbf{v}\circ\mathbf{Fl}^{\mathbf{u}}_{\mu})\\ &= \mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ T\mathbf{u}\circ\mathbf{v}-_{T\boldsymbol{\tau}_{\mathbb{M}}}T\mathbf{v}\circ\mathbf{u})\\ &= T\mathbf{u}\circ\mathbf{v}-_{\boldsymbol{\tau}_{\mathbb{TM}}}\mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ T\mathbf{v}\circ\mathbf{u}\,.\end{aligned}$$

The vectors $(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ T\mathbf{u}\circ\mathbf{v})(\mathbf{x})$ and $(T\mathbf{v}\circ\mathbf{u})(\mathbf{x})$ belong to the same fibre $\mathbb{T}_{\mathbf{v}_{\mathbf{x}}}\mathbb{TM}$ of $\boldsymbol{\tau}_{\mathbb{TM}}\in C^1(\mathbb{T}^2\mathbb{M};\mathbb{TM})$ and have the same base-velocity $\mathbf{u}_{\mathbf{x}}\in\mathbb{TM}$ and hence belong also to the same fibre in the vector bundle $T\boldsymbol{\tau}_{\mathbb{M}}\in C^1(\mathbb{T}^2\mathbb{M};\mathbb{TM})$. Their sum in this latter bundle is well-defined and is the one to be performed to get the vector $\partial_{\mu=0}\,\partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\circ\mathbf{Fl}^{\mathbf{u}}_{\mu}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda})(\mathbf{x})$ which is based at $\mathbf{0}_{\mathbf{x}}\in\mathbb{TM}$ in $\boldsymbol{\tau}_{\mathbb{TM}}\in C^1(\mathbb{T}^2\mathbb{M};\mathbb{TM})$ with base velocity equal to $\mathbf{u}_{\mathbf{x}}\in\mathbb{TM}$. Then the flipped vector is based at $\mathbf{u}_{\mathbf{x}}\in\mathbb{TM}$ and is vertical. ■

Lemma 1.3.2 (Lie-derivative vector field) *The* Lie*-derivative vector field* $\mathcal{L}_{\mathbf{v}}\mathbf{u}\in C^1(\mathbb{M};\mathbb{TM})$ *is well-defined by the relation*

$$\mathbf{Vl}_{\mathbb{TM}}(\mathbf{u})\cdot\mathcal{L}_{\mathbf{v}}\mathbf{u}=\partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u})=T\mathbf{u}\circ\mathbf{v}-\mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ T\mathbf{v}\circ\mathbf{u}\,.$$

Equivalently we may put:

$$\mathcal{L}_{\mathbf{v}}\mathbf{u}=\mathbf{vd}_{\mathbb{TM}}(\partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u})=\mathbf{vd}_{\mathbb{TM}}(T\mathbf{u}\circ\mathbf{v}-\mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ T\mathbf{v}\circ\mathbf{u})\,.$$

Proof. The statement holds by injectivity of the vertical lift at $\mathbf{u}_{\mathbf{x}}\in\mathbb{TM}$, i.e. the linear map $\mathbf{Vl}_{\mathbb{TM}}(\mathbf{u}_{\mathbf{x}})\in C^1(\mathbb{T}_{\mathbf{x}}\mathbb{M};\mathbb{V}_{\mathbf{u}_{\mathbf{x}}}\mathbb{TM})$ with $\mathbb{V}_{\mathbf{u}_{\mathbf{x}}}\mathbb{TM}=\mathbb{T}_{\mathbf{u}_{\mathbf{x}}}\mathbb{T}_{\mathbf{x}}\mathbb{M}$. ■

The result of Lemma 1.3.1 and 1.3.2 are proved in a chart in [80], Lemma 8.14. It is customary to drop the vertical drill and to write $\mathcal{L}_{\mathbf{v}}\mathbf{u}=\partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u})$.

The Lie derivative of a $(1,1)$ tensor field $\mathbf{A}\in C^1(\mathbb{M};BL(\mathbb{TM},\mathbb{T}^*\mathbb{M};\Re))$ (one time covariant and one time contravariant) is similarly defined by:

$$\mathcal{L}_{\mathbf{v}}\mathbf{A}:=\partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{A})\,.$$

That is, for $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ and $\boldsymbol{\omega} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ we have that

$$\begin{aligned}(\mathcal{L}_{\mathbf{v}}\mathbf{A})(\mathbf{v},\boldsymbol{\omega}) &= \partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!(\mathbf{A}(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\uparrow\!\mathbf{v},\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\uparrow\!\boldsymbol{\omega}))\\ &= \partial_{\lambda=0}\,(\mathbf{A}(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\uparrow\!\mathbf{v},\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\uparrow\!\boldsymbol{\omega}))\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\,.\end{aligned}$$

Proposition 1.3.1 (Pull back of the Lie derivative along a flow) *The pull back of the* Lie *derivative of a tensor field is equal to the time derivative of its pull back, that is*

$$\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!(\mathcal{L}_{\mathbf{v}}\mathbf{A}) = \partial_{\mu=\lambda}\,(\mathbf{Fl}^{\mathbf{v}}_{\mu}\!\downarrow\!\mathbf{A})\,.$$

Proof. We recall that $\mathbf{Fl}^{\mathbf{v}}_{\lambda+\mu}\!\downarrow\!\mathbf{A} = \big(\mathbf{Fl}^{\mathbf{v}}_{\mu}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\big)\!\downarrow\!\mathbf{A} = \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\big(\mathbf{Fl}^{\mathbf{v}}_{\mu}\!\downarrow\!\mathbf{A}\big)$. Observing that $\partial_{\mu=\lambda}\,(\mathbf{Fl}^{\mathbf{v}}_{\mu}\!\downarrow\!\mathbf{A}) = \partial_{\mu=0}\,\mathbf{Fl}^{\mathbf{v}}_{\lambda+\mu}\!\downarrow\!\mathbf{A} = \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!(\partial_{\mu=0}\,\mathbf{Fl}^{\mathbf{v}}_{\mu}\!\downarrow\!\mathbf{A}) = \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathcal{L}_{\mathbf{v}}\mathbf{A}$, the result is proved. ■

If the Lie derivative vanishes identically along the flow, Proposition 1.3.1 implies that

$$\partial_{\mu=\lambda}\,(\mathbf{Fl}^{\mathbf{v}}_{\mu}\!\downarrow\!\mathbf{A}) = 0\,, \quad \forall\,\lambda\in I\,,$$

that is

$$\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{A} = \mathbf{Fl}^{\mathbf{v}}_{0}\!\downarrow\!\mathbf{A} = \mathbf{A}\,, \quad \forall\,\lambda\in I\,.$$

Therefore we have that

- The Lie derivative $\mathcal{L}_{\mathbf{v}}\mathbf{A}$ vanishes identically if and only if the tensor field $\mathbf{A} \in \mathrm{C}^1(\mathbb{M}\,;BL\,(\mathbb{TM},\mathbb{T}^*\mathbb{M}\,;\Re))$ is dragged along the flow. In particular the Lie derivative $\mathcal{L}_{\mathbf{v}}f$ of a scalar field $f \in \mathrm{C}^1(\mathbb{M}\,;\Re)$ vanishes identically if and only if the scalar field is constant along the flow.

As a consequence of Propositions 1.2.1 (page 39) and 1.3.1 we have that

Proposition 1.3.2 (Lie derivative and commutation) *The* Lie *derivative of a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *along a vector field* $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *vanishes if and only if the flows of the two vector fields commute.*

Proposition 1.3.3 (Leibniz rule) *The* Lie *derivative fulfils the* Leibniz *rule*

$$\begin{aligned}&\mathcal{L}_{\mathbf{v}}(\mathbf{A}\,(\mathbf{u},\boldsymbol{\omega})) = (\mathcal{L}_{\mathbf{v}}\mathbf{A})\,(\mathbf{u},\boldsymbol{\omega}) + \mathbf{A}\,(\mathcal{L}_{\mathbf{v}}\mathbf{u},\boldsymbol{\omega}) + \mathbf{A}\,(\mathbf{u},\mathcal{L}_{\mathbf{v}}\boldsymbol{\omega})\,,\\ &\forall\,\mathbf{v}:\mathbb{M}\mapsto\mathbb{T}^*\mathbb{M}\,,\quad \forall\,\mathbf{u}:\mathbb{M}\mapsto\mathbb{TM}\,,\quad \forall\,\boldsymbol{\omega}:\mathbb{M}\mapsto\mathbb{T}^*\mathbb{M}\,.\end{aligned}$$

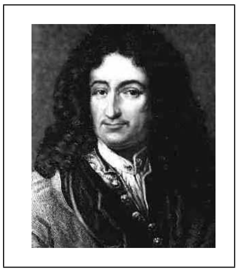

Figure 1.15: Gottfried Wilhelm von Leibniz (1646 - 1716)

Proof. Recalling that $(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{A})\,(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u},\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\boldsymbol{\omega}) = \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!(\mathbf{A}\,(\mathbf{u},\boldsymbol{\omega}))$, the result follows by taking the derivative $\partial_{\lambda=0}$ and applying the partial and chain differentiation rules. ■

Lemma 1.3.3 (A commutation property) *Let* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *be an invertible morphism and* $\mathbf{v} \in \mathrm{C}^1(\Re\,;\mathbb{TM})$ *be a differentiable one-parameter family of tangent vectors such that* $\partial_{\mu=\lambda}\,\mathbf{v}(\mu) \in \mathbb{VTM}$. *Then the following commutation property holds:*

$$\partial_{\mu=\lambda}\,\boldsymbol{\varphi}\!\downarrow\!\mathbf{v}(\mu) = \boldsymbol{\varphi}\!\downarrow\!\partial_{\mu=\lambda}\,\mathbf{v}(\mu)\,.$$

Proof. A direct computation gives:

$$\partial_{\mu=\lambda}\,\boldsymbol{\varphi}\!\downarrow\!\mathbf{v}(\mu) = T\boldsymbol{\varphi}\!\downarrow\!(\mathbf{v}(\lambda))\cdot\partial_{\mu=\lambda}\,\mathbf{v}(\mu) = \boldsymbol{\varphi}\!\downarrow\!\partial_{\mu=\lambda}\,\mathbf{v}(\mu)\,,$$

being $T\boldsymbol{\varphi}\!\downarrow\!(\mathbf{v}(\lambda)) = \boldsymbol{\varphi}\!\downarrow$ by the fibrewise linearity of $\boldsymbol{\varphi}\!\downarrow$. ■

Proposition 1.3.4 (Lie derivative of pull and push) *Let* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *be an invertible morphism between two manifolds* $\mathbb{M}$ *and* $\mathbb{N}$. *Then, for any given vector field* $\mathbf{u} \in \mathrm{C}^0(\mathbb{M},\mathbb{TM})$, *scalar field* $f \in \mathrm{C}^0(\mathbb{N}\,;\Re)$, *vector field* $\mathbf{v} \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}^*\mathbb{N})$ *and covector field* $\mathbf{v}^* \in \mathrm{C}^0(\mathbb{N}\,;\mathbb{T}^*\mathbb{N})$, *the following formulas holds*

$$\boldsymbol{\varphi}\!\downarrow\!(\mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{u}} f) = \mathcal{L}_{\mathbf{u}}\,\boldsymbol{\varphi}\!\downarrow\! f\,,$$

$$\boldsymbol{\varphi}\!\uparrow\!(\mathcal{L}_{\mathbf{u}}\mathbf{v}) = \mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{u}}\,\boldsymbol{\varphi}\!\uparrow\!\mathbf{v}\,,$$

$$\boldsymbol{\varphi}\!\downarrow\!(\mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{u}}\mathbf{v}^*) = \mathcal{L}_{\mathbf{u}}\,\boldsymbol{\varphi}\!\downarrow\!\mathbf{v}^*\,.$$

Proof. By proposition 1.2.5 we have that $\mathbf{Fl}_\lambda^{\varphi\uparrow\mathbf{u}} \circ \varphi = \varphi \circ \mathbf{Fl}_\lambda^{\mathbf{u}}$ and hence:

$$\varphi\downarrow(\mathbf{Fl}_\lambda^{\varphi\uparrow\mathbf{u}}\downarrow f) = (\mathbf{Fl}_\lambda^{\varphi\uparrow\mathbf{u}} \circ \varphi)\downarrow f = (\varphi \circ \mathbf{Fl}_\lambda^{\mathbf{u}})\downarrow f = \mathbf{Fl}_\lambda^{\mathbf{u}}\downarrow(\varphi\downarrow f)\,,$$

$$\varphi\uparrow(\mathbf{Fl}_\lambda^{\mathbf{u}}\downarrow \mathbf{v}) = \varphi\uparrow(\mathbf{Fl}_{-\lambda}^{\mathbf{u}}\uparrow \mathbf{v}) = (\varphi \circ \mathbf{Fl}_{-\lambda}^{\mathbf{u}})\uparrow \mathbf{v} = (\mathbf{Fl}_{-\lambda}^{\varphi\uparrow\mathbf{u}} \circ \varphi)\uparrow \mathbf{v} = \mathbf{Fl}_\lambda^{\varphi\uparrow\mathbf{u}}\downarrow(\varphi\uparrow\mathbf{v})\,,$$

$$\varphi\downarrow(\mathbf{Fl}_\lambda^{\varphi\uparrow\mathbf{u}}\downarrow \mathbf{v}^*) = (\mathbf{Fl}_\lambda^{\varphi\uparrow\mathbf{u}} \circ \varphi)\downarrow \mathbf{v}^* = (\varphi \circ \mathbf{Fl}_\lambda^{\mathbf{u}})\downarrow \mathbf{v}^* = \mathbf{Fl}_\lambda^{\mathbf{u}}\downarrow(\varphi\downarrow\mathbf{v}^*)\,.$$

The formulas in the statement are then a direct consequence of the definition of Lie derivative and of the commutation property of Lemma 1.3.3 and the similar one for scalar and covectors. ■

The result of Proposition 1.3.4 leads to the statement:

- The Lie derivative is *natural* with respect to the pull or push by a diffeomorphism.

1.3.1 Lie bracket

The next proposition provides a basic characterization of the Lie derivative of a vector field along a flow. This far reaching result shows that the directional derivative of a scalar field along a Lie derivative is equal to the gap of symmetry of the iterated directional derivative of the scalar field. This antisymmetric gap is in fact a first order derivation along the direction of the tangent vector detected by the Lie derivative which is thus expressed as an antisymmetric Lie bracket of the vector fields.

Proposition 1.3.5 (Lie bracket) *Let* $\mathbf{v}, \mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *be vector fields. Then*

$$(\mathcal{L}_{\mathbf{v}}\,\mathbf{u})\,f = [\mathbf{v}\,,\mathbf{u}]\,f := \mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}}\,f - \mathcal{L}_{\mathbf{u}}\,\mathcal{L}_{\mathbf{v}}\,f\,,\quad \forall\, f \in \mathrm{C}^2(\mathbb{M}\,;\Re)\,,$$

that is: the Lie *derivative is the gap of symmetry of the iterated directional derivative of a scalar field.*

Proof. Following [3] and denoting by $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ the flow of the vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$, we have that

$$\begin{aligned}(\mathcal{L}_{\mathbf{v}}\,\mathbf{u})\,f &= \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\!\mathbf{u})\,f = \partial_{\lambda=0}\,(\mathcal{L}_{(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\downarrow\mathbf{u})}\,f)\\ &= \partial_{\lambda=0}\,(\mathcal{L}_{(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\downarrow\mathbf{u})}\,\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow(\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\!\downarrow\! f))\\ &= \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow\mathcal{L}_{\mathbf{u}}\,(\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\!\downarrow\! f)) = \mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}}\,f + \partial_{\lambda=0}\,(\mathcal{L}_{\mathbf{u}}\,(\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\!\downarrow\! f))\\ &= \mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}}\,f + \mathcal{L}_{\mathbf{u}}\,\partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\!\downarrow\! f) = \mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}}\,f + \mathcal{L}_{\mathbf{u}}\,\partial_{\lambda=0}\,(f\circ\mathbf{Fl}^{\mathbf{v}}_{-\lambda})\\ &= \mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}}\,f + \mathcal{L}_{\mathbf{u}}\,\mathcal{L}_{(\partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{-\lambda})}\,f = (\mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}} - \mathcal{L}_{\mathbf{u}}\,\mathcal{L}_{\mathbf{v}})\,f\,,\end{aligned}$$

and the result follows. ■

In the proof of Proposition 1.3.5 we have made recourse to the property of invariance under an exchange of the order of time derivatives and to the formula for the derivative of the inverse which is derived hereafter:

$$\begin{aligned}&0 = \partial_{\mu=\lambda}\,(\mathbf{Fl}^{\mathbf{v}}_{-\mu}\circ\mathbf{Fl}^{\mathbf{v}}_{\mu}) = \mathbf{Fl}^{\mathbf{v}}_{-\lambda}\circ(\partial_{\mu=\lambda}\,\mathbf{Fl}^{\mathbf{v}}_{\mu}) + (\partial_{\mu=\lambda}\,\mathbf{Fl}^{\mathbf{v}}_{-\mu})\circ\mathbf{Fl}^{\mathbf{v}}_{\mu}\\ &\Longrightarrow\ \partial_{\mu=\lambda}\,\mathbf{Fl}^{\mathbf{v}}_{-\mu} = -\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\circ(\partial_{\mu=\lambda}\,\mathbf{Fl}^{\mathbf{v}}_{\mu})\circ\mathbf{Fl}^{\mathbf{v}}_{-\lambda}\\ &\Longrightarrow\ \partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{-\lambda} = -\partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{\lambda} = -\mathbf{v}\,.\end{aligned}$$

Proposition 1.3.5 reveals that, by exchanging the roles of the involved vector fields, the Lie derivative just changes its sign. This basic property is put into evidence by adopting the bracket notation of the commutators.

Then any property concerning one of the vector fields immediately holds also for the other one. From Proposition 1.3.5 it follows that the Lie derivative is a field of vector valued two-form $\mathcal{L} \in \mathrm{C}^k(\mathbb{M}\,;\Lambda(\mathbb{TM}^2\,;\mathbb{TM}))$. The lack of symmetry of the iterated directional derivative along two vector fields is strictly related to the lack of commutativity of the corresponding flows. Indeed the Lie derivative (and hence the commutator) vanishes if and only if the flows of the vector fields commute. The next proposition provides another proof that the Lie derivative is a Lie bracket.

Proposition 1.3.6 (Lie bracket) *The* Lie *derivative* $\mathcal{L}_{\mathbf{v}}\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *of a vector field* $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *along a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is equal to the* Lie *bracket* $[\mathbf{v}\,,\mathbf{u}] \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, *defined by*

$$\begin{aligned}(\mathcal{L}_{\mathbf{v}}\mathbf{u})\,f = [\mathbf{v}\,,\mathbf{u}]\,f &:= (\mathcal{L}_{\mathbf{v}}\circ\mathcal{L}_{\mathbf{u}} - \mathcal{L}_{\mathbf{u}}\circ\mathcal{L}_{\mathbf{v}})\,f\\ &= (\mathbf{vu}-\mathbf{uv})\,f\,,\quad \forall\, f \in \mathrm{C}^1(\mathbb{M}\,;\Re)\,.\end{aligned}$$

Proof. As for the diagram in Lemma 1.2.10 on page 57 the fibre linearity of the tangent map $Tf \in C^1(\mathbb{TM};\mathbb{T}\Re)$ ensures the commutativity of the following diagrams (the latter is the $\partial_{t=0}$ derivative of the former):

$$\begin{array}{ccccc} \mathbb{T}\Re & \xleftarrow{\mathbf{aff}^t_{\mathbb{T}\Re}} & \mathbb{T}\Re \times_\Re \mathbb{T}\Re & \xrightarrow{\mathbf{Vl}_{\mathbb{T}\Re}} & \mathbb{TT}\Re \\ {\scriptstyle Tf}\uparrow & & \uparrow{\scriptstyle Tf} & & \uparrow{\scriptstyle T^2 f} \\ \mathbb{TM} & \xleftarrow{\mathbf{aff}^t_{\mathbb{TM}}} & \mathbb{TM} \times_{\mathbb{M}} \mathbb{TM} & \xrightarrow{\mathbf{Vl}_{\mathbb{TM}}} & \mathbb{TTM} \end{array} \quad \text{with} \quad \begin{cases} Tf \circ \mathbf{aff}^t_{\mathbb{TM}} = \mathbf{aff}^t_{\mathbb{T}\Re} \circ Tf \,, \\ T^2 f \circ \mathbf{Vl}_{\mathbb{TM}} = \mathbf{Vl}_{\mathbb{T}\Re} \circ Tf \,. \end{cases}$$

Then, being

$$T^2 f \circ \mathbf{Vl}_{\mathbb{TM}} \circ (\mathbf{u}\,,[\mathbf{v}\,,\mathbf{u}]) = \mathbf{Vl}_{\mathbb{T}\Re} \circ Tf \circ (\mathbf{u}\,,[\mathbf{v}\,,\mathbf{u}])\,,$$

we have to prove that

$$\mathbf{Vl}_{\mathbb{T}\Re} \circ Tf \circ (\mathbf{u}\,,[\mathbf{v}\,,\mathbf{u}]) = T^2 f \circ \mathbf{Vl}_{\mathbb{TM}} \circ (\mathbf{u}\,,\mathcal{L}_{\mathbf{v}}\mathbf{u})\,.$$

To this end we recall that by Lemma 1.2.5 we have: $T^2 f \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} = \mathbf{k}_{\mathbb{TT}\Re} \circ T^2 f$ and that by Lemma 1.2.8 it is: $\mathbf{k}_{\mathbb{TT}\Re} \circ T^2 f = T^2 f$. Hence

$$\begin{aligned} T^2 f \circ \mathbf{Vl}_{\mathbb{TM}} \circ (\mathbf{u}\,,\mathcal{L}_{\mathbf{v}}\mathbf{u}) &= T^2 f \circ (T\mathbf{u} \circ \mathbf{v} - \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{u}) \\ &= T^2 f \circ T\mathbf{u} \circ \mathbf{v} - \mathbf{k}_{\mathbb{TT}\Re} \circ T^2 f \circ T\mathbf{v} \circ \mathbf{u} \\ &= T^2 f \circ T\mathbf{u} \circ \mathbf{v} - T^2 f \circ T\mathbf{v} \circ \mathbf{u} \\ &= \mathbf{Vl}_{\mathbb{T}\Re} \circ (\mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}}\, f - \mathcal{L}_{\mathbf{u}}\,\mathcal{L}_{\mathbf{v}}\, f) \\ &= \mathbf{Vl}_{\mathbb{T}\Re} \circ Tf \circ (\mathcal{L}_{\mathbf{v}}\,\mathcal{L}_{\mathbf{u}} - \mathcal{L}_{\mathbf{u}}\,\mathcal{L}_{\mathbf{v}}) \\ &= \mathbf{Vl}_{\mathbb{T}\Re} \circ Tf \circ (\mathbf{u}\,,[\mathbf{v}\,,\mathbf{u}]) \\ &= T^2 f \circ \mathbf{Vl}_{\mathbb{TM}} \circ (\mathbf{u}\,,[\mathbf{v}\,,\mathbf{u}])\,, \end{aligned}$$

which implies that $\mathcal{L}_{\mathbf{v}}\mathbf{u} = [\mathbf{v}\,,\mathbf{u}]$. ■

Let us now provide the proof of a basic result which generalizes to relatedness the second formula in Proposition 1.3.4.

Lemma 1.3.4 (Lie bracket of morphism-related vector fields) *Let the vector fields* $\mathbf{X}, \mathbf{Y} \in C^1(\mathbb{N};\mathbb{TN})$ *be related to the fields* $\mathbf{u}, \mathbf{v} \in C^1(\mathbb{M};\mathbb{TM})$ *by a* *morphism* $\varphi \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *, according to the commutative diagram*

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\varphi} & \mathbb{TN} \\ {\scriptstyle \mathbf{u},\mathbf{v}}\uparrow & & \uparrow{\scriptstyle \mathbf{X},\mathbf{Y}} \\ \mathbb{M} & \xrightarrow{\varphi} & \mathbb{N} \end{array} \quad \Longleftrightarrow \quad \begin{cases} \mathbf{X} \circ \varphi = T\varphi \circ \mathbf{u} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TN})\,, \\ \mathbf{Y} \circ \varphi = T\varphi \circ \mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TN})\,. \end{cases}$$

Then also their LIE *brackets are* φ*-related:*

$$[\mathbf{X}\,,\mathbf{Y}] \circ \varphi = T\varphi \circ [\mathbf{u}\,,\mathbf{v}]\,.$$

Given a morphism $\varphi \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *, setting* $T_{\mathbf{v}} \circ \varphi := T\varphi \circ \mathbf{v}$ *, we have that*

$$T\varphi \circ [\mathbf{u}\,,\mathbf{v}] = T_{[\mathbf{u}\,,\mathbf{v}]} \circ \varphi\,,$$

and the result may be stated in the form:

$$[T_{\mathbf{u}}\,,T_{\mathbf{v}}] = T_{[\mathbf{u}\,,\mathbf{v}]}\,.$$

Proof. By Proposition 1.2.5 we have that

$$\varphi \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} = \mathbf{Fl}^{\mathbf{X}}_{\lambda} \circ \varphi \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})\,,$$

and then, applying the tangent functor, also

$$T\varphi \circ T\mathbf{Fl}^{\mathbf{u}}_{\lambda} = T\mathbf{Fl}^{\mathbf{X}}_{\lambda} \circ T\varphi \in \mathrm{C}^0(\mathbb{TM}\,;\mathbb{TN})\,.$$

Moreover $T^2\varphi \circ \mathbf{Vl}_{\mathbb{TM}} = \mathbf{Vl}_{\mathbb{TN}} \circ T\varphi$ by the commutative diagram in Lemma 1.2.10 on page 57. The following equalities thus hold

$$\begin{aligned} \mathbf{Vl}_{\mathbb{TN}} \circ (\mathbf{Y}\,,[\mathbf{X}\,,\mathbf{Y}]) \circ \varphi &:= \partial_{\lambda=0}\, T\mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{Y} \circ \mathbf{Fl}^{\mathbf{X}}_{\lambda} \circ \varphi \\ &= \partial_{\lambda=0}\, T\mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{Y} \circ \varphi \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} \\ &= \partial_{\lambda=0}\, T\mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ T\varphi \circ \mathbf{v} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} \\ &= \partial_{\lambda=0}\, T\varphi \circ T\mathbf{Fl}^{\mathbf{u}}_{-\lambda} \circ \mathbf{v} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} \\ &= T^2\varphi \circ \partial_{\lambda=0}\, (T\mathbf{Fl}^{\mathbf{u}}_{-\lambda} \circ \mathbf{v} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda}) \\ &= T^2\varphi \circ \mathbf{Vl}_{\mathbb{TM}} \circ (\mathbf{v}\,,[\mathbf{u}\,,\mathbf{v}]) \\ &= \mathbf{Vl}_{\mathbb{TN}} \circ T\varphi \circ (\mathbf{v}\,,[\mathbf{u}\,,\mathbf{v}]) \\ &= \mathbf{Vl}_{\mathbb{TN}} \circ (T\varphi \circ \mathbf{v}\,, T\varphi \circ [\mathbf{u}\,,\mathbf{v}])\,. \end{aligned}$$

The result then follows by observing that

$$\begin{aligned}\mathbf{Vl}_{\mathbb{TN}} \circ (\mathbf{Y}\,,[\mathbf{X}\,,\mathbf{Y}]) \circ \boldsymbol{\varphi} &= \mathbf{Vl}_{\mathbb{TN}} \circ (\mathbf{Y} \circ \boldsymbol{\varphi}\,,[\mathbf{X}\,,\mathbf{Y}] \circ \boldsymbol{\varphi}) \\ &= \mathbf{Vl}_{\mathbb{TN}} \circ (T\boldsymbol{\varphi} \circ \mathbf{v}\,,[\mathbf{X}\,,\mathbf{Y}] \circ \boldsymbol{\varphi})\,.\end{aligned}$$

By comparing the last terms of the two equalities above, the $\boldsymbol{\varphi}$-relatedness of the Lie brackets follows: $[\mathbf{X}\,,\mathbf{Y}] \circ \boldsymbol{\varphi} = T\boldsymbol{\varphi} \circ [\mathbf{u}\,,\mathbf{v}]$. ■

1.3.2 Lie algebra and Jacobi's identity

The Lie derivative $\mathcal{L}_{\mathbf{v}}\mathbf{u}$ is apparently $\Re$-linear in the vector field $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$. Indeed for any $\mathbf{u}_1, \mathbf{u}_2 \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ and any $\alpha \in \Re$, we have that

$$\begin{aligned}&i) \quad \mathcal{L}_{\mathbf{v}}(\mathbf{u}_1 + \mathbf{u}_2) = \mathcal{L}_{\mathbf{v}}(\mathbf{u}_1) + \mathcal{L}_{\mathbf{v}}(\mathbf{u}_2)\,,\\ &ii) \quad \mathcal{L}_{\mathbf{v}}(\alpha\mathbf{u}) = \alpha\,\mathcal{L}_{\mathbf{v}}(\mathbf{u})\,.\end{aligned}$$

By Proposition 1.3.5 we infer that the Lie derivative $\mathcal{L}_{\mathbf{v}}\,\mathbf{u}$ is $\Re$-linear in $\mathbf{v}$, so that, for any $\mathbf{v}_1, \mathbf{v}_2 \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ and any $\alpha \in \Re$, we have that

$$\begin{aligned}&i) \quad \mathcal{L}_{(\mathbf{v}_1+\mathbf{v}_2)}\mathbf{u} = \mathcal{L}_{\mathbf{v}_1}(\mathbf{u}) + \mathcal{L}_{\mathbf{v}_2}(\mathbf{u})\,,\\ &ii) \quad \mathcal{L}_{(\alpha\mathbf{v})}(\mathbf{u}) = \alpha\,\mathcal{L}_{\mathbf{v}}(\mathbf{u})\,.\end{aligned}$$

It is easy to verify that properties *i*) and *ii*) still hold for the Lie derivative of a tensor field. A result more general than *ii*) will be proven in Proposition 1.3.11. Due to the antisymmetry of the commutator we have that

$$\mathcal{L}_{\mathbf{v}}\,\mathbf{v} = [\mathbf{v}\,,\mathbf{v}] = 0\,,$$

a result which follows also form the definition of the Lie derivative since a vector field is dragged by its flow.

The Lie bracket defines, in the real Banach space of indefinitely continuously differentiable vector fields $\mathbf{u} \in \mathrm{C}^\infty(\mathbb{M}\,;\mathbb{TM})$, a Lie *algebra* which enjoys the following properties:

$$\begin{aligned}&i) \quad [-\,,-] \quad \text{is } \Re\text{-bilinear}\,,\\ &ii) \quad [\mathbf{u}\,,\mathbf{u}] = 0\,, \quad \forall\,\mathbf{u} \in \mathrm{C}^\infty(\mathbb{M}\,;\mathbb{TM})\,,\\ &iii) \quad [\mathbf{v}\,,[\mathbf{u}\,,\mathbf{w}]] + [\mathbf{w}\,,[\mathbf{v}\,,\mathbf{u}]] + [\mathbf{u}\,,[\mathbf{w}\,,\mathbf{v}]] = 0\,, \quad \forall\,\mathbf{v},\mathbf{u},\mathbf{w} \in \mathrm{C}^\infty(\mathbb{M}\,;\mathbb{TM})\,.\end{aligned}$$

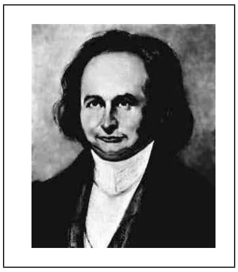

Figure 1.16: Carl Gustav Jacob Jacobi (1804 - 1851)

The identity *iii*) is named JACOBI*'s identity.* It can be proven by a direct computation based on the observation that

$$[\mathbf{v}\,,[\mathbf{u}\,,\mathbf{w}]] = \mathbf{v}\,[\mathbf{u}\,,\mathbf{w}] - [\mathbf{u}\,,\mathbf{w}]\,\mathbf{v} = \mathbf{v}\,\mathbf{u}\,\mathbf{w} - \mathbf{v}\,\mathbf{w}\,\mathbf{u} - \mathbf{u}\,\mathbf{w}\,\mathbf{v} + \mathbf{w}\,\mathbf{u}\,\mathbf{v}\,.$$

Then, summing up the twelve terms resulting from a cyclic permutation, we recognize that each one appers twice with opposite signs. The LIE bracket is also called the LIE *commutator*. JACOBI's identity, rewritten as

$$[[\mathbf{u}\,,\mathbf{w}]\,,\mathbf{v}] = [\mathbf{u}\,,[\mathbf{w}\,,\mathbf{v}]] - [\mathbf{w}\,,[\mathbf{u}\,,\mathbf{v}]]\,,$$

gives:

$$\mathcal{L}_{[\mathbf{u}\,,\mathbf{w}]}\,\mathbf{v} = (\mathcal{L}_{\mathbf{u}}\mathcal{L}_{\mathbf{w}} - \mathcal{L}_{\mathbf{w}}\mathcal{L}_{\mathbf{u}})\,\mathbf{v} = [\mathcal{L}_{\mathbf{u}}\,,\mathcal{L}_{\mathbf{w}}]\,\mathbf{v}\,, \quad \forall\,\mathbf{v} \in \mathrm{C}^2(\mathbb{M}\,;\mathbb{TM})\,.$$

Toghether with the formula provided in Proposition 1.3.5

$$\mathcal{L}_{[\mathbf{u}\,,\mathbf{w}]}\,f := (\mathcal{L}_{\mathbf{u}}\,\mathcal{L}_{\mathbf{w}} - \mathcal{L}_{\mathbf{w}}\,\mathcal{L}_{\mathbf{u}})\,f = [\mathcal{L}_{\mathbf{u}}\,,\mathcal{L}_{\mathbf{w}}]\,f\,, \quad \forall\,f \in \mathrm{C}^2(\mathbb{M}\,;\Re)\,.$$

it implies an analogous formula for the LIE derivative of any tensor field $\boldsymbol{\alpha}$:

$$\mathcal{L}_{[\mathbf{u}\,,\mathbf{w}]}\,\boldsymbol{\alpha} = (\mathcal{L}_{\mathbf{u}}\mathcal{L}_{\mathbf{w}} - \mathcal{L}_{\mathbf{w}}\mathcal{L}_{\mathbf{u}})\,\boldsymbol{\alpha} = [\mathcal{L}_{\mathbf{u}}\,,\mathcal{L}_{\mathbf{w}}]\,\boldsymbol{\alpha}\,.$$

Proposition 1.3.7 *For any vector field* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *, the* LIE *derivative* $\mathcal{L}_{\mathbf{v}}$ *is a derivation. Indeed*

$$\mathcal{L}_{\mathbf{v}}(f\,\mathbf{u}) = f\,\mathcal{L}_{\mathbf{v}}\mathbf{u} + (\mathcal{L}_{\mathbf{v}}f)\,\mathbf{u}\,,$$

for any $f \in \mathrm{C}^1(\mathbb{M}\,;\Re)$ *and any vector field* $\mathbf{u} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *.*

Proof. By LEIBNIZ rule of differentiation we have that

$$\mathcal{L}_{\mathbf{v}}(f\,\mathbf{u}) = \partial_{\lambda=0}\ \mathbf{Fl}^{\mathbf{v}}_{\lambda}{\downarrow}(f\,\mathbf{u}) = \partial_{\lambda=0}\ \mathbf{Fl}^{\mathbf{v}}_{\lambda}{\downarrow}f\,\mathbf{Fl}^{\mathbf{v}}_{\lambda}{\downarrow}\mathbf{u} = f\,\mathcal{L}_{\mathbf{v}}\mathbf{u} + (\mathcal{L}_{\mathbf{v}}f)\,\mathbf{u}\,,$$

and hence the result. By rewriting the JACOBI identity as

$$[\mathbf{v}\,,[\mathbf{u}\,,\mathbf{w}]] = [[\mathbf{v}\,,\mathbf{u}]\,,\mathbf{w}] + [\mathbf{u}\,,[\mathbf{v}\,,\mathbf{w}]]\,,$$

we see that the adjoint field $\mathrm{ADJ}_{\mathbf{v}} := [\mathbf{v}\,,\cdot\,]$ is a LIE-algebra derivation since

$$\mathrm{ADJ}_{\mathbf{v}}([\mathbf{u}\,,\mathbf{w}]) = [\mathrm{ADJ}_{\mathbf{v}}(\mathbf{u})\,,\mathbf{w}] + [\mathbf{u}\,,\mathrm{ADJ}_{\mathbf{v}}(\mathbf{w})]\,.$$

1.3.3 Frames and coordinate systems

Frames and coordinate systems are respectively named *repére mobile* and *repére naturel* in the french literature [27].

- A *local frame* on an n-dimensional manifold $\mathbb{M}$ is a set of n vector fields $\mathbf{E}_i : \mathbb{M} \mapsto \mathbb{TM}$, $i = 1, \dots, n$ whose values $\mathbf{E}_i(\mathbf{x})$ at any point $\mathbf{x} \in U_{\mathbb{M}}$ of a neighborhood $U_{\mathbb{M}} \subset \mathbb{M}$ form a basis of the tangent space $\mathbb{TM}(\mathbf{x})$.
- A *local coordinate system* on a n-dimensional manifold $\mathbb{M}$ is a local frame whose vector fields $\mathbf{E}_i \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, $i = 1, \dots, n$ are the velocities of a local coordinate map $\varphi \in \mathrm{C}^1(U_E\,;U_{\mathbb{M}})$.

From Proposition 1.3.2 on page 63, we infer the following results.

Proposition 1.3.8 *A local frame is a local coordinate system if and only if the* LIE *bracket of any pair of vector fields of the local frame vanishes:*

$$[\mathbf{E}_i\,,\mathbf{E}_j] = 0\,, \quad \forall\, i, j \in \{1, \dots, n\}\,.$$

The commutativity of the flows ensures that to any point $\mathbf{x} \in \mathbb{M}$ there corresponds a unique set of coordinates.

This property is preserved under a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ since the transformation rule of a flow $\boldsymbol{\varphi}{\uparrow}\boldsymbol{\chi} = \boldsymbol{\varphi} \circ \boldsymbol{\chi} \circ \boldsymbol{\varphi}^{-1}$ pushed by $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ implies that

$$\boldsymbol{\varphi}{\uparrow}\boldsymbol{\chi} \circ \boldsymbol{\varphi}{\uparrow}\boldsymbol{\psi} = \boldsymbol{\varphi} \circ \boldsymbol{\chi} \circ \boldsymbol{\varphi}^{-1} \circ \boldsymbol{\varphi} \circ \boldsymbol{\psi} \circ \boldsymbol{\varphi}^{-1} = \boldsymbol{\varphi} \circ \boldsymbol{\chi} \circ \boldsymbol{\psi} \circ \boldsymbol{\varphi}^{-1}\,,$$

and hence commutativity interchanges with the push:

$$\boldsymbol{\varphi}{\uparrow}\boldsymbol{\chi} \circ \boldsymbol{\varphi}{\uparrow}\boldsymbol{\psi} = \boldsymbol{\varphi}{\uparrow}\boldsymbol{\psi} \circ \boldsymbol{\varphi}{\uparrow}\boldsymbol{\chi} \iff \boldsymbol{\chi} \circ \boldsymbol{\psi} = \boldsymbol{\psi} \circ \boldsymbol{\chi}\,.$$

The next simple corollary of Proposition 1.3.8 will be recalled in sections 1.9.1 and 1.9.6 in defining the coordinate components of the torsion tensor and of the RIEMANN-CHRISTOFFEL curvature tensor.

Proposition 1.3.9 *Any* n*-tuple of tangent vectors* $\mathbf{v}_i(\mathbf{x}) \in \mathbb{TM}(\mathbf{x})$ *with* $i = 1, \ldots, n$ *can be extended to an* n*-tuple of tangent vector fields* $\mathbf{v}_i : \mathbb{M} \mapsto \mathbb{TM}$ *such that*

$$[\mathbf{v}_i\,, \mathbf{v}_j] = 0\,, \quad i, j = 1, \ldots, n\,.$$

Proof. It is sufficient to observe that an n-tuple of tangent vectors $\mathbf{v}_i(\mathbf{x}) \in \mathbb{TM}(\mathbf{x})$ is diffeomorphically transformed by a local chart into an n-tuple of vectors $(\varphi{\uparrow}\mathbf{v}_i)(\mathbf{x}) \in E$ where E is the model space of $\mathbb{M}$.

Then each vector $(\varphi{\uparrow}\mathbf{v}_i)(\mathbf{x}) \in E$ defines a straight line-flow thru any point in the linear space E and these commuting flows are mapped by the inverse local charts to flow on $\mathbb{M}$ which still commute. ∎

Given two vector fields $X, Y \in \mathrm{C}^1(E\,; E)$ in a BANACH space E, the second directional derivative of a scalar field $f \in \mathrm{C}^2(E\,; \Re)$ is the twice covariant tensor field defined according to the LEIBNIZ formula

$$\partial^2_{XY}\, f := \partial_X\, \partial_Y\, f - \partial_{(\partial_X\, Y)}\, f\,.$$

Tensoriality is easily verified by the criterion provided by Proposition 1.2.1.

If the vector fields $X, Y \in \mathrm{C}^1(E\,; E)$ are constant in E, the derivatives $\partial_X\, Y$ and $\partial_Y\, X$ vanish identically and the LIE bracket $[X\,, Y]$ vanishes too since the flows of constant vector fields commute. This in accordance with the fact that the second directional derivative of any scalar field is symmetric:

$$\partial^2_{XY}\, f = \partial^2_{YX}\, f\,.$$

For arbitrary vector fields $X, Y \in \mathrm{C}^1(E\,; E)$ we have that

$$[X\,, Y]\, f := \partial_X\, \partial_Y\, f - \partial_Y\, \partial_X\, f = \partial_{(\partial_X Y)}\, f - \partial_{(\partial_Y X)}\, f\,,$$

and the LIE bracket takes the expression: $[X\,, Y] = \partial_X Y - \partial_Y X$.

We may then state that

Proposition 1.3.10 *In terms of local coordinates* $\{\mathcal{U}\,, \varphi\}$ *the* LIE *bracket* $[\mathbf{v}\,, \mathbf{u}]$ *may be written as*

$$[\mathbf{v}\,, \mathbf{u}] = (Y^i_{/j}\, X^j - X^i_{/j}\, Y^j)\, \mathbf{E}_i\,.$$

Proof. Denoting by $X = \varphi{\uparrow}\mathbf{u}$ and $Y = \varphi{\uparrow}\mathbf{v}$ the expression of $\mathbf{u}$ and $\mathbf{v}$ in coordinates, by Proposition 1.3.4 we have that:

$$\begin{aligned}[\mathbf{v}\,,\mathbf{u}] &= \varphi{\downarrow}[Y\,,X] = \varphi{\downarrow}(\partial_Y X - \partial_X Y)\\ &= (X^i_{/j}\,Y^j - Y^i_{/j}\,X^j)\,\varphi{\downarrow}\mathbf{e}_i = (X^i_{/j}\,Y^j - Y^i_{/j}\,X^j)\,\mathbf{E}_i\,.\end{aligned}$$

where $\varphi{\downarrow}\mathbf{e}_i = \mathbf{E}_i$, i.e. $\mathbf{e}_i$ is the base vector, in the model space E, corresponding to the coordinate vector $\mathbf{E}_i$. ■

1.3.4 Properties of the Lie derivative

Proposition 1.3.11 *The* Lie *derivative fulfills the relations*

$$\begin{aligned}
&i) && \mathcal{L}_{(g\,\mathbf{v})}\,f = g\,\mathcal{L}_{\mathbf{v}}f\,,\\
&ii) && \mathcal{L}_{(f\,\mathbf{v})}\,\mathbf{u} = -[\mathbf{u}, f\,\mathbf{v}\,,=]f\,[\mathbf{v}\,,\mathbf{u}] - (\mathcal{L}_{\mathbf{u}}f)\,\mathbf{v} = f\,\mathcal{L}_{\mathbf{v}}\,\mathbf{u} - (\mathcal{L}_{\mathbf{u}}f)\,\mathbf{v}\,,\\
&iii) && [f\,\mathbf{v}, g\,\mathbf{u}\,,=]f\,g\,[\mathbf{v}\,,\mathbf{u}] + (\mathcal{L}_{\mathbf{v}}g)\,f\,\mathbf{u} - (\mathcal{L}_{\mathbf{u}}f)\,g\,\mathbf{v}\,,\\
&iv) && \mathcal{L}_{(f\,\mathbf{v})}\,\boldsymbol{\alpha} = f\,\mathcal{L}_{\mathbf{v}}\boldsymbol{\alpha} + (df \star \boldsymbol{\alpha})\mathbf{v}\,,\\
&v) && (\mathcal{L}_{\mathbf{v}}f)\,\boldsymbol{\mu} = df \wedge \boldsymbol{\mu}\mathbf{v}\,,\\
&vi) && \mathcal{L}_{(f\,\mathbf{v})}\,\boldsymbol{\mu} = \mathcal{L}_{\mathbf{v}}(f\,\boldsymbol{\mu})\,,\\
&vii) && \mathcal{L}_{\mathbf{v}}(df) = d(\mathcal{L}_{\mathbf{v}}f)\,,\\
&viii) && \mathcal{L}_{\mathbf{v}}(\boldsymbol{\alpha}\otimes\boldsymbol{\beta}) = (\mathcal{L}_{\mathbf{v}}\boldsymbol{\alpha}\otimes\boldsymbol{\beta}) + (\boldsymbol{\alpha}\otimes\mathcal{L}_{\mathbf{v}}\boldsymbol{\beta})\,,\\
&ix) && \mathcal{L}_{\mathbf{u}}\,(\boldsymbol{\alpha}\mathbf{v}) = (\mathcal{L}_{\mathbf{u}}\boldsymbol{\alpha})\mathbf{v} + \boldsymbol{\alpha}\,(\mathcal{L}_{\mathbf{u}}\mathbf{v})\,,\\
&x) && [\mathcal{L}_{\mathbf{u}}\,,\mathbf{i}_{\mathbf{v}}] := \mathcal{L}_{\mathbf{u}}\circ\mathbf{i}_{\mathbf{v}} - \mathbf{i}_{\mathbf{v}}\circ\mathcal{L}_{\mathbf{u}} = \mathbf{i}_{[\mathbf{u}\,,\mathbf{v}]}\,,\quad \mathcal{L}_{\mathbf{v}}\circ\mathbf{i}_{\mathbf{v}} = \mathbf{i}_{\mathbf{v}}\circ\mathcal{L}_{\mathbf{v}}\,,\\
&xi) && \mathcal{L}_{[\mathbf{u}\,,\mathbf{v}]} = [\mathcal{L}_{\mathbf{u}}\,,\mathcal{L}_{\mathbf{v}}] = \mathcal{L}_{\mathbf{u}}\circ\mathcal{L}_{\mathbf{v}} - \mathcal{L}_{\mathbf{v}}\circ\mathcal{L}_{\mathbf{u}}\,,
\end{aligned}$$

where $f\,,g \in \mathrm{C}^1(\mathbb{M}\,;\Re)$ *and* $\mathbf{v},\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *are scalar and vector fields,* $\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM}^k)$ *are tensor fields and* $\boldsymbol{\mu} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM}^{(\dim\mathbb{M})})$ *is a volume form. The operator* $\star$ *is defined in the proof of formula* $iv)$.

Proof. Property $i)$ follows from the definition of partial derivative. Property $ii)$ follows from Propositions 1.3.5 and 1.3.7. Indeed by virtue of the antisymmetry of the Lie bracket we infer that

$$\begin{aligned}[f\,\mathbf{v}\,,\mathbf{u}] = -[\mathbf{u}\,,f\,\mathbf{v}] = -\mathcal{L}_{\mathbf{u}}(f\,\mathbf{v}) &= -f\,\mathcal{L}_{\mathbf{u}}(\mathbf{v}) - (\mathcal{L}_{\mathbf{u}}f)\,\mathbf{v}\\ &= f\,[\mathbf{v}\,,\mathbf{u}] - (\mathcal{L}_{\mathbf{u}}f)\,\mathbf{v}\,.\end{aligned}$$

From $i)$ and $ii)$ we obtain formula $iii)$ as follows

$$\begin{aligned}[f\,\mathbf{v}\,,g\,\mathbf{u}] &= f\,[\mathbf{v}\,,g\,\mathbf{u}] - (\mathcal{L}_{(g\,\mathbf{u})}f)\,\mathbf{v} = f\,[\mathbf{v}\,,g\,\mathbf{u}] - (\mathcal{L}_{\mathbf{u}}f)\,g\,\mathbf{v} \\ &= -f\,[g\,\mathbf{u}\,,\mathbf{v}] - (\mathcal{L}_{\mathbf{u}}f)\,g\,\mathbf{v} \\ &= f\,g\,[\mathbf{v}\,,\mathbf{u}] + (\mathcal{L}_{\mathbf{v}}g)\,f\,\mathbf{u} - (\mathcal{L}_{\mathbf{u}}f)\,g\,\mathbf{v}\,.\end{aligned}$$

Formula $iv)$ is inferred from property $i)$, the LEIBNIZ rule of Proposition 1.3.3 and Proposition 1.3.5 as follows.

$$\begin{aligned}&(\mathcal{L}_{(f\,\mathbf{v})}\,\boldsymbol{\alpha})\,(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k) = \\ &\qquad = \mathcal{L}_{(f\,\mathbf{v})}\,(\boldsymbol{\alpha}(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k)) - \sum_{i=1}^{k}\boldsymbol{\alpha}(\mathbf{v}_1,\dots\mathcal{L}_{(f\,\mathbf{v})}\mathbf{v}_i,\dots\mathbf{v}_k) \\ &\qquad = f\,\mathcal{L}_{\mathbf{v}}\,(\boldsymbol{\alpha}(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k)) + \sum_{i=1}^{k}\boldsymbol{\alpha}(\mathbf{v}_1,\dots\mathcal{L}_{\mathbf{v}_i}(f\,\mathbf{v}),\dots\mathbf{v}_k)\,.\end{aligned}$$

Then we observe that

$$\begin{aligned}&\boldsymbol{\alpha}(\mathbf{v}_1,\dots\mathcal{L}_{\mathbf{v}_i}\,(f\,\mathbf{v}),\dots\mathbf{v}_k) = \\ &\qquad = \boldsymbol{\alpha}(\mathbf{v}_1,\dots(\mathcal{L}_{\mathbf{v}_i}f)\,\mathbf{v},\dots\mathbf{v}_k) + \boldsymbol{\alpha}(\mathbf{v}_1,\dots f\,\mathcal{L}_{\mathbf{v}_i}\mathbf{v},\dots\mathbf{v}_k) \\ &\qquad = \boldsymbol{\alpha}(\mathbf{v}_1,\dots(\mathcal{L}_{\mathbf{v}_i}f)\,\mathbf{v},\dots\mathbf{v}_k) - \boldsymbol{\alpha}(\mathbf{v}_1,\dots f\,\mathcal{L}_{\mathbf{v}}\mathbf{v}_i,\dots\mathbf{v}_k)\,,\end{aligned}$$

and that

$$\begin{aligned}f\,(\mathcal{L}_{\mathbf{v}}\,(\boldsymbol{\alpha}(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k)) - \sum_{i=1}^{k}\boldsymbol{\alpha}(\mathbf{v}_1,\dots\mathcal{L}_{\mathbf{v}}\mathbf{v}_i,\dots\mathbf{v}_k)) = \\ = f\,(\mathcal{L}_{\mathbf{v}}\,\boldsymbol{\alpha})(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k))\,.\end{aligned}$$

Then

$$\begin{aligned}&(\mathcal{L}_{(f\,\mathbf{v})}\,\boldsymbol{\alpha})(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k) = \\ &\qquad = f\,(\mathcal{L}_{\mathbf{v}}\,\boldsymbol{\alpha})(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k)) + \sum_{i=1}^{k}(\mathcal{L}_{\mathbf{v}_i}f)\,\boldsymbol{\alpha}(\mathbf{v}_1,\dots,\mathbf{v},\dots\mathbf{v}_k) \\ &\qquad = (f\,\mathcal{L}_{\mathbf{v}}\,\boldsymbol{\alpha} + (df\star\boldsymbol{\alpha})\mathbf{v})(\mathbf{v}_1,\dots\mathbf{v}_i,\dots\mathbf{v}_k)\,,\end{aligned}$$

where, denoting by $\mathbb{A}_{1i}$ the operator which exchanges the first and the i-th element of a list, the $\star$ operation is defined by

$$(df\star\boldsymbol{\alpha})\mathbf{v} := \sum_{i=1}^{k}((df\otimes\boldsymbol{\alpha})\circ\mathbb{A}_{1i})\,\mathbf{v}\,.$$

A simpler proof of property *iv)* for k-forms is based on the homotopy formula and will be given in section 1.6.10.
To get formula *v)*, let $\{\mathbf{v}_1, \dots \mathbf{v}_i, \dots \mathbf{v}_n\}$ be a frame and $\{\mathbf{v}^1, \dots \mathbf{v}^i, \dots \mathbf{v}^n\}$ be the dual co-frame. Then $\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}^i, \mathbf{v} \rangle \, \mathbf{v}_i$ so that:

$$
\begin{aligned}
&(df \wedge \boldsymbol{\mu}\mathbf{v})(\mathbf{v}_1, \dots \mathbf{v}_i, \dots \mathbf{v}_n) = \\
&= \sum_{i=1}^{n} (\mathcal{L}_{\mathbf{v}_i} f) \, \boldsymbol{\mu}(\mathbf{v}_1, \dots, \mathbf{v}, \dots \mathbf{v}_n) = \sum_{i=1}^{n} (\mathcal{L}_{\mathbf{v}_i} f) \, \langle \mathbf{v}^i, \mathbf{v} \rangle \, \boldsymbol{\mu}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots \mathbf{v}_n) \\
&= (\mathcal{L}_{\mathbf{v}} f) \, \boldsymbol{\mu}(\mathbf{v}_1, \dots \mathbf{v}_i, \dots \mathbf{v}_n) \,.
\end{aligned}
$$

Formula *vi)* then follows from formulas *iv)* and *v)*:

$$
\mathcal{L}_{(f\,\mathbf{v})} \, \boldsymbol{\mu} = f \, \mathcal{L}_{\mathbf{v}} \, \boldsymbol{\mu} + df \wedge \boldsymbol{\mu}\mathbf{v} = f \, \mathcal{L}_{\mathbf{v}} \, \boldsymbol{\mu} + (d_{\mathbf{v}} f) \, \boldsymbol{\mu} = \mathcal{L}_{\mathbf{v}} \, (f \, \boldsymbol{\mu}) \,.
$$

Formula *vii)*, stating that the Lie derivative commutes with the derivation of a scalar function, can be inferred from Proposition 1.2.4 by exchanging the order of derivation with respect to time and position, indeed:

$$
\mathcal{L}_{\mathbf{v}}(df) = \partial_{\lambda=0} \, \varphi_{\lambda}{\downarrow}(df) = \partial_{\lambda=0} \, d(\varphi_{\lambda}{\downarrow} f) = d(\partial_{\lambda=0} \, \varphi_{\lambda}{\downarrow} f) = d(\mathcal{L}_{\mathbf{v}} f) \,.
$$

This result is a special case of a more general property, concerning the commutativity between the Lie derivative and the exterior derivative, that will be proved in Proposition 1.6.2.

Formulas *viii)* and *ix)* can be inferred from remark 1.2.2, by exchanging the pull-back with the contraction and with the tensor product, and applying Leibniz rule.

In *x)* the first formula is just a rewriting of *ix)* and the second is a simple corollary of the first formula, since $\mathcal{L}_{\mathbf{v}} \, \mathbf{v} = 0$.

Formula *xi)* for scalar fields follows by the definition of Lie bracket, for vector fields follows from Jacobi's identity and for general tensor field by Leibniz rule (see section 1.3.2). ■

Remark 1.3.1 *The linear isomorphism* $\mathbf{g} \in BL(\mathbb{TM}; \mathbb{T}^*\mathbb{M})$ *induced by a metric tensor field doesn't commute in general with the convective derivative, since by* Leibniz *rule:*

$$
\mathcal{L}_{\mathbf{v}}(\mathbf{gu}) = (\mathcal{L}_{\mathbf{v}}\mathbf{g})\mathbf{u} + \mathbf{g}(\mathcal{L}_{\mathbf{v}}\mathbf{u}) \,.
$$

However from the formula above we infer that

$$
\mathcal{L}_{\mathbf{v}}\mathbf{g} = 0 \iff \mathcal{L}_{\mathbf{v}} \circ \mathbf{g} = \mathbf{g} \circ \mathcal{L}_{\mathbf{v}} \,.
$$

1.3.5 Method of characteristics

The properties of the Lie derivative enunciated in Propositions 1.3.1 and 1.3.4 provide a powerful tool for the solution of a class of partial differential equations by computing the flow of a vector field.

Let us consider on a manifold $\mathbb{M}$ a time independent scalar field $\mathbf{v} \in C^0(\mathbb{M}\,;\mathbb{TM})$ which is Lipschitz continuous:

$$\|\mathbf{v}(\mathbf{x}) - \mathbf{v}(\mathbf{b})\| \leq \mathrm{Lip}\, \|\mathbf{x} - \mathbf{b}\|\,, \quad \mathrm{Lip} > 0\,,$$

and the partial differential equation:

$$\partial_{\tau=t}\; f(\mathbf{x}, \tau) = \mathcal{L}_{\mathbf{v}} f(\mathbf{x}, t)\,,$$

under the initial condition $f(\mathbf{x}, 0) = g(\mathbf{x})$.

Denoting by $\mathbf{Fl}^{\mathbf{v}}_t$ the flow of $\mathbf{v}$, the solution is given by $f(\mathbf{x}, t) = (\mathbf{Fl}^{\mathbf{v}}_t \!\downarrow\! g)(\mathbf{x}) = (g \circ \mathbf{Fl}^{\mathbf{v}}_t)(\mathbf{x})$. Indeed we have that

$$\partial_{\tau=t}\; (\mathbf{Fl}^{\mathbf{v}}_\tau \!\downarrow\! g)(\mathbf{x}) = (\mathbf{Fl}^{\mathbf{v}}_t \!\downarrow\! \mathcal{L}_{\mathbf{v}} g)(\mathbf{x}) = \mathcal{L}_{\mathbf{v}} (\mathbf{Fl}^{\mathbf{v}}_t \!\downarrow\! g)(\mathbf{x})\,.$$

The solution is then obtained by dragging the initial condition along the flow associated with the vector field $\mathbf{v} \in C^0(\mathbb{M}\,;\mathbb{TM})$. The integral curves of the vector field are called the *characteristic curves* and the solution methodology is called the *method of characteristics.*

1.3.6 Time dependent fields

Let us now consider the general case of time dependent vector fields which, denoting by $\mathbf{pr}_i$ the projection on the i-th component of a cartesian product, are defined as follows.

- A *time dependent vector field* is a mapping $\mathbf{v} \in C^0(\mathbb{M} \times I\,;\mathbb{TM})$ with $\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} = \mathbf{pr}_1$ and $I = [-\varepsilon, +\varepsilon\,,]$, $\varepsilon > 0$.

The integral curve of $\mathbf{v} \in C^0(\mathbb{M} \times I\,;\mathbb{TM})$ passing thru $\mathbf{x} \in \mathbb{M}$ at time $t = 0$ is the unique curve $\mathbf{c} \in C^1(I\,;\mathbb{M})$ solution of the differential equation

$$\partial_{\tau=t}\; \mathbf{c}(\tau) = \mathbf{v}(\mathbf{c}(t), t)\,, \quad t \in I\,,$$

under the initial condition $\mathbf{c}(0) = \mathbf{x} \in \mathbb{M}$.

- The *evolution operator* or *time dependent flow* associated with the time dependent vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M} \times I\,;\mathbb{TM})$ is the smooth map $\mathbf{Fl}^{\mathbf{v}} \in \mathrm{C}^1(\mathbb{M} \times I \times I\,;\mathbb{M})$ such that $\mathbf{Fl}^{\mathbf{v}}_{t,s}(\mathbf{x}) = \mathbf{c}(t)$ is the integral curve of the vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M} \times I\,;\mathbb{TM})$ passing thru $\mathbf{x} \in \mathbb{M}$ at time $s \in I$.

To a time dependent vector field $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\times I\,;\mathbb{TM})$ on $\mathbb{M}$ we may associate a time independent tangent vector field on the product manifold $\mathbb{M}\times I$, denoted by $\bar{\mathbf{v}} \in \mathrm{C}^0(\mathbb{M} \times I\,;\mathbb{TM} \times \mathbb{T}I)$, with $\boldsymbol{\tau}_{\mathbb{M}\times I} \circ \bar{\mathbf{v}} = \mathbf{id}_{\mathbb{M}\times I}$, and defined by the relation:

$$\bar{\mathbf{v}}(\mathbf{x}, t) := \{\mathbf{v}(\mathbf{x}, t)\,, 1_t\} \in \mathbb{T}_{\mathbf{x}}\mathbb{M} \times \mathbb{T}_t I\,.$$

Then it is

$$\mathbf{Fl}^{\bar{\mathbf{v}}}_{t-s}(\mathbf{x}, s) = \{\mathbf{Fl}^{\mathbf{v}}_{t,s}(\mathbf{x})\,, t\}\,.$$

The uniqueness of the integral curves implies the validity of the Chapman-Kolmogorov *law of determinism*:

$$\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \circ \mathbf{Fl}^{\mathbf{v}}_{t,s} = \mathbf{Fl}^{\mathbf{v}}_{\tau,s}\,.$$

Since $\mathbf{Fl}^{\mathbf{v}}_{s,t} \circ \mathbf{Fl}^{\mathbf{v}}_{t,s} = \mathbf{Fl}^{\mathbf{v}}_{s,s}$ is the identity map $\mathbf{id}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$, we have that $\mathbf{Fl}^{\mathbf{v}}_{s,t} = {\mathbf{Fl}^{\mathbf{v}}_{t,s}}^{-1}$ and, by differentiating with respect to time, we get the relation

$$\partial_{\tau=t}\, (\mathbf{Fl}^{\mathbf{v}}_{s,\tau} \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,s}) = (\partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{s,\tau}) \circ \mathbf{Fl}^{\mathbf{v}}_{t,s} + T\mathbf{Fl}^{\mathbf{v}}_{s,t} \circ \partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{\tau,s} = 0\,.$$

Being $\partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{\tau,s} = \mathbf{v}_t \circ \mathbf{Fl}^{\mathbf{v}}_{t,s}$, we get that

$$\partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{s,\tau} = -T\mathbf{Fl}^{\mathbf{v}}_{s,t} \circ \mathbf{v}_t = -(\mathbf{Fl}^{\mathbf{v}}_{s,t}{\uparrow}\mathbf{v}_t) \circ \mathbf{Fl}^{\mathbf{v}}_{s,t}\,,$$

or equivalently:

$$\partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{t,\tau} = -\mathbf{v}_t\,,$$

i.e. the velocity of the inverse evolution is the opposite of the velocity of the direct evolution.

1.3.7 Convective time derivative

Definition 1.3.2 *The* ***convective time derivative*** *of a time-dependent tensor field* $\mathbf{A}_t$ *on* $\mathbb{M}$*, along the evolution* $\mathbf{Fl}^{\mathbf{v}}$ *associated with a time-dependent vector field* $\mathbf{v}_t \in \mathrm{C}^0(\mathbb{M} \times I\,;\mathbb{TM})$ *is defined, according to the* Leibniz *rule, by*

$$\mathcal{L}_{t,\mathbf{v}}\mathbf{A}_t := \partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{\tau,t}{\downarrow}\mathbf{A}_\tau = \partial_{\tau=t}\, \mathbf{A}_\tau + \partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{\tau,t}{\downarrow}\mathbf{A}_t\,.$$

It is then the sum of two terms:

- the *partial time-derivative*
$$\partial_{\tau=t}\ \mathbf{A}_\tau$$
which takes account only of the changes induced by time on the tensor field $\mathbf{A}_\tau$, by considering the evolution as frozen-in at time t,
- the LIE *derivative* or *convective derivative*
$$\mathcal{L}_{\mathbf{v}_t}\mathbf{A}_t := \partial_{\tau=t}\ \mathbf{Fl}^{\mathbf{v}}_{\tau,t}{\downarrow}\mathbf{A}_t\,,$$
which takes account only of the changes induced by the evolution on the tensor field $\mathbf{A}_t$ considered as frozen-in at time t.

The next proposition extends the result stated in Proposition 1.3.1 to time dependent fields.

Proposition 1.3.12 *Along the evolution* $\mathbf{Fl}^{\mathbf{v}} \in \mathrm{C}^1(\mathbb{M}\times I\times I\,;\mathbb{M})$ *of a time dependent vector field* $\mathbf{v}\in\mathrm{C}^0(\mathbb{M}\times I\,;\mathbb{TM})$ *we have that*
$$\mathbf{Fl}^{\mathbf{v}}_{t,s}{\downarrow}\mathcal{L}_{t,\mathbf{v}}\mathbf{A}_t = \partial_{\tau=t}\ (\mathbf{Fl}^{\mathbf{v}}_{\tau,s}{\downarrow}\mathbf{A}_\tau)\,.$$

Proof. Recalling the expression
$$\mathbf{Fl}^{\mathbf{v}}_{\tau,s}{\downarrow}\mathbf{A}_\tau = (\mathbf{Fl}^{\mathbf{v}}_{\tau,t}\circ\mathbf{Fl}^{\mathbf{v}}_{t,s}){\downarrow}\mathbf{A}_\tau = \mathbf{Fl}^{\mathbf{v}}_{t,s}{\downarrow}(\mathbf{Fl}^{\mathbf{v}}_{\tau,t}{\downarrow}\mathbf{A}_\tau)\,,$$
and observing that
$$\partial_{\tau=t}\ (\mathbf{Fl}^{\mathbf{v}}_{\tau,s}{\downarrow}\mathbf{A}_\tau) = \mathbf{Fl}^{\mathbf{v}}_{t,s}{\downarrow}(\partial_{\tau=t}\ (\mathbf{Fl}^{\mathbf{v}}_{\tau,t}{\downarrow}\mathbf{A}_\tau)) = \mathbf{Fl}^{\mathbf{v}}_{t,s}{\downarrow}\mathcal{L}_{t,\mathbf{v}}\mathbf{A}_t\,,$$
we get the result. ■

If $\mathcal{L}_{t,\mathbf{v}}\mathbf{A}_t$ vanishes identically on the temporal domain of the flow, by Proposition 1.3.12 we infer that $\partial_{\tau=t}\ (\mathbf{Fl}^{\mathbf{v}}_{\tau,s}{\downarrow}\mathbf{A}_\tau)=0$ and then
$$\mathbf{Fl}^{\mathbf{v}}_{t,s}{\downarrow}\mathbf{A}_t = \mathbf{Fl}^{\mathbf{v}}_{s,s}{\downarrow}\mathbf{A}_s = \mathbf{A}_s\,,\quad \forall\, t\in I\,,$$
that is $\mathbf{Fl}^{\mathbf{v}}_{t,s}{\uparrow}\mathbf{A}_s = \mathbf{A}_t\,,\quad \forall\, t\in I$. We may then state that

- The convective time derivative $\mathcal{L}_{t,\mathbf{v}}\mathbf{A}_t$ vanishes identically if and only if the time dependent tensor field $\mathbf{A}_t \in BL(\mathbb{TM},\mathbb{T}^*\mathbb{M}\,;\Re)$ is dragged along the flow. In particular the LIE derivative $\mathcal{L}_{t,\mathbf{v}}f_t$ of a scalar field $f_t\in\mathrm{C}^1(\mathbb{M}\,;\Re)$ vanishes identically if and only if the time dependent scalar field is constant along the flow.

1.3.8 Time dependent diffeomorphisms

The results of Propositions 1.2.5 and 1.3.4 can be extended to flows of time dependent diffeomorphisms, as illustrated in Propositions 1.3.13 and 1.3.14.

Proposition 1.3.13 (Flows of time dependent pushes) *Given a time dependent diffeomorphism* $\varphi_t \in C^1(\mathbb{M}\,;\mathbb{N})$ *, let us consider the evolution operator*

$$\mathbf{Fl}^{\mathbf{v}}_{t,s} := \varphi_t \circ \varphi_s^{-1} \in C^1(\mathbb{N}\,;\mathbb{N})\,,$$

and let $\mathbf{v}_t \in C^1(\mathbb{N}\,;\mathbb{TN})$ *be the relevant time dependent velocity vector field:*

$$\partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{\tau,s} = \mathbf{v}_t \circ \mathbf{Fl}^{\mathbf{v}}_{t,s}\,, \qquad \mathbf{Fl}^{\mathbf{v}}_{s,s}(\mathbf{x}) = \mathbf{x}\,, \quad \forall\, \mathbf{x} \in \mathbb{N}\,.$$

Moreover let $\mathbf{u}_t \in C^1(\mathbb{M}\,;\mathbb{TM})$ *and* $\mathbf{w}_t \in C^1(\mathbb{N}\,;\mathbb{TN})$ *be time dependent vector fields. The following equivalence then holds*

$$\boxed{\mathbf{w}_t = \mathbf{v}_t + \varphi_t{\uparrow}\mathbf{u}_t \iff \varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} = \mathbf{Fl}^{\mathbf{w}}_{t,s} \circ \varphi_s\,.}$$

Proof. By differentiating the expression $\varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} = \mathbf{Fl}^{\mathbf{w}}_{t,s} \circ \varphi_s$ with respect to time, we have that

$$\begin{aligned}
\partial_{\tau=t}\, (\varphi_\tau \circ \mathbf{Fl}^{\mathbf{u}}_{\tau,s}) &= (\partial_{\tau=t}\, \varphi_\tau) \circ \mathbf{Fl}^{\mathbf{u}}_{\tau,s} + T\varphi_t \circ (\partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{u}}_{\tau,s}) \\
&= \mathbf{v}_t \circ \varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} + T\varphi_t \circ \mathbf{u}_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s}\,, \\
&= \mathbf{v}_t \circ \varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} + (\varphi_t{\uparrow}\mathbf{u}_t) \circ \varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s}\,, \\
\partial_{\tau=t}\, (\mathbf{Fl}^{\mathbf{w}}_{\tau,s} \circ \varphi_s) &= \partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{w}}_{\tau,s} \circ \varphi_s = \mathbf{w}_t \circ \mathbf{Fl}^{\mathbf{w}}_{t,s} \circ \varphi_s \\
&= \mathbf{w}_t \circ \varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s}\,.
\end{aligned}$$

By equating the two expressions above we get that $\mathbf{v}_t + \varphi_t{\uparrow}\mathbf{u}_t = \mathbf{w}_t$. Vice versa if this equality holds, we have that

$$\partial_{\tau=t}\, (\varphi_\tau \circ \mathbf{Fl}^{\mathbf{u}}_{\tau,s}) = \mathbf{w}_t \circ (\varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s})\,.$$

Hence the curve $(\varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{\tau,s})(\mathbf{x})$ with $\mathbf{x} \in \mathbb{M}$ is the integral curve of the vector field $\mathbf{w}_t$ passing thru $\varphi_s(\mathbf{x}) \in \mathbb{N}$ at the time s . The uniqueness of the integral implies that $\varphi_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} = \mathbf{Fl}^{\mathbf{w}}_{t,s} \circ \varphi_s$. ∎

The result of Proposition 1.3.13 can be expressed as follows

- The velocity of a pushed flow is equal to the velocity of the pushing flow plus the push of the velocity of the flow.

Proposition 1.3.14 (Lie derivative of the time dependent push) *Let* $\mathbf{A}_t : \mathbb{M} \mapsto BL(\mathbb{TM}, \mathbb{T}^*\mathbb{M}; \Re)$ *be a tensor field and* $\mathbf{u}_t : \mathbb{M} \mapsto \mathbb{TM}$ *be a vector field on the manifold* $\mathbb{M}$. *For each diffeomorphism* $\boldsymbol{\varphi}_t : \mathbb{M} \mapsto \mathbb{N}$ *the following formula then holds*

$$\mathcal{L}_{t,\mathbf{w}}(\boldsymbol{\varphi}_t\!\uparrow\!\mathbf{A}_t) = \boldsymbol{\varphi}_t\!\uparrow\!(\mathcal{L}_{t,\mathbf{u}}\mathbf{A}_t),$$

where $\mathbf{w}_t = \mathbf{v}_t + \boldsymbol{\varphi}_t\!\uparrow\!\mathbf{u}_t$ *is the velocity of the flow* $\mathbf{Fl}^{\mathbf{w}}_{t,s} = \boldsymbol{\varphi}_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} \circ \boldsymbol{\varphi}_s^{-1}$.

Proof. By Proposition 1.3.13 we have that

$$\begin{aligned}
\mathbf{Fl}^{\mathbf{w}}_{t,s}\!\downarrow\!(\boldsymbol{\varphi}_t\!\uparrow\!\mathbf{A}_t) &= (\boldsymbol{\varphi}_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} \circ \boldsymbol{\varphi}_s^{-1})\!\downarrow\!(\boldsymbol{\varphi}_t\!\uparrow\!\mathbf{A}_t) \\
&= (\boldsymbol{\varphi}_s \circ {\mathbf{Fl}^{\mathbf{u}}_{t,s}}^{-1} \circ \boldsymbol{\varphi}_t^{-1})\!\uparrow\!(\boldsymbol{\varphi}_t\!\uparrow\!\mathbf{A}_t) = \\
&= \boldsymbol{\varphi}_s\!\uparrow\!(\mathbf{Fl}^{\mathbf{u}}_{t,s}\!\downarrow\!\mathbf{A}_t).
\end{aligned}$$

The result then follows from the definition of Lie derivative. Indeed we have that

$$\begin{aligned}
\mathcal{L}_{s,\mathbf{w}}(\boldsymbol{\varphi}_s\!\uparrow\!\mathbf{A}_s) &:= \partial_{t=s}\, \mathbf{Fl}^{\mathbf{w}}_{t,s}\!\downarrow\!(\boldsymbol{\varphi}_t\!\uparrow\!\mathbf{A}_t), \\
\boldsymbol{\varphi}_s\!\uparrow\!(\mathcal{L}_{s,\mathbf{u}}\mathbf{A}_s) &:= \partial_{t=s}\, \boldsymbol{\varphi}_s\!\uparrow\!(\mathbf{Fl}^{\mathbf{u}}_{t,s}\!\downarrow\!\mathbf{A}_t),
\end{aligned}$$

which proves the result at time $s \in I$. ■

The formula in Proposition 1.3.13 expresses in a generalized form the relation between the velocity fields measured by two observers in relative motion.

The formula in Proposition 1.3.14 has important applications in mechanics for the definition of objective time derivatives of the stress tensor [106], [164].

1.3.9 Generalized Lie derivative

Let us now introduce, by making reference to the treatment in [80], a more general definition of the Lie derivative proposed in [185]. The definition of the Lie derivative of a cross section of a fibre bundle will soon be useful in introducing the concept of a connection in a fibre bundle (see section 1.4).

Let $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}; \mathbb{M})$ and $\boldsymbol{\tau}_{\mathbb{N}} \in C^1(\mathbb{TN}; \mathbb{N})$ be tangent bundles, and let $\mathbf{f} \in C^1(\mathbb{M}; \mathbb{N})$ be a smooth map with tangent map $T\mathbf{f} \in C^1(\mathbb{TM}; \mathbb{TN})$.

Definition 1.3.3 *The generalized* LIE *derivative of* $\mathbf{f} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$ *along the pair of vector fields* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ *and* $\mathbf{X} \in \mathrm{C}^1(\mathbb{N};\mathbb{TN})$ *is the gap of commutativity of the diagram:*

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\mathbf{f}} & \mathbb{TN} \\ \mathbf{v}\uparrow & & \uparrow\mathbf{X} \\ \mathbb{M} & \xrightarrow{\mathbf{f}} & \mathbb{N} \end{array} \qquad \textit{that is} \qquad \mathcal{L}_{(\mathbf{X},\mathbf{v})}\,\mathbf{f} := T\mathbf{f} \circ \mathbf{v} - \mathbf{X} \circ \mathbf{f}\,.$$

The definition is well-posed since $(T\mathbf{f}\circ\mathbf{v})(\mathbf{x}), (\mathbf{X}\circ\mathbf{f})(\mathbf{x}) \in \mathbb{T}_{\mathbf{f}(\mathbf{x})}\mathbb{N}$ *for all* $\mathbf{x} \in \mathbb{M}$ *.*

From the definition it follows that $\mathcal{L}_{(\mathbf{X},\mathbf{v})}\,\mathbf{f} \in \mathrm{C}^1(\mathbb{M};\mathbb{TN})$ vanishes if and only if the vector fields $\mathbf{X}$ and $\mathbf{v}$ are $\mathbf{f}$-related. Moreover $\mathcal{L}_{(\mathbf{X},\mathbf{v})}\,\mathbf{f} \in \mathrm{C}^1(\mathbb{M};\mathbb{TN})$ is a vector field along $\mathbf{f} \in \mathrm{C}^1(\mathbb{M};\mathbb{N})$. Indeed

$$\left.\begin{array}{l} \boldsymbol{\tau}_{\mathbb{N}} \circ T\mathbf{f} \circ \mathbf{v} = \mathbf{f} \circ \boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{v} = \mathbf{f} \\ \boldsymbol{\tau}_{\mathbb{N}} \circ \mathbf{X} \circ \mathbf{f} = \mathbf{f} \end{array}\right\} \Longrightarrow \boldsymbol{\tau}_{\mathbb{N}} \circ \mathcal{L}_{(\mathbf{X},\mathbf{v})}\,\mathbf{f} = \mathbf{f}\,.$$

A direct computation shows that the previous definition is equivalent to:

$$\mathcal{L}_{(\mathbf{X},\mathbf{v})}\,\mathbf{f} := \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{f} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda})\,.$$

Section of fibre bundles

An important special case is met when the manifold $\mathbb{N}$ is the total space of a fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$ and the map $\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ is a section of the fibre bundle so that $\mathbf{p}\circ\mathbf{s} = \mathbf{id}_{\mathbb{M}}$. Indeed, let $\mathbf{X} \in \mathrm{C}^1(\mathbb{E};\mathbb{TE})$ be a vector field which admits a projected vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$. This means that the following diagram is commutative:

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{\mathbf{X}} & \mathbb{TE} \\ \mathbf{p}\downarrow & & \downarrow T\mathbf{p} \\ \mathbb{M} & \xrightarrow{\mathbf{v}} & \mathbb{TM} \end{array} \qquad \text{with} \qquad T\mathbf{p} \circ \mathbf{X} = \mathbf{v} \circ \mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{TM})\,,$$

i.e. that the vector fields $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ and $\mathbf{X} \in \mathrm{C}^1(\mathbb{E};\mathbb{TE})$ are $\mathbf{p}$-related.

Then the LIE derivative is defined by

$$\mathcal{L}_{(\mathbf{X},\mathbf{v})}\mathbf{s} := \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}) \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TE})\,.$$

By Proposition 1.2.5 it is $\mathbf{p} \circ \mathbf{Fl}^{\mathbf{X}}_{-\lambda} = \mathbf{Fl}^{\mathbf{v}}_{-\lambda} \circ \mathbf{p}$ and hence, being $\mathbf{p} \circ \mathbf{s} = \mathbf{id}_{\mathbb{M}}$, we have that $\mathbf{p} \circ \mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{s} = \mathbf{Fl}^{\mathbf{v}}_{-\lambda}$ which may be written

$$\mathbf{p} \circ \mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} = \mathbf{id}_{\mathbb{M}}\,.$$

Taking the derivative $\partial_{\lambda=0}$ we get:

$$T\mathbf{p} \circ \mathcal{L}_{(\mathbf{X},\mathbf{v})}\mathbf{s} = 0\,.$$

Then the LIE derivative $\mathcal{L}_{(\mathbf{X},\mathbf{v})}\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{VE})$ is a section of the vertical fibre bundle $\mathbf{p} \circ \boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{VE}\,;\mathbb{M})$ over the bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, with $\mathbb{VE} := \mathbf{ker}(T\mathbf{p}) \subset \mathbb{TE}$.

We have the following result which will be referred to in Lemma 1.4.17.

Proposition 1.3.15 (Pull back along related flows) *The* LIE *derivative of a cross section of a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *has the property*

$$\mathbf{Fl}^{(\mathbf{X},\mathbf{v})}_{\lambda}\!\downarrow\!(\mathcal{L}_{(\mathbf{X},\mathbf{v})}\mathbf{s}) = \partial_{\mu=\lambda}\,(\mathbf{Fl}^{(\mathbf{X},\mathbf{v})}_{\mu}\!\downarrow\!\mathbf{s})\,,$$

with the pull back defined by

$$\mathbf{Fl}^{(\mathbf{X},\mathbf{v})}_{\lambda}\!\downarrow\!\mathbf{s} := \mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}\,.$$

Proof. The result is a direct consequence of the equality

$$\begin{aligned}\mathbf{Fl}^{(\mathbf{X},\mathbf{v})}_{\lambda}\!\downarrow\!(\partial_{\mu=0}\,(\mathbf{Fl}^{\mathbf{X}}_{-\mu} \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{\mu})) &= \partial_{\mu=0}\,\mathbf{Fl}^{\mathbf{X}}_{-\lambda} \circ \mathbf{Fl}^{\mathbf{X}}_{-\mu} \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{\mu} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \\ &= \partial_{\mu=\lambda}\,\mathbf{Fl}^{\mathbf{X}}_{-\mu} \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{\mu}\,,\end{aligned}$$

and of the definition of pull back given in the statement. ■

Section of vector bundles

In a vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ the elements of the vertical space $\mathbb{V}_{\mathbf{e}}\mathbb{E} := \mathbf{ker}(T_{\mathbf{e}}\mathbf{p}) \subset \mathbb{T}_{\mathbf{e}}\mathbb{E}$ at a point $\mathbf{e} \in \mathbb{E}$ may be identified with vectors of the linear fibre $\mathbb{E}_{\mathbf{e}} := \mathbf{p}^{-1}(\mathbf{e})$. Accordingly, the fibre bundle $\mathbf{p} \circ \boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{VE}\,;\mathbb{M})$ may be identified with the vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ and the LIE derivative $\mathcal{L}_{(\mathbf{X},\mathbf{v})}\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{VE})$ may be regarded as a section $\mathcal{L}_{(\mathbf{X},\mathbf{v})}\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ of the vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$.

Section of tangent bundles

The LIE derivative of a section $\mathbf{s} \in C^1(\mathbb{M};\mathbb{TM})$ of the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM};\mathbb{M})$ is recovered from the definition of the generalized LIE derivative by setting $\mathbf{X} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v}$ so that

$$\begin{array}{ccccc} \mathbb{M} & \xrightarrow{\mathbf{s}} & \mathbb{TM} & \xleftarrow{\mathbf{s}} & \mathbb{M} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{M}}}\uparrow & & \downarrow{\scriptstyle \mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ T\mathbf{v}} & & \downarrow{\scriptstyle \mathbf{v}} \\ \mathbb{TM} & \xrightarrow{T\mathbf{s}} & \mathbb{T}^2\mathbb{M} & \xrightarrow{T\boldsymbol{\tau}_{\mathbb{M}}} & \mathbb{TM} \end{array} \qquad \text{with} \qquad T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{s} = \mathbf{v}\,,$$

and we have that

$$\begin{aligned} \mathcal{L}_{\mathbf{v}}\, \mathbf{s} &:= \partial_{\lambda=0}\, (\mathbf{Fl}_{-\lambda}^{\mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ T\mathbf{v}} \circ \mathbf{s} \circ \mathbf{Fl}_{\lambda}^{\mathbf{v}}) \\ &= \partial_{\lambda=0}\, (T\mathbf{Fl}_{-\lambda}^{\mathbf{v}} \circ \mathbf{s} \circ \mathbf{Fl}_{\lambda}^{\mathbf{v}}) \\ &= T\mathbf{s} \circ \mathbf{v} - \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{v} \circ \mathbf{s}\,. \end{aligned}$$

Scalar functions

The LIE derivative of a scalar function $f \in C^1(\mathbb{M};\Re)$ on the manifold $\mathbb{M}$ is recovered from the definition of the generalized LIE derivative by setting $\mathbf{X} \in C^1(\Re;\mathbb{T}\Re)$ equal to the sero section of $\Re$ so that $\mathbf{Fl}_{-\lambda}^{\mathbf{X}} = \mathbf{Fl}_{-\lambda}^{0} = \mathbf{id}_{\Re}$:

$$\begin{array}{ccc} \mathbb{M} & \xrightarrow{f} & \Re \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{M}}}\uparrow & & \downarrow{\scriptstyle \{\mathbf{id}_{\Re},0\}} \\ \mathbb{TM} & \xrightarrow{Tf} & \mathbb{T}\Re \end{array}$$

and we have that

$$\mathcal{L}_{\mathbf{v}}\, f := \partial_{\lambda=0}\, (\mathbf{Fl}_{-\lambda}^{0} \circ f \circ \mathbf{Fl}_{\lambda}^{\mathbf{v}}) = \partial_{\lambda=0}\, (f \circ \mathbf{Fl}_{\lambda}^{\mathbf{v}}) = Tf \circ \mathbf{v}\,.$$

Morphism of fibre bundles

Let us consider two fibre bundles $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ and $\mathbf{q} \in C^1(\mathbb{Z};\mathbb{M})$ over the same base manifold $\mathbb{M}$ and two vector fields $\mathbf{u} \in C^1(\mathbb{E};\mathbb{TE})$ and $\mathbf{w} \in C^1(\mathbb{Z};\mathbb{TZ})$ which project on the same base vector field $\mathbf{v} \in C^1(\mathbb{M};\mathbb{TM})$:

$$T\mathbf{p} \circ \mathbf{u} = \mathbf{v} \circ \mathbf{p}\,, \quad T\mathbf{q} \circ \mathbf{w} = \mathbf{v} \circ \mathbf{q}\,,$$

that is $\mathbf{v} = \mathbf{p}{\uparrow}\mathbf{u} = \mathbf{q}{\uparrow}\mathbf{w}$.

Let $\mathbf{f} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{Z})$ be a base preserving morphism: $\mathbf{q} \circ \mathbf{f} = \mathbf{p}$. Then we have the following commutative diagram:

$$
\begin{array}{ccccccc}
\mathbb{TE} & \xleftarrow{\;\mathbf{u}\;} & \mathbb{E} & \xrightarrow{\;\mathbf{f}\;} & \mathbb{Z} & \xrightarrow{\;\mathbf{w}\;} & \mathbb{TZ} \\
\Big\downarrow T\mathbf{p} & & \Big\downarrow \mathbf{p} & & \Big\downarrow \mathbf{q} & & \Big\downarrow T\mathbf{q} \\
\mathbb{TM} & \xleftarrow{\;\mathbf{v}\;} & \mathbb{M} & \xrightarrow{\;\mathbf{id}_{\mathbb{M}}\;} & \mathbb{M} & \xrightarrow{\;\mathbf{v}\;} & \mathbb{TM}
\end{array}
$$

The generalized LIE derivative of $\mathbf{f}$ along the pair $\{\mathbf{u}\,,\mathbf{w}\}$ is defined as

$$
\mathcal{L}_{\{\mathbf{u}\,,\mathbf{w}\}}\,\mathbf{f} := T\mathbf{f} \circ \mathbf{u} - \mathbf{w} \circ \mathbf{f} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{VZ})\,,
$$

with $\mathbb{VZ} := \mathbf{ker}(T\mathbf{q}) \subset \mathbb{TZ}$. Indeed, the LIE derivative is also defined by

$$
\mathcal{L}_{\{\mathbf{u}\,,\mathbf{w}\}}\,\mathbf{f} := \partial_{\lambda=0}\,\left(\mathbf{Fl}^{\mathbf{w}}_{-\lambda} \circ \mathbf{f} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda}\right) \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{TZ})\,.
$$

We claim that $\mathbf{q} \circ \mathbf{Fl}^{\mathbf{w}}_{-\lambda} \circ \mathbf{f} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} = \mathbf{p}$.

$$
\begin{array}{ccccc}
\mathbb{M} & \xleftarrow{\;\mathbf{q}\;} & & \xleftarrow{\;\mathbf{q}\;} & \mathbb{Z} \\
\mathbf{id}_{\mathbb{M}}\Big\uparrow & & \mathbf{id}_{\mathbb{E}}\Big\uparrow & & \mathbf{id}_{\mathbb{Z}}\Big\uparrow \\
\mathbb{M} & \xleftarrow{\;\mathbf{p}\;} & \mathbb{E} & \xrightarrow{\;\mathbf{f}\;} & \mathbb{Z} \\
\mathbf{Fl}^{\mathbf{v}}_{\lambda}\Big\uparrow & & \Big\uparrow\mathbf{Fl}^{\mathbf{u}}_{\lambda} & & \Big\downarrow\mathbf{Fl}^{\mathbf{w}}_{-\lambda} \\
\mathbb{M} & \xleftarrow{\;\mathbf{p}\;} & \mathbb{E} & \xrightarrow{\;\mathbf{f}\;} & \mathbb{Z}
\end{array}
$$

Indeed $\mathbf{q} \circ \mathbf{Fl}^{\mathbf{w}}_{-\lambda} = \mathbf{Fl}^{\mathbf{q}\uparrow\mathbf{w}}_{\lambda} \circ \mathbf{q} = \mathbf{Fl}^{\mathbf{p}\uparrow\mathbf{u}}_{\lambda} \circ \mathbf{q}$, so that

$$
\begin{aligned}
\mathbf{q} \circ \mathbf{Fl}^{\mathbf{w}}_{-\lambda} \circ \mathbf{f} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} &= \mathbf{Fl}^{\mathbf{p}\uparrow\mathbf{u}}_{-\lambda} \circ \mathbf{q} \circ \mathbf{f} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} \\
&= \mathbf{Fl}^{\mathbf{p}\uparrow\mathbf{u}}_{-\lambda} \circ \mathbf{p} \circ \mathbf{Fl}^{\mathbf{u}}_{\lambda} \\
&= \mathbf{Fl}^{\mathbf{p}\uparrow\mathbf{u}}_{-\lambda} \circ \mathbf{Fl}^{\mathbf{p}\uparrow\mathbf{u}}_{\lambda} \circ \mathbf{p} = \mathbf{p}\,.
\end{aligned}
$$

Taking the derivative $\partial_{\lambda=0}$ we infer that $T\mathbf{q} \circ \mathcal{L}_{\{\mathbf{u}\,,\mathbf{w}\}}\,\mathbf{f} = 0$.

Then the LIE derivative takes values into the *vertical bundle* over the bundle $\mathbf{q} \in \mathrm{C}^1(\mathbb{Z}\,;\mathbb{M})$, defined by $\mathbb{VZ} := \mathbf{ker}(T\mathbf{q}) \subset \mathbb{TZ}$.

1.4 Connection on a fibre bundle

The notion of connection on a fibre bundle was introduced by CHARLES EHRESMANN in 1950 , [45] and investigated upon by PAULETTE LIBERMANN in [96], [97], [98], [99].

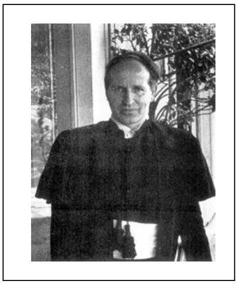

Figure 1.17: Charles Ehresmann (1905 - 1979)

We develop here, with an original methodology, a treatment based on the exposition of the main concepts given in [80]. The most significant new contributions are the results in Theorems 1.4.2 and 1.4.5, which respectively provide a simplest proof of FROBENIUS integrability condition based on clear geometrical arguments and a direct bridge between this condition and the expression of the curvature in terms of covariant derivatives. Let us give the following definitions.

Definition 1.4.1 *A* ***connection*** *in a fibre bundle* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ *is a tangent valued one-form* $P_V \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$ *which is a pointwise projector on the vertical bundle.*

Then $P_V \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$ is **idempotent**: $P_V \circ P_V = P_V$, with $\mathbf{im}(P_V) = \mathbf{ker}(T\mathbf{p})$.

Definition 1.4.2 ***Horizontal*** *vector fields are the ones in the kernel* $\mathbf{ker}(P_V)$ *of the connection. The projector* $P_H \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$ *on the horizontal bundle is given by* $P_H = \mathbf{id}_{\mathbb{TE}} - P_V$ *and it is* $P_H \circ P_H = P_H$ *and* $P_H \circ P_V = P_V \circ P_H = 0$.

Lemma 1.4.1 (Characterization of connections) *A connection is characterized by an involutive tangent valued one-form* $\mathbf{\Gamma} \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$ *related to the* *projector* $P_{\mathrm{V}} \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$ *by the equivalent properties:*

$$\begin{cases} P_{\mathrm{H}} + P_{\mathrm{V}} = \mathbf{I} \\ P_{\mathrm{H}} - P_{\mathrm{V}} = \mathbf{\Gamma} \end{cases} \iff \begin{cases} 2\,P_{\mathrm{H}} = \mathbf{I} + \mathbf{\Gamma} \\ 2\,P_{\mathrm{V}} = \mathbf{I} - \mathbf{\Gamma} \end{cases} \iff \begin{cases} \mathbf{\Gamma}^2 = \mathbf{I} \\ \boldsymbol{ker}(\mathbf{I} + \mathbf{\Gamma}) = \boldsymbol{ker}(T\mathbf{p}) \end{cases}$$

Definition 1.4.3 *The **natural derivative** of a section* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *of a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *according to a vector field* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *is defined by*

$$T_{\mathbf{v}} \circ \mathbf{s} := T\mathbf{s} \circ \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TE})\,.$$

Lemma 1.4.2 (Natural derivative) *In a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$*, for any given section* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$*, the natural derivative according to a tangent vector field* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *is a vector field* $T_{\mathbf{v}} \in \mathrm{C}^0(\mathbb{E}\,;\mathbb{TE})$ *in the tangent bundle* $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{E})$ *which projects on* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$*, i.e.*

$$\boldsymbol{\tau}_{\mathbb{E}} \circ T_{\mathbf{v}} = \mathbf{id}_{\mathbb{E}}\,,$$

$$T\mathbf{p} \circ T_{\mathbf{v}} = \mathbf{v} \circ \mathbf{p}\,.$$

Proof. The former property is inferred from the commutativity of the diagram

$$\begin{array}{ccc} \mathbb{TM} & \xrightarrow{T\mathbf{s}} & \mathbb{TE} \\ \mathbf{v}\big\uparrow & & \big\downarrow\boldsymbol{\tau}_{\mathbb{E}} \\ \mathbb{M} & \xrightarrow{\mathbf{s}} & \mathbb{E} \end{array} \qquad \text{with} \qquad \boldsymbol{\tau}_{\mathbb{E}} \circ T\mathbf{s} \circ \mathbf{v} = \mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})\,,$$

which follows since $T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E}$ so that $\boldsymbol{\tau}_{\mathbb{E}}(T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}}) = \mathbf{s}(\mathbf{x})$. Then $\boldsymbol{\tau}_{\mathbb{E}} \circ T_{\mathbf{v}} \circ \mathbf{s} = \mathbf{s}$ for all $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ and this gives the result. The latter property is inferred from the commutativity of the diagram

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{T_{\mathbf{v}}} & \mathbb{TE} \\ \mathbf{s}\big\uparrow & & \big\downarrow T\mathbf{p} \\ \mathbb{M} & \xrightarrow{\mathbf{v}} & \mathbb{TM} \end{array} \qquad \text{with} \qquad T\mathbf{p} \circ T_{\mathbf{v}} \circ \mathbf{s} = T\mathbf{p} \circ T\mathbf{s} \circ \mathbf{v} = \mathbf{v}\,,$$

for all sections $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$, which, being $\mathbf{p} \circ \mathbf{s} = \mathbf{id}_{\mathbb{M}}$, may be written as $T\mathbf{p} \circ T_{\mathbf{v}} \circ \mathbf{s} = \mathbf{v} \circ \mathbf{p} \circ \mathbf{s}$. Then $T\mathbf{p} \circ T_{\mathbf{v}} = \mathbf{v} \circ \mathbf{p}$. ■

The $\mathbf{p}$-relatedness between the natural derivative $T_{\mathbf{v}} \in C^0(\mathbb{E}\,;\mathbb{TE})$ and the vector field $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ gives the commutativity of the diagram:

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{\mathbf{Fl}^{T_{\mathbf{v}}}_{\lambda}} & \mathbb{E} \\ \mathbf{p}\Big\downarrow & & \mathbf{p}\Big\downarrow \\ \mathbb{M} & \xrightarrow{\mathbf{Fl}^{\mathbf{v}}_{\lambda}} & \mathbb{M} \end{array} \qquad \iff \qquad \mathbf{p} \circ \mathbf{Fl}^{T_{\mathbf{v}}}_{\lambda} = \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})\,.$$

The natural derivative $T_{\mathbf{v}} \in C^0(\mathbb{E}\,;\mathbb{TE})$ is tensorial in $\mathbf{v} \in C^0(\mathbb{M}\,;\mathbb{TM})$ since the vector $T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E}$ depends linearly on the vector $\mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$.

1.4.1 Horizontal lift in a fibre bundle

Definition 1.4.4 *The horizontal lift of a vector field* $\mathbf{v} \in C^0(\mathbb{M}\,;\mathbb{TM})$ *along a section* $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ *of the fibre bundle* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$*, is the horizontal component of the natural derivative:*

$$\mathbf{H}_{\mathbf{v}}\mathbf{s} := P_{\mathrm{H}} \circ T_{\mathbf{v}} \circ \mathbf{s} = P_{\mathrm{H}} \circ T\mathbf{s} \circ \mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TE})\,.$$

Then, for a given $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$, $P_{\mathrm{H}} \circ T\mathbf{s} \in C^1(\mathbb{TM}\,;\mathbb{HE})$ *and* $\mathbf{H}_{\mathbf{v}} \in C^1(\mathbb{E}\,;\mathbb{HE})$.

Lemma 1.4.3 (Injectivity) *For any section* $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ *of a fibre bundle* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$*, the map* $P_{\mathrm{H}} \circ T\mathbf{s} \in C^1(\mathbb{TM}\,;\mathbb{TE})$ *is a fibre bundle homomorphism from the bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}\,;\mathbb{M})$ *to the bundle* $\mathbf{p} \in C^1(\mathbb{HE}\,;\mathbb{E})$ *which is fibre-wise injective. This means that the linear map* $P_{\mathrm{H}} \circ T_{\mathbf{x}}\mathbf{s} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E})$ *is injective at each point* $\mathbf{x} \in \mathbb{M}$.

Proof. We have to prove that $\mathbf{ker}(P_{\mathrm{H}} \circ T_{\mathbf{x}}\mathbf{s}) = \{0\}$. We first investigate the linear differential $T_{\mathbf{x}}\mathbf{s} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E})$. From the characteristic property of a section, $\mathbf{p} \circ \mathbf{s} = \mathbf{id}_{\mathbb{M}}$, we get:

$$T_{\mathbf{s}(\mathbf{x})}\mathbf{p} \cdot T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}} = T_{\mathbf{x}}(\mathbf{p} \circ \mathbf{s}) \cdot \mathbf{v}_{\mathbf{x}} = \mathbf{v}_{\mathbf{x}}\,, \quad \forall\, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,.$$

Then $\mathbf{ker}(T_{\mathbf{x}}\mathbf{s}) = \{0\}$ and $\mathbf{im}(T_{\mathbf{x}}\mathbf{s}) \cap \mathbf{ker}(T_{\mathbf{s}(\mathbf{x})}\mathbf{p}) = \{0\}$. The injectivity of $T_{\mathbf{x}}\mathbf{s}$ implies that: $\dim \mathbf{im}(T_{\mathbf{x}}\mathbf{s}) = \dim \mathbb{T}_{\mathbf{x}}\mathbb{M}$. Being $T_{\mathbf{x}}\mathbf{s} = \nabla_{\mathbf{x}}\mathbf{s} + P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s}$ with $\mathbf{im}(\nabla_{\mathbf{x}}\mathbf{s}) \subseteq \mathbf{ker}(T_{\mathbf{s}(\mathbf{x})}\mathbf{p})$, we have that

$$T_{\mathbf{s}(\mathbf{x})}\mathbf{p} \cdot P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}} = T_{\mathbf{s}(\mathbf{x})}\mathbf{p} \cdot T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}} = \mathbf{v}_{\mathbf{x}}\,, \quad \forall\, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,.$$

We may conclude that $P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E})$ is a right inverse of $T_{\mathbf{s}(\mathbf{x})}\mathbf{p} \in BL(\mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E}\,;\mathbb{T}_{\mathbf{x}}\mathbb{M})$, that is

$$T_{\mathbf{s}(\mathbf{x})}\mathbf{p} \circ P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s} = \mathbf{id}_{\mathbb{T}_{\mathbf{x}}\mathbb{M}}\,.$$

It follows that $\mathbf{ker}(P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s}) = \{0\}$ and $\mathbf{im}(P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s}) \cap \mathbf{ker}(T_{\mathbf{s}(\mathbf{x})}\mathbf{p}) = \{0\}$ with $\dim \mathbf{im}(P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s}) = \dim \mathbb{T}_{\mathbf{x}}\mathbb{M}$. ■

Theorem 1.4.1 (Tensoriality of the horizontal lift) *Given a section* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ *of a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$, *the map* $P_{\mathrm{H}} \circ T\mathbf{s} \in \mathrm{C}^1(\mathbb{TM};\mathbb{TE})$ *is a vector bundle homomorphism from the bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM};\mathbb{M})$ *to the bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{HE};\mathbb{E})$ *which is fibrewise invertible and tensorial in* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$.

Proof. Let $\dim \mathbb{M} = \dim \mathbb{T}_{\mathbf{x}}\mathbb{M} = m$ and $\dim \mathbb{F} = f$ where $\mathbb{F}$ is the typical fibre. Being $\dim \mathbb{E} = \dim \mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E} = m + f$ we have that $\dim \mathbb{V}_{\mathbf{s}(\mathbf{x})}\mathbb{E} = f$ and $\dim \mathbb{H}_{\mathbf{s}(\mathbf{x})}\mathbb{E} = m$. By reasons of dimensions the injectivity of $P_{\mathrm{H}} \circ T_{\mathbf{x}}\mathbf{s} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M};\mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E})$ implies the surjectivity of $P_{\mathrm{H}} \circ T_{\mathbf{x}}\mathbf{s} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M};\mathbb{H}_{\mathbf{s}(\mathbf{x})}\mathbb{E})$. Let us now consider in the bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$ two sections $\mathbf{s}, \bar{\mathbf{s}} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ such that $\bar{\mathbf{s}}(\mathbf{x}) = \mathbf{s}(\mathbf{x})$. For any vector $\mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$, being $T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}}, T_{\mathbf{x}}\bar{\mathbf{s}} \cdot \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E}$, we have that

$$T\mathbf{p} \circ (T_{\mathbf{x}}\mathbf{s} - T_{\mathbf{x}}\bar{\mathbf{s}}) \cdot \mathbf{v}_{\mathbf{x}} = 0\,,$$

and hence that $P_{\mathrm{H}} \circ T_{\mathbf{x}}\mathbf{s} = P_{\mathrm{H}} \circ T_{\mathbf{x}}\bar{\mathbf{s}} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M};\mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E})$. Then to a tangent vector $\mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ there corresponds a horizontal vector $P_{\mathrm{H}} \circ T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v}_{\mathbf{x}} \in \mathbb{H}_{\mathbf{s}(\mathbf{x})}\mathbb{E}$ depending only on the value of $\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ at $\mathbf{x} \in \mathbb{M}$. ■

Lemma 1.4.4 (Projectability of horizontal lifts) *The vector field* $\mathbf{H}_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E};\mathbb{HE})$ *of horizontal lifts of a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ *is* $\mathbf{p}$*-related to the field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ *according to the commutative diagram*

$$\begin{array}{ccc} \mathbb{HE} & \xrightarrow{T\mathbf{p}} & \mathbb{TM} \\ \mathbf{H}_{\mathbf{v}} \uparrow & & \uparrow \mathbf{v} \\ \mathbb{E} & \xrightarrow{\mathbf{p}} & \mathbb{M} \end{array} \quad \iff \quad T\mathbf{p} \circ \mathbf{H}_{\mathbf{v}} = \mathbf{v} \circ \mathbf{p} \in \mathrm{C}^0(\mathbb{E};\mathbb{TM})\,.$$

Proof. From the decomposition $T_{\mathbf{v}} = \nabla_{\mathbf{v}} + \mathbf{H}_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E};\mathbb{TE})$ we infer that:

$$T\mathbf{p} \circ T_{\mathbf{v}} = T\mathbf{p} \circ \nabla_{\mathbf{v}} + T\mathbf{p} \circ \mathbf{H}_{\mathbf{v}} = T\mathbf{p} \circ \mathbf{H}_{\mathbf{v}}\,,$$

being, by definition $T\mathbf{p} \circ \nabla_{\mathbf{v}} = 0$. Then the projectability of the natural derivative stated in Lemma 1.4.2 implies that the horizontal lift $\mathbf{H}_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E};\mathbb{HE})$ also projects on $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$. ■

Lemma 1.4.5 *If the tangent field* $\mathbf{X} \in \mathrm{C}^1(\mathbb{E};\mathbb{TE})$ *projects on the tangent field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ *then* $P_{\mathrm{H}}\mathbf{X} \in \mathrm{C}^1(\mathbb{E};\mathbb{HE})$ *also projects on* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ *and* $P_{\mathrm{H}}\mathbf{X} = \mathbf{H}_{\mathbf{v}}$.

Proof. Since $T\mathbf{p} \cdot (P_{\mathrm{H}}\mathbf{X} - \mathbf{H_v}) = T\mathbf{p} \cdot P_{\mathrm{H}}\mathbf{X} - T\mathbf{p} \cdot \mathbf{H_v} = \mathbf{v} \circ \mathbf{p} - \mathbf{v} \circ \mathbf{p} = 0$, the difference of the horizontal fields $P_{\mathrm{H}}\mathbf{X}$ and $\mathbf{H_v}$, being vertical, vanishes. ■

The flow of the horizontal lift $\mathbf{H_v} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{H}\mathbb{E})$ fulfils the commutative diagram:

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{\mathbf{Fl}_\lambda^{\mathbf{H_v}}} & \mathbb{E} \\ {\scriptstyle \mathbf{p}}\big\downarrow & & \big\downarrow{\scriptstyle \mathbf{p}} \\ \mathbb{M} & \xrightarrow{\mathbf{Fl}_\lambda^{\mathbf{v}}} & \mathbb{M} \end{array} \quad \Longleftrightarrow \quad \mathbf{p} \circ \mathbf{Fl}_\lambda^{\mathbf{H_v}} = \mathbf{Fl}_\lambda^{\mathbf{v}} \circ \mathbf{p}\,.$$

The $\mathbf{p}$-relatedness of brackets of $\mathbf{p}$-related vector fields gives the commutative diagram:

$$\begin{array}{ccc} \mathbb{H}\mathbb{E} & \xrightarrow{T\mathbf{p}} & \mathbb{T}\mathbb{M} \\ {\scriptstyle [\mathbf{H_u}\,,\mathbf{H_v}]}\big\uparrow & & \big\uparrow{\scriptstyle [\mathbf{u}\,,\mathbf{v}]} \\ \mathbb{E} & \xrightarrow{\mathbf{p}} & \mathbb{M} \end{array} \quad \Longleftrightarrow \quad T\mathbf{p} \circ [\mathbf{H_u}\,, \mathbf{H_v}] = [\mathbf{u}\,, \mathbf{v}] \circ \mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{T}\mathbb{M})\,.$$

On the basis of the previous results we may state the following definitions and properties.

Definition 1.4.5 (Horizontal lift) *In a bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *the horizontal lift* $\mathbf{H} \in \mathrm{C}^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{T}\mathbb{M}\,;\mathbb{T}\mathbb{E})$ *is a right inverse of* $(\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \in \mathrm{C}^1(\mathbb{T}\mathbb{E}\,;\mathbb{E} \times_{\mathbb{M}} \mathbb{T}\mathbb{M})$ *such that the map* $\mathbf{H}_{\mathbf{s_x}} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{T}\mathbb{E})$ *defined by* $\mathbf{H}_{\mathbf{s_x}}(\mathbf{v_x}) = \mathbf{H}(\mathbf{s_x}\,, \mathbf{v_x})$ *is a linear homomorphism from the tangent bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{M})$ *to the tangent bundle* $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{T}\mathbb{E}\,;\mathbb{E})$, *i.e:*

$$(\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \circ \mathbf{H} = \mathbf{id}_{\mathbb{E} \times_{\mathbb{M}} \mathbb{T}\mathbb{M}}\,,$$

$$\mathbf{H}_{\mathbf{s_x}}(\alpha\, \mathbf{u_x} + \beta\, \mathbf{v_x}) = \alpha\, \mathbf{H}_{\mathbf{s_x}}(\mathbf{u_x}) + \beta\, \mathbf{H}_{\mathbf{s_x}}(\mathbf{v_x}) \in \mathbb{T}_{\mathbf{s_x}}\mathbb{E}\,,$$

with $\mathbf{s_x} \in \mathbb{E}_{\mathbf{x}}$ *and* $\mathbf{u_x}, \mathbf{v_x} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ *and* $\alpha, \beta \in \Re$.

Lemma 1.4.6 (Horizontal lifts and horizontal projectors) *Given a horizontal projector* $P_{\mathrm{H}} \in \mathrm{C}^1(\mathbb{T}\mathbb{E}\,;\mathbb{T}\mathbb{E})$, *the induced horizontal lift is defined by*

$$\mathbf{H}(\mathbf{s_x}\,, \mathbf{v_x}) := P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v_x} \in \mathbb{H}_{\mathbf{s_x}}\mathbb{E}\,, \quad \forall\, \mathbf{s_x} \in \mathbb{E}_{\mathbf{x}}, \quad \mathbf{v_x} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,,$$

where $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *is an arbitrary section extension of* $\mathbf{s_x} \in \mathbb{E}_{\mathbf{x}}$. *Vice versa, given a horizontal lift* $\mathbf{H} \in \mathrm{C}^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{T}\mathbb{M}\,;\mathbb{T}\mathbb{E})$, *the corresponding horizontal projector is given by*

$$P_{\mathrm{H}} := \mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p})\,.$$

Proof. The former formula yields a horizontal lift since:

$$((\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \circ \mathbf{H})(\mathbf{s_x}\,, \mathbf{v_x}) = (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \cdot P_{\mathrm{H}} \cdot T_{\mathbf{x}}\mathbf{s} \cdot \mathbf{v_x} = (\mathbf{s_x}\,, \mathbf{v_x})\,.$$

The latter formula yields a horizontal projector since idempotency follows from:

$$P_{\mathrm{H}} \circ P_{\mathrm{H}} = \mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \circ \mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) = \mathbf{H} \circ \mathbf{id}_{\mathbb{E}\times_{\mathbb{M}}\mathbb{TM}} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) = P_{\mathrm{H}}\,.$$

Horizontality of $P_{\mathrm{H}} := \mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p})$ is expressed by the implication:

$$(\mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}))(\mathbf{X}) = \mathbf{0}_{\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})} \implies T\mathbf{p} \cdot \mathbf{X} = \mathbf{0}_{\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})}\,,$$

which is inferred from

$$(\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})\,, \mathbf{0}) = ((\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \circ \mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}))(\mathbf{X}) = (\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})\,, T\mathbf{p}(\mathbf{X}))\,,$$

and the result is proved. ■

1.4.2 Splitting of dual exact sequences

Definition 1.4.6 (Connnection and splitting) *A connnection on a fibre bundle* $(\mathbb{E}\,, \mathbf{p}\,, \mathbb{M})$ *is a splitting of the exact sequence (see section 1.2.10, page 47):*

$$0 \longrightarrow \mathbb{VE} \xrightarrow{\ \mathbf{i}\ } \mathbb{TE} \xrightarrow{(\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p})} \mathbf{p}{\downarrow}\mathbb{TM} \longrightarrow 0$$

where $\mathbf{i} \in \mathrm{C}^1(\mathbb{VE}\,; \mathbb{TE})$ *is the inclusion and* $(\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \in \mathrm{C}^1(\mathbb{TE}\,; \mathbf{p}{\downarrow}\mathbb{TM})$ *is the canonical surjection.*

By definition, a splitting $\mathbf{H} \in \mathrm{C}^1(\mathbf{p}{\downarrow}\mathbb{TM}\,; \mathbb{TE})$ is such that $(\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \circ \mathbf{H} = \mathbf{id}_{\,\mathbf{p}\downarrow\mathbb{TM}}$, that is:

$$((\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \circ \mathbf{H})(\mathbf{e}\,, \mathbf{v}) = (\mathbf{e}\,, \mathbf{v})\,,$$

with $(\mathbf{e}\,, \mathbf{v}) \in \mathbf{p}{\downarrow}\mathbb{TM} = \mathbb{E} \times_{\mathbb{M}} \mathbb{TM}$. Hence, the splitting $\mathbf{H} \in \mathrm{C}^1(\mathbf{p}{\downarrow}\mathbb{TM}\,; \mathbb{TE})$ is a horizontal lifting if it induces a linear homomorphism from the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,; \mathbb{M})$ to the tangent bundle $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{TE}\,; \mathbb{E})$.

Lemma 1.4.7 (Dual exact sequence) *A dual exact sequence can be associated with the previous one:*

$$0 \longrightarrow \mathbf{p}{\downarrow}\mathbb{TM}^* \xrightarrow{\ T^*\mathbf{p}\ } \mathbb{TE}^* \xrightarrow{\ \mathbf{j}\ } \mathbb{VE}^* \longrightarrow 0.$$

Proof. The canonical injection $T^*\mathbf{p} \in \mathrm{C}^1(\mathbf{p}{\downarrow}\mathbb{T}^*\mathbb{M}\,;\mathbb{TE}^*)$ is defined by:

$$T^*\mathbf{p}(\mathbf{e}\,,\mathbf{v}^*) := T^*_{\mathbf{e}}\mathbf{p}\cdot\mathbf{v}^*\,,$$

with $(\mathbf{e}\,,\mathbf{v}^*) \in \mathbf{p}{\downarrow}\mathbb{T}^*\mathbb{M} = \mathbb{E}\times_{\mathbb{M}}\mathbb{T}^*\mathbb{M}$. Then, for all $\mathbf{X_e} \in \mathbb{T}_{\mathbf{e}}\mathbb{E}$:

$$T^*\mathbf{p}(\mathbf{e}\,,\mathbf{v}^*)\cdot\mathbf{X_e} := \langle\, T^*_{\mathbf{e}}\mathbf{p}\cdot\mathbf{v}^*, \mathbf{X_e}\,\rangle = \langle\, \mathbf{v}^*, T_{\mathbf{e}}\mathbf{p}\cdot\mathbf{X_e}\,\rangle\,,$$

Let us now verify that the dual sequence is exact. The surjectivity of $T_{\mathbf{e}}\mathbf{p} \in BL(\mathbb{T}_{\mathbf{e}}\mathbb{E}\,;\mathbb{T}_{\mathbf{p(e)}}\mathbb{M})$ implies that $\mathbf{ker}(T^*_{\mathbf{e}}\mathbf{p}) = \mathbf{im}(T_{\mathbf{e}}\mathbf{p})^{\circ} = \{0\}$ and hence the injectivity of $T^*_{\mathbf{e}}\mathbf{p} \in BL(\mathbb{T}^*_{\mathbf{p(e)}}\mathbb{M}\,;\mathbb{T}^*_{\mathbf{e}}\mathbb{E})$. Moreover, by BANACH's closed range theorem, it implies also that $\mathbf{im}(T^*_{\mathbf{e}}\mathbf{p}) = \mathbf{ker}(T_{\mathbf{e}}\mathbf{p})^{\circ}$. On the other hand, the canonical surjection $\mathbf{j} \in \mathrm{C}^1(\mathbb{TE}^*\,;\mathbb{VE}^*)$ is the pointwise dual of $\mathbf{i} \in \mathrm{C}^1(\mathbb{VE}\,;\mathbb{TE})$ according to the identity:

$$\langle\, \mathbf{j}(\boldsymbol{\alpha}), \mathbf{V}\,\rangle = \langle\, \boldsymbol{\alpha}, \mathbf{i}(\mathbf{V})\,\rangle\,, \quad \forall\,(\boldsymbol{\alpha}\,,\mathbf{V}) \in \mathbb{TE}^*\times_{\mathbb{E}}\mathbb{VE}\,.$$

The linear space $\mathbb{V}^*_{\mathbf{e}}\mathbb{E}$, dual of the subspace $\mathbb{V}_{\mathbf{e}}\mathbb{E} \subset \mathbb{T}_{\mathbf{e}}\mathbb{E}$, is isometrically isomorphic to the quotient space $\mathbb{T}^*_{\mathbf{e}}\mathbb{E}/(\mathbb{V}_{\mathbf{e}}\mathbb{E})^{\circ}$, and hence $\mathbf{ker}(\mathbf{j}(\mathbf{e})) = (\mathbb{V}_{\mathbf{e}}\mathbb{E})^{\circ} = \mathbf{ker}(T_{\mathbf{e}}\mathbf{p})^{\circ} = \mathbf{im}(T^*_{\mathbf{e}}\mathbf{p})$. The surjectivity of $\mathbf{j} \in \mathrm{C}^1(\mathbb{TE}^*\,;\mathbb{VE}^*)$ follows from $\mathbf{im}(\mathbf{j}) = \mathbf{ker}(\mathbf{i})^{\circ}$ due to the closedness of $\mathbf{im}(\mathbf{i})$. ∎

A dual splitting $P^*_{\mathrm{V}} \in \mathrm{C}^1(\mathbb{VE}^*\,;\mathbb{TE}^*)$ is pointwise defined by the linear projector $P_{\mathrm{V}}(\mathbf{e})^* \in BL(\mathbb{V}^*_{\mathbf{e}}\mathbb{E}\,;\mathbb{T}^*_{\mathbf{e}}\mathbb{E})$ dual to the vertical projector $P_{\mathrm{V}}(\mathbf{e}) \in BL(\mathbb{T}_{\mathbf{e}}\mathbb{E}\,;\mathbb{V}_{\mathbf{e}}\mathbb{E})$ according to:

$$\langle\, \boldsymbol{\beta}, P_{\mathrm{V}}\cdot\mathbf{X}\,\rangle = \langle\, P^*_{\mathrm{V}}\cdot\boldsymbol{\beta}, \mathbf{X}\,\rangle\,,$$

with $\{\boldsymbol{\beta}\,,\mathbf{X}\} \in \mathbb{VE}^*\times_{\mathbb{E}}\mathbb{TE}$.

1.4.3 Horizontal lift in a vector bundle

Definition 1.4.7 (Linear connections) *In a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *a connection is linear if the pair made of the horizontal lift* $\mathbf{H_v} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{HE})$ *and of the vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is a linear vector bundle homomorphism from the vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *to the vector bundle* $T\mathbf{p} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TM})$ *:*

$$\begin{array}{ccc} \mathbb{E} & \xrightarrow{\mathbf{H_v}} & \mathbb{HE} \\ \mathbf{p}\big\downarrow & & \big\downarrow T\mathbf{p} \\ \mathbb{M} & \xrightarrow{\mathbf{v}} & \mathbb{TM} \end{array} \qquad \Longleftrightarrow \qquad T\mathbf{p}\circ\mathbf{H_v} = \mathbf{v}\circ\mathbf{p}\,,$$

so that additivity $\mathbf{H_{v_x}}(\mathbf{s}_1(\mathbf{x}) +_{\mathbf{p}} \mathbf{s}_2(\mathbf{x})) = \mathbf{H_{v_x}}(\mathbf{s}_1(\mathbf{x})) +_{T\mathbf{p}} \mathbf{H_{v_x}}(\mathbf{s}_2(\mathbf{x}))$ *and homogeneity* $\mathbf{H_{v_x}}(\alpha \cdot_{\mathbf{p}} \mathbf{s}(\mathbf{x})) = \alpha \cdot_{T\mathbf{p}} \mathbf{H_{v_x}}(\mathbf{s}(\mathbf{x}))$ *properties hold.*

Lemma 1.4.8 (Fiberwise $\Re$-linearity of horizontal lifts) *In a vector bundle $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ a horizontal projector $P_H \in C^1(\mathbb{TE};\mathbb{HE})$ defines a linear connection iff it is fibrewise linear in the bundle $T\mathbf{p} \in C^1(\mathbb{TE};\mathbb{TM})$:*

$$P_H \cdot (\mathbf{X}_1 +_{T\mathbf{p}} \mathbf{X}_2) = P_H \cdot \mathbf{X}_1 +_{T\mathbf{p}} P_H \cdot \mathbf{X}_2$$
$$P_H \cdot (\alpha \cdot_{T\mathbf{p}} \mathbf{X}) = \alpha \cdot_{T\mathbf{p}} \mathbf{X}\,.$$

Proof. Given two sections $\mathbf{s}_1, \mathbf{s}_2 \in C^1(\mathbb{M};\mathbb{E})$ of the vector bundle $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$, fibrewise additivity follows from the tensoriality of the horizontal lift and the equality:

$$\begin{aligned}\mathbf{H}_{\mathbf{v}}(\mathbf{s}_1(\mathbf{x})) +_{T\mathbf{p}} \mathbf{H}_{\mathbf{v}}(\mathbf{s}_2(\mathbf{x})) &= P_H \cdot T_{\mathbf{x}}\mathbf{s}_1 \cdot \mathbf{v}_{\mathbf{x}} +_{T\mathbf{p}} P_H \cdot T_{\mathbf{x}}\mathbf{s}_2 \cdot \mathbf{v}_{\mathbf{x}} \\ &= P_H \cdot (T_{\mathbf{x}}\mathbf{s}_1 \cdot \mathbf{v}_{\mathbf{x}} +_{T\mathbf{p}} T_{\mathbf{x}}\mathbf{s}_2 \cdot \mathbf{v}_{\mathbf{x}}) \\ &= P_H \cdot T_{\mathbf{x}}(\mathbf{s}_1 +_{\mathbf{p}} \mathbf{s}_2) \cdot \mathbf{v}_{\mathbf{x}} = \mathbf{H}_{\mathbf{v}}(\mathbf{s}_1(\mathbf{x}) +_{\mathbf{p}} \mathbf{s}_2(\mathbf{x}))\,,\end{aligned}$$

where the map $\mathbf{s}_1 +_{\mathbf{p}} \mathbf{s}_2 \in C^1(\mathbb{M};\mathbb{E})$ is pointwise defined by $(\mathbf{s}_1 +_{\mathbf{p}} \mathbf{s}_2)(\mathbf{x}) = \mathbf{s}_1(\mathbf{x}) +_{\mathbf{p}} \mathbf{s}_2(\mathbf{x}) \in \mathbb{E}_{\mathbf{x}}$. Fiberwise $\Re$-linearity is likewise inferred. ■

Corollary 1.4.1 (Horizontal isomorphism) *A horizontal lift in a vector bundle $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ is a vector bundle isomorphism $\mathbf{H} \in C^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{TM};\mathbb{HE})$ between the* Whitney *sum $(\mathbf{p}, \boldsymbol{\tau}_{\mathbb{M}}) \in C^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{TM};\mathbb{M})$ and the horizontal bundle $\boldsymbol{\tau}_{\mathbb{E}} \in C^1(\mathbb{HE};\mathbb{E})$.*

Lemma 1.4.9 (Spray of a linear connection) *In the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM};\mathbb{M})$ let $\mathbf{H} \in C^1(\mathbb{TM} \times_{\mathbb{M}} \mathbb{TM};\mathbb{T}^2\mathbb{M})$ be the horizontal lift induced by a linear connection. Then, denoting by $\text{DIAG} \in C^1(\mathbb{TM};\mathbb{TM} \times_{\mathbb{M}} \mathbb{TM})$ the map $\text{DIAG} := (\mathbf{id}_{\mathbb{TM}}, \mathbf{id}_{\mathbb{TM}})$, the tangent vector field $\mathbf{S} \in C^1(\mathbb{TM};\mathbb{T}^2\mathbb{M})$ defined by:*

$$\mathbf{S} := \mathbf{H} \circ \text{DIAG} \iff \mathbf{S}(\mathbf{v}) := \mathbf{H}(\mathbf{v}, \mathbf{v})\,, \quad \forall\, \mathbf{v} \in \mathbb{TM}\,,$$

is a quadratic spray, that is: $T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{S} = \mathbf{id}_{\mathbb{TM}}$ or $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{S} = \mathbf{S}$ and

$$\mathbf{S} \circ \boldsymbol{mult}^t_{\mathbb{TM}} = T(\boldsymbol{mult}^t_{\mathbb{TM}}) \circ \boldsymbol{mult}^t_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{S}\,, \quad \forall\, t \in \Re\,.$$

Any connection $\mathbf{H} \in C^1(\mathbb{TM} \times_{\mathbb{M}} \mathbb{TM};\mathbb{T}^2\mathbb{M})$ such that $\mathbf{S} := \mathbf{H} \circ \text{DIAG}$ is said to be compatible with the spray.

Proof. The latter formula follows from the relation:

$$\begin{aligned}\mathbf{H}(\alpha\mathbf{u}\,,\beta\mathbf{v}) = P_{\mathrm{H}}(T(\alpha\mathbf{u})\cdot\beta\mathbf{v}) &= (P_{\mathrm{H}}\circ T(\mathbf{mult}\,^{\alpha}_{\mathbb{TM}}\circ\mathbf{u}))\cdot\mathbf{mult}\,^{\beta}_{\mathbb{TM}}(\mathbf{v})\\ &= (P_{\mathrm{H}}\circ\mathbf{mult}\,^{\beta}_{\mathbb{T}^2\mathbb{M}}\circ T(\mathbf{mult}\,^{\alpha}_{\mathbb{TM}}\circ\mathbf{u}))\cdot\mathbf{v}\\ &= (\mathbf{mult}\,^{\beta}_{\mathbb{T}^2\mathbb{M}}\circ P_{\mathrm{H}}\circ T(\mathbf{mult}\,^{\alpha}_{\mathbb{TM}}\circ\mathbf{u}))\cdot\mathbf{v}\\ &= (\mathbf{mult}\,^{\beta}_{\mathbb{T}^2\mathbb{M}}\circ P_{\mathrm{H}}\circ T\mathbf{mult}\,^{\alpha}_{\mathbb{TM}}\circ T\mathbf{u})\cdot\mathbf{v}\\ &= (T\mathbf{mult}\,^{\alpha}_{\mathbb{TM}}\circ\mathbf{mult}\,^{\beta}_{\mathbb{T}^2\mathbb{M}}\circ P_{\mathrm{H}}\circ T\mathbf{u})\cdot\mathbf{v}\\ &= (T\mathbf{mult}\,^{\alpha}_{\mathbb{TM}}\circ\mathbf{mult}\,^{\beta}_{\mathbb{T}^2\mathbb{M}}\circ\mathbf{H})(\mathbf{u}\,,\mathbf{v})\,.\end{aligned}$$

The former formula is trivial. ■

Definition 1.4.8 (Symmetric connection) *A linear connection is symmetric if the horizontal lift fulfils the condition:*

$$\mathbf{H} = \mathbf{k}_{\mathbb{T}^2\mathbb{M}}\circ\mathbf{H}\circ\mathbf{flip}_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}}\,,$$

with $\mathbf{flip}_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}}$ *the involution on* $\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}$ *defined by*

$$\mathbf{flip}_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}}(\mathbf{s_x}\,,\mathbf{v_x}) = (\mathbf{v_x}\,,\mathbf{s_x})\,,\quad \forall\,\mathbf{s_x},\mathbf{v_x}\in\mathbb{T}_{\mathbf{x}}\mathbb{M}\,.$$

1.4.4 Covariant derivative

Definition 1.4.9 *The* ***covariant derivative*** $\nabla_{\mathbf{v}}\in\mathrm{C}^1(\mathbb{E}\,;\mathbb{VE})$ *is the vertical component of the natural derivative, defined by*

$$\nabla_{\mathbf{v}}\mathbf{s} := P_{\mathrm{V}}\circ T_{\mathbf{v}}\mathbf{s}\in\mathrm{C}^1(\mathbb{M}\,;\mathbb{VE})\,,$$

where $\mathbf{v}\in\mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *and* $\mathbf{s}\in\mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *is a section of the fibre bundle* $\mathbf{p}\in\mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *and we write* $\nabla\mathbf{s}\circ\mathbf{v}$ *for* $\nabla_{\mathbf{v}}\mathbf{s}$ *with* $\nabla\mathbf{s}\in\mathrm{C}^1(\mathbb{TM}\,;\mathbb{VE})$ *.*

Then $T\mathbf{s} = \nabla\mathbf{s}+\mathbf{Hs}\in\mathrm{C}^1(\mathbb{TM}\,;\mathbb{TE})$ and $T_{\mathbf{v}} = \nabla_{\mathbf{v}}+\mathbf{H_v}\in\mathrm{C}^1(\mathbb{E}\,;\mathbb{TE})$.

Lemma 1.4.10 (Covariant derivative as a generalized Lie derivative) *The covariant derivative of a section* $\mathbf{s}\in\mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *of the fibre bundle* $\mathbf{p}\in\mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *may be defined as a generalized* LIE *derivative:*

$$\nabla_{\mathbf{v}}\mathbf{s} := \mathcal{L}_{(\mathbf{H_v}\,,\mathbf{v})}\mathbf{s} = \partial_{\lambda=0}\,\mathbf{Fl}^{(\mathbf{H_v}\,,\mathbf{v})}_{\lambda}\!\downarrow\!\mathbf{s} = \partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{H_v}}_{-\lambda}\circ\mathbf{s}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\,.$$

Proof. By LEIBNIZ rule $\mathcal{L}_{(\mathbf{H_v},\mathbf{v})}\mathbf{s} = T\mathbf{s} \circ \mathbf{v} - \mathbf{H_v s} = T_\mathbf{v}\mathbf{s} - \mathbf{H_v s}$. Then, being $\mathcal{L}_{(\mathbf{H_v},\mathbf{v})}\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{VE})$ and $\mathbf{H_v s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{HE})$, by uniqueness of the vertical-horizontal split, we get that $\nabla_\mathbf{v}\mathbf{s} := P_\mathrm{V} \circ T_\mathbf{v}\mathbf{s} = \mathcal{L}_{(\mathbf{H_v},\mathbf{v})}\mathbf{s}$. ∎

Definition 1.4.10 (Geodesic) *A curve* $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{M})$ *in a manifold with a connection is a geodesic if the velocity field of the curve* $\mathbf{v} \in \mathrm{C}^0(\mathbf{c}(I)\,;\mathbb{TM})$ *fulfils the condition*

$$\nabla_{\mathbf{v_x}}\mathbf{v} = 0$$

for all $\mathbf{x} \in \mathbf{c}(I)$.

Lemma 1.4.11 (Geodesic of a quadratic spray) *Let* $\mathbf{S} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *be a quadratic spray and* $\mathbf{v_x} \in \mathbb{T_x M}$ *a tangent vector. Then the base curve:*

$$\mathrm{GEO}_\lambda(\mathbf{v_x}) := (\boldsymbol{\tau}_\mathbb{M} \circ \mathbf{Fl}^\mathbf{S}_\lambda)(\mathbf{v_x})\,,$$

below the flow-line of the spray through $\mathbf{v_x} \in \mathbb{T_x M}$ *is a geodesic for any connection compatible with the quadratic spray and fulfils the properties:*

$$\begin{aligned}
&\mathrm{GEO}_0(\mathbf{v_x}) = \mathbf{x}\,, \quad \partial_{\lambda=0}\,\mathrm{GEO}_\lambda(\mathbf{v_x}) = \mathbf{v_x}\,,\\
&\mathrm{GEO}_\lambda(\alpha\mathbf{v_x}) = \mathrm{GEO}_{\alpha\lambda}(\mathbf{v_x})\,,\\
&\mathrm{GEO}_\lambda(\partial_{t=\mu}\,\mathrm{GEO}_t(\mathbf{v_x})) = \mathrm{GEO}_{\lambda+\mu}(\mathbf{v_x})\,.
\end{aligned}$$

Given a geodesic GEO *the spray can be evaluated as* $\mathbf{S} = \partial_{\lambda=0}\,\partial_{\mu=\lambda}\,\mathrm{GEO}_\mu$.

Proof. The velocity field of the curve $\mathbf{c}_\lambda(\mathbf{x}) = (\boldsymbol{\tau}_\mathbb{M} \circ \mathbf{Fl}^\mathbf{S}_\lambda)(\mathbf{v_x})$ is given by

$$\begin{aligned}
(\mathbf{v} \circ \boldsymbol{\tau}_\mathbb{M} \circ \mathbf{Fl}^\mathbf{S}_\lambda)(\mathbf{v_x}) &= \mathbf{v}(\mathbf{c}_\lambda(\mathbf{x})) = \partial_{\mu=\lambda}\,\mathbf{c}_\mu(\mathbf{x})\\
&= \partial_{\mu=\lambda}\,(\boldsymbol{\tau}_\mathbb{M} \circ \mathbf{Fl}^\mathbf{S}_\mu)(\mathbf{v_x}) = T\boldsymbol{\tau}_\mathbb{M} \cdot \mathbf{S}(\mathbf{Fl}^\mathbf{S}_\lambda(\mathbf{v_x})) = \mathbf{Fl}^\mathbf{S}_\lambda(\mathbf{v_x})\,.
\end{aligned}$$

Then, being $\mathbf{Fl}^\mathbf{v}_\lambda(\mathbf{x}) = \mathbf{c}_\lambda(\mathbf{x})$, the formula for the covariant derivative yields:

$$\begin{aligned}
\nabla_{\mathbf{v_x}}\mathbf{v} &= \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{H_v}}_{-\lambda} \circ \mathbf{v} \circ \mathbf{Fl}^\mathbf{v}_\lambda)(\mathbf{x})\\
&= \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{H_v}}_{-\lambda} \circ \mathbf{v} \circ \boldsymbol{\tau}_\mathbb{M} \circ \mathbf{Fl}^\mathbf{S}_\lambda)(\mathbf{v_x})\\
&= \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{H_v}}_{-\lambda} \circ \mathbf{Fl}^\mathbf{S}_\lambda)(\mathbf{v_x}) = \mathbf{S}(\mathbf{v_x}) - \mathbf{H}_{\mathbf{v_x}}(\mathbf{v_x})\,.
\end{aligned}$$

The curve $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{M})$ is a geodesic iff $\mathbf{S}(\mathbf{v_x}) = \mathbf{H}_{\mathbf{v_x}}(\mathbf{v_x})$. ∎

Definition 1.4.11 (Geodesic exponential) *To a given geodesic there corresponds an exponential map* $\text{EXP} \in C^1(\mathbb{TM}\,;\mathbb{M})$ *defined by*

$$\text{EXP}\,(\mathbf{v_x}) := \text{GEO}_1(\mathbf{v_x})\,,$$

in an open neighborhood of the zero section in $\mathbb{TM}$ *.*

The exponential map fulfils the properties:

$$\begin{aligned}&\text{EXP}\,(t\mathbf{v_x}) = \text{GEO}_t(\mathbf{v_x})\,,\\ &T_{0_\mathbf{x}}\text{EXP}\,\cdot\mathbf{v_x} = \mathbf{v_x}\,,\quad \forall\,\mathbf{v_x}\in\mathbb{T_x M}\quad (\mathbb{T}_{0_\mathbf{x}}\mathbb{T_x M}\simeq\mathbb{T_x M})\,,\end{aligned}$$

since

$$\begin{aligned}T_{0_\mathbf{x}}\text{EXP}\,\cdot\mathbf{v_x} &= \partial_{t=0}\,\text{EXP}\,(0_\mathbf{x}+t\mathbf{v_x}) = \partial_{t=0}\,\text{EXP}\,(t\mathbf{v_x})\\ &= \partial_{t=0}\,\text{GEO}_1(t\mathbf{v_x}) = \partial_{t=0}\,\text{GEO}_t(\mathbf{v_x}) = \mathbf{v_x}\,.\end{aligned}$$

Then the map $(\boldsymbol{\tau}_\mathbb{M}\,,\text{EXP}\,) \in C^1(\mathbb{TM}\,;\mathbb{M}\times\mathbb{M})$ is a diffeomorphism from an open neighborhood of the zero section in $\mathbb{TM}$ to an open neighborhood of the diagonal in $\mathbb{M}\times\mathbb{M}$.

1.4.5 Frobenius integrability theorem

The integrability theorem of FROBENIUS concerns a local vector sub-bundle $\mathcal{A}$ of $\mathbb{TM}$, called a *distribution*, with n-D base manifold $\mathbb{M}$ and fibres which are linear k-D subspaces of the tangent spaces to $\mathbb{M}$ such that in the neighborhood of a point of $\mathbb{M}$ there a family of k vector fields which form a frame for the local vector bundle. Such a family is called a *local basis.*

Figure 1.18: Ferdinand Georg Frobenius (1849 - 1917)

Definition 1.4.12 *The vector subbundle* $\mathcal{A}$ *is said to be* ***integrable*** *at* $\mathbf{x} \in \mathbb{M}$ *if there exists a (local) submanifold, the* ***integral manifold*** $\mathbb{I}_{\mathcal{A}} \subset \mathbb{M}$ *through* $\mathbf{x}$*, such that its tangent manifold is the subbundle* $\mathcal{A}$ *restricted to* $\mathbb{I}_{\mathcal{A}}$*.*

In terms of the *inclusion operator* $\mathbf{i} \in \mathrm{C}^1(\mathbb{I}_{\mathcal{A}}\,;\mathbb{M})$, the integral manifold is characterized by: $\mathbb{T}(\mathbf{i}(\mathbb{I}_{\mathcal{A}})) = (\mathcal{A} \circ \mathbf{i})(\mathbb{I}_{\mathcal{A}})$. Equivalently, integrability requires the existence of a local chart for $\mathbb{M}$ such that the velocities of k of the n coordinate lines form a local basis for the vector sub-bundle. Such a chart is called a *flat chart* for the local vector sub-bundle. We develop here an original approach to the proof of Frobenius integrability condition which is the simplest to be grasped in its geometrical aspects.

Let us provide a propædeutic result (see [80], Theorem 3.17).

Lemma 1.4.12 (Local frames and coordinates) *Let* $\mathbb{M}$ *be a manifold modeled on a* n*-D linear space* E *and* $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *be a set of vector fields* $\mathbf{v}_i \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *in a neighborhood of* $\mathbf{x} \in \mathbb{M}$ *such that* $\{\mathbf{v}_1(\mathbf{x}), \dots, \mathbf{v}_n(\mathbf{x})\}$ *is a frame at* $\mathbf{x} \in \mathbb{M}$ *with* $[\mathbf{v}_i\,, \mathbf{v}_j] = 0$ *for all* $i, j = 1, \dots, n$*. Then the vector fields* $\mathbf{v}_i$ *are the velocities of the coordinate lines associated with a coordinate map* $\varphi \in \mathrm{C}^1(E\,;\mathbb{M})$ *centered at* $\mathbf{x} \in \mathbb{M}$*.*

Proof. By Proposition 1.3.2 the flows of the vector fields $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ commute pairwise. Then, if $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is a basis of E and $\mathbf{t} = \sum_{i=1}^n t_i \mathbf{e}_i$, we may set

$$\begin{aligned} \boldsymbol{\varphi}(\mathbf{t}) &:= (\mathbf{Fl}^{\mathbf{v}_1}_{t_1} \circ \mathbf{Fl}^{\mathbf{v}_2}_{t_2} \circ \dots \circ \mathbf{Fl}^{\mathbf{v}_n}_{t_n})(\mathbf{x})\,, \quad \mathbf{t} \in E\,, \\ T_{\mathbf{e}_i}\boldsymbol{\varphi}(\mathbf{t}) &= \partial_{\tau_i = t_i}\, \mathbf{Fl}^{\mathbf{v}_i}_{\tau_i} \circ (\mathbf{Fl}^{\mathbf{v}_1}_{t_1} \circ \mathbf{Fl}^{\mathbf{v}_2}_{t_2} \circ \dots \circ \mathbf{Fl}^{\mathbf{v}_n}_{t_n})_i(\mathbf{x}) \\ &= \mathbf{v}_i(\mathbf{Fl}^{\mathbf{v}_1}_{t_1} \circ \mathbf{Fl}^{\mathbf{v}_2}_{t_2} \circ \dots \circ \mathbf{Fl}^{\mathbf{v}_n}_{t_n})(\mathbf{x})\,, \end{aligned}$$

with $\boldsymbol{\varphi}(0) = \mathbf{x}$ and $T_{\mathbf{e}_i}\boldsymbol{\varphi}(0) = \mathbf{v}_i(\mathbf{x})$. The subscript $()_i$ denotes that the i-th term is missing. ∎

The proof of Frobenius sufficient condition of integrability will be carried out with reference to a horizontal subbundle of a fibre bundle. Subsequently, this result is readily adapted to deal with the general case of a distribution on a manifold. Preliminarily we prove that an integrable subbundle of a tangent bundle $\mathbb{TM}$ is involutive, according to the following definition.

Definition 1.4.13 *A vector subbundle* $\mathcal{A}$ *of the tangent bundle* $\mathbb{TM}$ *is said to be* ***involutive*** *if, for any pair of vector fields of* $\mathcal{A}$*, their bracket belongs to* $\mathcal{A}$*.*

Lemma 1.4.13 (Necessity of the involutivity property) *An integrable vector subbundle* $\mathcal{A}$ *of the tangent bundle* $\mathbb{TM}$ *is involutive.*

Proof. Since $\mathcal{A}$ is integrable, we have that $(\mathcal{A} \circ \mathbf{i})(\mathbb{I}_{\mathcal{A}}) = \mathbb{T}(\mathbf{i}(\mathbb{I}_{\mathcal{A}})) = T\mathbf{i}(\mathbb{T}\mathbb{I}_{\mathcal{A}})$, where $\mathbf{i} \in \mathrm{C}^1(\mathbb{I}_{\mathcal{A}}\,;\mathbb{M})$ is the inclusion. Then the property $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbf{i}(\mathbb{I}_{\mathcal{A}})\,;\mathcal{A})$ is equivalent to require their $\mathbf{i}$-relatedness to vector fields $\mathbf{u}_{\mathcal{A}}, \mathbf{v}_{\mathcal{A}} \in \mathrm{C}^1(\mathbb{I}_{\mathcal{A}}\,;\mathbb{T}\mathbb{I}_{\mathcal{A}})$ according to the commutative diagram

$$\begin{array}{ccc} \mathbb{T}\mathbb{I}_{\mathcal{A}} & \xrightarrow{T\mathbf{i}} & \mathbb{T}\mathbb{M} \\ {\scriptstyle \mathbf{u}_{\mathcal{A}},\mathbf{v}_{\mathcal{A}}}\uparrow & & \uparrow{\scriptstyle \mathbf{u},\mathbf{v}} \\ \mathbb{I}_{\mathcal{A}} & \xrightarrow{\mathbf{i}} & \mathbb{M} \end{array} \qquad \Longleftrightarrow \qquad \begin{cases} \mathbf{u} \circ \mathbf{i} = T\mathbf{i} \circ \mathbf{u}_{\mathcal{A}} \in \mathrm{C}^0(\mathbb{I}_{\mathcal{A}}\,;\mathbb{T}\mathbb{M})\,, \\ \mathbf{v} \circ \mathbf{i} = T\mathbf{i} \circ \mathbf{v}_{\mathcal{A}} \in \mathrm{C}^0(\mathbb{I}_{\mathcal{A}}\,;\mathbb{T}\mathbb{M})\,. \end{cases}$$

By Proposition 1.3.4 on page 64 the bracket $[\mathbf{u}\,,\mathbf{v}] \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ is $\mathbf{i}$-related to $[\mathbf{u}_{\mathcal{A}}\,,\mathbf{v}_{\mathcal{A}}] \in \mathrm{C}^1(\mathbb{I}_{\mathcal{A}}\,;\mathbb{T}\mathbb{I}_{\mathcal{A}})$. Hence $[\mathbf{u}\,,\mathbf{v}] \in \mathrm{C}^1(\mathbf{i}(\mathbb{I}_{\mathcal{A}})\,;\mathcal{A})$. ■

Theorem 1.4.2 (Frobenius theorem for horizontal subbundles) *A horizontal subbundle* $\mathbb{H}\mathbb{E}$ *of the tangent bundle* $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{T}\mathbb{E}\,;\mathbb{E})$ *to a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *is integrable if it is involutive, i.e.*

$$\mathbf{X}, \mathbf{Y} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{H}\mathbb{E}) \implies [\mathbf{X}\,,\mathbf{Y}] \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{H}\mathbb{E})\,.$$

Proof. For any pair of vector fields $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ such that $[\mathbf{u}\,,\mathbf{v}] = 0$ the bracket of the horizontal lifts $\mathbf{H}_{\mathbf{u}}, \mathbf{H}_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{T}\mathbb{E})$ is a vertical field since, by Lemma 1.4.4, $T\mathbf{p} \circ [\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}] = [\mathbf{u}\,,\mathbf{v}] \circ \mathbf{p} = 0$. By the involutivity assumption it is also horizontal and hence vanishes. Given a set of coordinate lines on $\mathbb{M}$ with velocities $\mathbf{v}_i$, $i = 1, \ldots, n$ we have that $[\mathbf{v}_i\,,\mathbf{v}_j] = 0$, $i, j = 1 \ldots, n$ and also $[\mathbf{H}_{\mathbf{v}_i}\,,\mathbf{H}_{\mathbf{v}_j}] = 0$, $i, j = 1, \ldots, n$. Then, as in Lemma 1.4.12, the map $\boldsymbol{\varphi} \in \mathrm{C}^1(E\,;\mathbb{E})$ defined by $\boldsymbol{\varphi}(\mathbf{t}) = (\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}_1}}_{t_1} \circ \mathbf{Fl}^{\mathbf{H}_{\mathbf{v}_2}}_{t_2} \circ \ldots \circ \mathbf{Fl}^{\mathbf{H}_{\mathbf{v}_n}}_{t_n})(\mathbf{e})$ transforms an open neighborhood $U(0) \subset E$ in a submanifold $\boldsymbol{\varphi}(U(0)) \subset \mathbb{E}$ which is the horizontal leaf passing through $\boldsymbol{\varphi}(0) = \mathbf{e} \in \mathbb{E}$. ■

The integral manifolds provide a *foliation* of $\mathbb{E}$ into a family of disjoint connected horizontal *leaves* [3], [80].

Theorem 1.4.3 (Frobenius theorem for distributions) *A vector subbundle* $\mathcal{A}$ *of the tangent bundle* $\mathbb{T}\mathbb{M}$ *is integrable if it is involutive.*

Proof. The proof is directly inferred from Theorem 1.4.2 by the following trick. Let us consider a decomposition of the model linear space E of the manifold $\mathbb{M}$ into two supplementary linear subspaces: $E = H + V$ with $H := T_{\mathbf{m}_0}\varphi(\mathcal{A}(\mathbf{m}_0))$ for a chart $\varphi \in \mathrm{C}^1(U(\mathbf{m}_0)\,;E)$ with $U(\mathbf{m}_0) \subset \mathbb{M}$ open neighborhood of $\mathbf{m}_0 \in \mathbb{M}$. Denoting by $\mathbf{P}_H \in BL(E\,;H)$ the linear projector on the subspace H, we define the vector bundle $\mathbf{p} := \mathbf{P}_H \circ \varphi \in \mathrm{C}^1(\mathbb{M}\,;H)$ with typical fibre V.

The vertical subbundle of the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}\,;\mathbb{M})$ is then given by $T\varphi^{-1}(V)$ since $T\mathbf{p} \circ T\varphi^{-1}(V) = \mathbf{P}_H \circ T\varphi \circ T\varphi^{-1}(V) = \mathbf{P}_H(V) = 0$. The horizontal bundle is taken to be $\mathcal{A}$. ■

Subbundles with 1D fibres are always integrable. Indeed, the involutive property is trivially verified since $\mathbf{Y} = f \cdot \mathbf{X}$ with $f \in C^1(\mathbb{M}\,;\Re)$, so that

$$[\mathbf{X}\,,\mathbf{Y}] = [\mathbf{X}, f \cdot \mathbf{X}\,,=]\mathcal{L}_{\mathbf{X}} f \cdot \mathbf{X} + f \cdot \mathcal{L}_{\mathbf{X}}\mathbf{X} = \mathcal{L}_{\mathbf{X}} f \cdot \mathbf{X} \in \mathbb{TI}_{\mathcal{A}}\,.$$

Various proofs of sufficiency of the involutive property for integrability of a subbundle are given in the literature. The case of even non finite-dimensional BANACH spaces is considered in [38] Theorem 10.9.4, where FROBENIUS theorem is formulated as an integrability condition for a total differential equations.

In the same general context, FROBENIUS theorem is proved in [3] Theorem 4.4.3, as an integrability condition for subbundles of manifolds modeled on BANACH spaces. In the latter proof the role, which in the finite dimensional context is played by coordinate maps, is instead played by a skillful application of the LIE *transform method*. Other proofs are given in [80] section 3.23, in [91] chapter VI and in [27] chapter V.

FROBENIUS theorem can also be stated as an integrability condition for a total differential equation:

$$\mathbf{y}' = \mathbf{f}(\mathbf{x}, \mathbf{y})\,.$$

Here H, V are BANACH spaces, $\{\mathbf{x}\,,\mathbf{y}\} \in H \times V$ and $\mathbf{f}(\mathbf{x}, \mathbf{y}) \in BL\,(H\,;V)$ a bounded linear map. A solution is a differentiable map $\mathbf{u} \in C^1(U_H\,;U_V)$, with $U_H \subset H$ and $U_V \subset V$ open subsets, such that

$$\nabla\mathbf{u}(\mathbf{x}) = \mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}))\,.$$

The total differential equation is *completely integrable* in $U_H \times U_V \subset H \times V$ if at any point $\{\mathbf{x}_0\,,\mathbf{y}_0\} \in U_H \times U_V$ there is an open neighborhood $U(\mathbf{x}_0)$ of $\mathbf{x}_0 \in U_H$ such that there is a unique solution $\mathbf{u} \in C^1(U(\mathbf{x}_0)\,;U_V)$ fulfilling the condition $\mathbf{u}(\mathbf{x}_0) = \mathbf{y}_0$.

The equivalence with the integrability problem in Theorem 1.4.3 is revealed by the following observation. If $H := T_{\mathbf{m}_0}\varphi(\mathcal{A}(\mathbf{m}_0))$, then $T_{\mathbf{m}}\varphi(\mathcal{A}(\mathbf{m})) = \{(\mathbf{h}\,,\mathbf{f}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{h}) \mid \mathbf{h} \in H\}$ for any $\mathbf{m} \in U(\mathbf{m}_0)$ with $\{\mathbf{x}\,,\mathbf{y}\} = \varphi(\mathbf{m})$ according to a chart $\varphi \in C^1(U(\mathbf{m}_0)\,;H \times V)$.

Thus $\mathbf{y} = \mathbf{u}(\mathbf{x})$ is a parametric representation of the integral manifold in the model BANACH space $E = H \times V$. Setting $X_{\mathbf{h}} := \{(\mathbf{h}\,,\mathbf{f}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{h})\}$, we have that

$$\begin{aligned}[X_{\mathbf{h}_1}\,, X_{\mathbf{h}_2}] &= \nabla_{X_{\mathbf{h}_1}} X_{\mathbf{h}_2} - \nabla_{X_{\mathbf{h}_2}} X_{\mathbf{h}_1} \\ &= (0\,, \nabla_{X_{\mathbf{h}_1}}\mathbf{f}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{h}_2 - \nabla_{X_{\mathbf{h}_2}}\mathbf{f}(\mathbf{x}, \mathbf{y}) \cdot \mathbf{h}_1)\,,\end{aligned}$$

and the involutivity condition writes as in [3], Theorem 4.4.3:

$$\nabla_{X_{\mathbf{h}_1}} \mathbf{f}(\mathbf{x},\mathbf{y}) \cdot \mathbf{h}_2 = \nabla_{X_{\mathbf{h}_2}} \mathbf{f}(\mathbf{x},\mathbf{y}) \cdot \mathbf{h}_1 \,,$$

or more explicitly as in [38], Theorem 10.9.4:

$$\nabla_{\mathbf{h}_1} \mathbf{f}_{\mathbf{y}}(\mathbf{x}) \cdot \mathbf{h}_2 + \nabla_{\mathbf{f}(\mathbf{x},\mathbf{y})\cdot\mathbf{h}_1} \mathbf{f}_{\mathbf{x}}(\mathbf{y}) \cdot \mathbf{h}_2 = \nabla_{\mathbf{h}_2} \mathbf{f}_{\mathbf{y}}(\mathbf{x}) \cdot \mathbf{h}_1 + \nabla_{\mathbf{f}(\mathbf{x},\mathbf{y})\cdot\mathbf{h}_2} \mathbf{f}_{\mathbf{x}}(\mathbf{y}) \cdot \mathbf{h}_1 \,,$$

which expresses the symmetry of the second derivative of the solution map $\mathbf{u} \in \mathrm{C}^1(U_H\,; U_V)$. Indeed, being $\nabla \mathbf{u}(\mathbf{x}) = \mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x}))$, we have that

$$\begin{aligned}\nabla^2_{\mathbf{h}_1,\mathbf{h}_2} \mathbf{u}(\mathbf{x}) &= \nabla_{\mathbf{h}_2} (\mathbf{f}(\mathbf{x}, \mathbf{u}(\mathbf{x})) \cdot \mathbf{h}_1) \\ &= \nabla_{\mathbf{h}_2} \mathbf{f}_{\mathbf{u}(\mathbf{x})}(\mathbf{x}) \cdot \mathbf{h}_1 + \nabla_{\nabla_{\mathbf{h}_2} \mathbf{u}(\mathbf{x})} \mathbf{f}_{\mathbf{x}}(\mathbf{u}(\mathbf{x})) \cdot \mathbf{h}_1 \\ &= \nabla_{\mathbf{h}_2} \mathbf{f}_{\mathbf{u}(\mathbf{x})}(\mathbf{x}) \cdot \mathbf{h}_1 + \nabla_{\mathbf{f}(\mathbf{x},\mathbf{u}(\mathbf{x}))\cdot\mathbf{h}_2} \mathbf{f}_{\mathbf{x}}(\mathbf{u}(\mathbf{x})) \cdot \mathbf{h}_1 \,.\end{aligned}$$

Frobenius theorem provides the basis for the introduction of the notion of curvature of a connection on a fibre bundle [171].

1.4.6 Curvature of a connection

The next proposition states that the vertical subbundle of a fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ is integrable. The leaves of the induced foliation are the fibres of the bundle.

Proposition 1.4.1 (Integrability of the vertical subbundle) *The vertical subbundle* $\mathbb{VE} := \mathbf{ker}(T\mathbf{p})$ *of a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *is integrable.*

Proof. By definition, vertical vector fields are projectable to zero. Then Lemma 1.3.4 tells us that

$$T\mathbf{p} \circ [\mathbf{V}_1\,, \mathbf{V}_2] = 0\,, \quad \forall\, \mathbf{V}_1, \mathbf{V}_2 \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{VE} = \mathbf{ker}(T\mathbf{p}))\,.$$

and integrability follows from Frobenius Theorem 1.4.3. ∎

The integrability of the vertical subbundle may be expressed by the vanishing of the *cocurvature*:

$$\mathbf{R^c}(\mathbf{X}, \mathbf{Y}) := P_{\mathrm{H}} \circ [P_{\mathrm{V}}\mathbf{X}\,, P_{\mathrm{V}}\mathbf{Y}] = 0\,, \quad \forall\, \mathbf{X}, \mathbf{Y} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{TE})\,.$$

Frobenius Theorem 1.4.2 provides the following (necessary and sufficient) involutivity condition for the integrability of the horizontal subbundle of a fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ in which a connection has been fixed, [80]:

$$[P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}] \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{HE})\,, \quad \forall\, \mathbf{X}, \mathbf{Y} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{TE})\,.$$

equivalently expressed by the vanishing of the *curvature*:

$$\mathbf{R}(\mathbf{X},\mathbf{Y}) := P_{\mathrm{V}} \circ [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}] = 0\,, \quad \forall\, \mathbf{X},\mathbf{Y} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{TE})\,.$$

Proposition 1.4.2 (Tensoriality of the curvature) *The curvature of a connection* $P_{\mathrm{V}} \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$ *in a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *is tensorial. It is a vertical-valued horizontal two-form* $\mathbf{R} \in \Lambda^2(\mathbb{E}\,;\mathbb{VE})$*, that is a two-form vanishing on vertical vectors and taking values in the vertical bundle.*

Proof. From the vanishing of the *cocurvature* we get that $\frac{1}{2}[P_{\mathrm{V}}\,, P_{\mathrm{V}}] = \mathbf{R}$ and hence tensoriality follows from the tensoriality of the FN bracket. A direct verification, based on Lemma 1.2.1, also yields the result:

$$\begin{aligned}\mathbf{R}(\mathbf{X}, f\mathbf{Y}) &:= P_{\mathrm{V}} \circ [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}} f\mathbf{Y}] \\ &= f\, P_{\mathrm{V}} \circ [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}] + (\mathcal{L}_{P_{\mathrm{H}}\mathbf{X}} f)\,(P_{\mathrm{V}} \circ P_{\mathrm{H}})(\mathbf{Y}) \\ &= f\,\mathbf{R}(\mathbf{X},\mathbf{Y})\,, \quad \forall\, f \in \mathrm{C}^1(\mathbb{E}\,;\Re)\,,\end{aligned}$$

since $P_{\mathrm{V}} \circ P_{\mathrm{H}} = 0$. Similarly $\mathbf{R}(f\mathbf{X},\mathbf{Y}) = f\,\mathbf{R}(\mathbf{X},\mathbf{Y})$. ■

Theorem 1.4.4 (Curvature tensor in terms of horizontal lifts) *The curvature of a connection* $P_{\mathrm{V}} \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$ *in a fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$*, is expressed in terms of vector fields* $\mathbf{u},\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *on the tangent bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ *by setting:*

$$\mathrm{CURV}(\mathbf{s})(\mathbf{v},\mathbf{u}) := [\mathbf{H}_{\mathbf{u}}\,, \mathbf{H}_{\mathbf{v}}](\mathbf{s}) - \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]}(\mathbf{s})\,,$$

with $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *section of the fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$*. The differential two-form* $\mathrm{CURV}(\mathbf{s}) \in \Lambda^2(\mathbb{M}\,;\mathbb{VE})$ *is tensorial in* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$*.*

Proof. We rely on the properties of tensoriality and horizontality of the curvature two-form $\mathbf{R} \in \Lambda^2(\mathbb{E}\,;\mathbb{VE})$ stated in Proposition 1.4.2 and on the tensorial isomorphism of the horizontal liftings stated in Theorem 1.4.1.

Accordingly, the point value of the curvature $\mathbf{R}(\mathbf{X},\mathbf{Y}) := P_{\mathrm{V}} \circ [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}]$ at $\mathbf{b} \in \mathbb{E}_{\mathbf{x}}$ depends only on the vectors $P_{\mathrm{H}}\mathbf{X}_{\mathbf{b}}, P_{\mathrm{H}}\mathbf{Y}_{\mathbf{b}} \in \mathbb{T}_{\mathbf{b}}\mathbb{E}$.

Moreover, by Theorem 1.4.1, fixed any section $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ such that $\mathbf{s}_{\mathbf{x}} = \mathbf{b}$, there exists a uniquely determined pair of vectors $\mathbf{u}_{\mathbf{x}}, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$, such that $\mathbf{H}_{\mathbf{u}_{\mathbf{x}}}\mathbf{s} = (P_{\mathrm{H}}\mathbf{X})(\mathbf{s}_{\mathbf{x}})$ and $\mathbf{H}_{\mathbf{v}_{\mathbf{x}}}\mathbf{s} = (P_{\mathrm{H}}\mathbf{Y})(\mathbf{s}_{\mathbf{x}})$. The pair $\mathbf{u}_{\mathbf{x}}, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ does not depend on the choice of the field $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$, provided that $\mathbf{s}_{\mathbf{x}} = \mathbf{b}$.

We may then conclude that the curvature two-form $\mathbf{R} \in \Lambda^2(\mathbb{E}\,;\mathbb{V}\mathbb{E})$, evaluated on pairs of horizontal lifts, defines the vertical-valued two-form:

$$\text{CURV}(\mathbf{s})(\mathbf{v},\mathbf{u}) := P_{\mathrm{V}} \circ [\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}] \circ \mathbf{s} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{V}\mathbb{E})\,.$$

for any pair of vector fields $\mathbf{u},\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ on the tangent bundle and any section $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ of the fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$.

By tensoriality, for any section $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ the field $\text{CURV}(\mathbf{s}) \in \Lambda^2(\mathbb{M}\,;\mathbb{V}\mathbb{E})$ is a vertical-valued two-form on $\mathbb{M}$ with values in $\mathbb{V}\mathbb{E}$ and for any pair $\mathbf{u},\mathbf{v} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ the field $\text{CURV}(\mathbf{u},\mathbf{v}) \in \Lambda^1(\mathbb{M}\,;\mathbb{V}\mathbb{E})$ is a vertical tangent field on $\mathbb{E}$ along $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$. The projectability property of the horizontal lifts stated in Lemma 1.4.4 yields the relations

$$\left.\begin{aligned} T\mathbf{p} \circ [\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}] &= [\mathbf{u}\,,\mathbf{v}] \circ \mathbf{p} \\ T\mathbf{p} \circ \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]} &= [\mathbf{u}\,,\mathbf{v}] \circ \mathbf{p} \end{aligned}\right\} \implies T\mathbf{p} \circ ([\mathbf{H}_{\mathbf{u}},\mathbf{H}_{\mathbf{v}}\,,-]\mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]}) = 0\,.$$

Then $\mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]}$ is the horizontal component of $[\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}]$, i.e.

$$\mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]} = P_{\mathrm{H}} \cdot [\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}]\,,$$

and we get the equality: $[\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}] - \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]} = P_{\mathrm{V}} \circ [\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}]\,.$ ■

Theorem 1.4.5 (Curvature and covariant derivatives) *For any pair of vector fields* $\mathbf{u},\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ *the following identity holds:*

$$[\nabla_{\mathbf{u}}\,,\nabla_{\mathbf{v}}] - \nabla_{[\mathbf{u}\,,\mathbf{v}]} + [\mathbf{H}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}] - \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]} = [\mathbf{H}_{\mathbf{v}}\,,\nabla_{\mathbf{u}}] + [\nabla_{\mathbf{v}}\,,\mathbf{H}_{\mathbf{u}}] = 0\,.$$

Hence, for any section $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$*, the curvature form is given by*

$$\text{CURV}(\mathbf{s})(\mathbf{u},\mathbf{v}) = [\nabla_{\mathbf{u}}\,,\nabla_{\mathbf{v}}](\mathbf{s}) - \nabla_{[\mathbf{u}\,,\mathbf{v}]}(\mathbf{s})\,.$$

Proof. By definition, for any fixed section $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ of the fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, the natural derivative $T_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{T}\mathbb{E})$ along a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ and the vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ itself, are $\mathbf{s}$-related, i.e.:

$$\begin{array}{ccc} \mathbb{T}\mathbb{M} & \xrightarrow{T\mathbf{s}} & \mathbb{T}\mathbb{E} \\ {\scriptstyle \mathbf{u},\mathbf{v}}\uparrow & & \uparrow{\scriptstyle T_{\mathbf{u}},T_{\mathbf{v}}} \\ \mathbb{M} & \xrightarrow{\mathbf{s}} & \mathbb{E} \end{array} \iff \begin{cases} T_{\mathbf{u}} \circ \mathbf{s} := T\mathbf{s} \circ \mathbf{u} \in \mathrm{C}^0(\mathbb{E}\,;\mathbb{T}\mathbb{E})\,, \\ T_{\mathbf{v}} \circ \mathbf{s} := T\mathbf{s} \circ \mathbf{v} \in \mathrm{C}^0(\mathbb{E}\,;\mathbb{T}\mathbb{E})\,. \end{cases}$$

Then Lemma 1.3.4 tells us that $[T_{\mathbf{u}}\,, T_{\mathbf{v}}] \circ \mathbf{s} = T\mathbf{s} \circ [\mathbf{u}\,, \mathbf{v}]$ which may be written as $[T_{\mathbf{u}}, T_{\mathbf{v}}\,, \circ]\mathbf{s} = T_{[\mathbf{u}\,,\mathbf{v}]} \circ \mathbf{s}$. Then, being

$$T_{\mathbf{u}} = \nabla_{\mathbf{u}} + \mathbf{H}_{\mathbf{u}}\,, \qquad T_{\mathbf{v}} = \nabla_{\mathbf{v}} + \mathbf{H}_{\mathbf{v}}\,, \qquad T_{[\mathbf{u}\,,\mathbf{v}]} = \nabla_{[\mathbf{u}\,,\mathbf{v}]} + \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]}\,,$$

by the bilinearity of the LIE bracket we get

$$\begin{aligned}[\nabla_{\mathbf{u}} + \mathbf{H}_{\mathbf{u}}\,, \nabla_{\mathbf{v}} + \mathbf{H}_{\mathbf{v}}] &= [\nabla_{\mathbf{u}}\,, \nabla_{\mathbf{v}}] + [\mathbf{H}_{\mathbf{u}}\,, \mathbf{H}_{\mathbf{v}}] \\ &\quad + [\nabla_{\mathbf{u}}\,, \mathbf{H}_{\mathbf{v}}] + [\mathbf{H}_{\mathbf{u}}\,, \nabla_{\mathbf{v}}] = \nabla_{[\mathbf{u}\,,\mathbf{v}]} + \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]}\,,\end{aligned}$$

which, being $[\mathbf{H}_{\mathbf{u}}\,, \mathbf{H}_{\mathbf{v}}] - \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]} = P_{\mathrm{V}} \cdot [\mathbf{H}_{\mathbf{u}}\,, \mathbf{H}_{\mathbf{v}}]$, can be written as:

$$\begin{aligned}[\nabla_{\mathbf{u}}\,, \nabla_{\mathbf{v}}] - \nabla_{[\mathbf{u}\,,\mathbf{v}]} + [\mathbf{H}_{\mathbf{u}}\,, \mathbf{H}_{\mathbf{v}}] - \mathbf{H}_{[\mathbf{u}\,,\mathbf{v}]} &= [\nabla_{\mathbf{u}}\,, \nabla_{\mathbf{v}}] - \nabla_{[\mathbf{u}\,,\mathbf{v}]} + P_{\mathrm{V}} \cdot [\mathbf{H}_{\mathbf{u}}\,, \mathbf{H}_{\mathbf{v}}] \\ &= [\mathbf{H}_{\mathbf{v}}\,, \nabla_{\mathbf{u}}] + [\nabla_{\mathbf{v}}\,, \mathbf{H}_{\mathbf{u}}]\,.\end{aligned}$$

The tensoriality of the curvature two-form $\mathbf{R} \in \Lambda^2(\mathbb{E}\,; \mathbb{V}\mathbb{E})$ has the following implication. Let the vector fields $\mathcal{F}^{\mathbf{x}}_{\mathbf{u}}, \mathcal{F}^{\mathbf{x}}_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{T}\mathbb{E})$ be generated by dragging the vectors $\mathbf{H}_{\mathbf{u}_{\mathbf{x}}}, \mathbf{H}_{\mathbf{v}_{\mathbf{x}}} \in \mathbb{T}_{\mathbf{s}_{\mathbf{x}}}\mathbb{E}$ along the flows of $\nabla_{\mathbf{u}}, \nabla_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{T}\mathbb{E})$:

$$\begin{aligned}\mathcal{F}^{\mathbf{x}}_{\mathbf{u}} \circ \mathbf{Fl}^{\nabla_{\mathbf{v}}}_{\lambda} &:= T\mathbf{Fl}^{\nabla_{\mathbf{v}}}_{\lambda} \circ \mathbf{H}_{\mathbf{u}_{\mathbf{x}}}\,, \\ \mathcal{F}^{\mathbf{x}}_{\mathbf{v}} \circ \mathbf{Fl}^{\nabla_{\mathbf{u}}}_{\lambda} &:= T\mathbf{Fl}^{\nabla_{\mathbf{u}}}_{\lambda} \circ \mathbf{H}_{\mathbf{v}_{\mathbf{x}}}\,.\end{aligned}$$

By tensoriality, in evaluating the r.h.s. in the previous equality at a point $\mathbf{s}(\mathbf{x}) \in \mathbb{E}$, the horizontal lifts $\mathbf{H}_{\mathbf{u}}, \mathbf{H}_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{T}\mathbb{E})$ can be substituted by the vector fields $\mathcal{F}^{\mathbf{x}}_{\mathbf{u}}, \mathcal{F}^{\mathbf{x}}_{\mathbf{v}} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{T}\mathbb{E})$. Then $[\mathcal{F}^{\mathbf{x}}_{\mathbf{v}}\,, \nabla_{\mathbf{u}}]_{\mathbf{x}} = 0$ and $[\nabla_{\mathbf{v}}\,, \mathcal{F}^{\mathbf{x}}_{\mathbf{u}}]_{\mathbf{x}} = 0$ by definition. Being $[\mathbf{H}_{\mathbf{v}}\,, \nabla_{\mathbf{u}}]_{\mathbf{x}} + [\nabla_{\mathbf{v}}\,, \mathbf{H}_{\mathbf{u}}]_{\mathbf{x}} = [\mathcal{F}^{\mathbf{x}}_{\mathbf{v}}\,, \nabla_{\mathbf{u}}]_{\mathbf{x}} + [\nabla_{\mathbf{v}}\,, \mathcal{F}^{\mathbf{x}}_{\mathbf{u}}]_{\mathbf{x}}$ the result follows. ■

1.4.7 Connection on a vector bundle

Definition 1.4.14 (Connector) *In a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{M})$ *a connector* $\mathbf{K}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{T}\mathbb{E}\,; \mathbb{E})$ *is a linear homomorphism from the tangent bundle* $\boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{T}\mathbb{E}\,; \mathbb{E})$ *to the bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{M})$ *:*

$$\begin{array}{ccc} \mathbb{T}\mathbb{E} & \xrightarrow{\mathbf{K}_{\mathbb{E}}} & \mathbb{E} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{E}}}\big\downarrow & & \big\downarrow{\scriptstyle \mathbf{p}} \\ \mathbb{E} & \xrightarrow{\mathbf{p}} & \mathbb{M} \end{array} \qquad \iff \qquad \mathbf{p} \circ \mathbf{K}_{\mathbb{E}} = \mathbf{p} \circ \boldsymbol{\tau}_{\mathbb{E}} \in \mathrm{C}^0(\mathbb{T}\mathbb{E}\,; \mathbb{M})\,,$$

which is a left inverse to the vertical lift: $\mathbf{K}_{\mathbb{E}} \circ \mathbf{vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})} = \mathbf{id}_{\mathbb{E}}$ *. The linear map* $\mathbf{K}_{\mathbb{E}}(\mathbf{e}) \in BL(\mathbb{T}_{\mathbf{e}}\mathbb{E}\,; \mathbb{E}_{\mathbf{p}(\mathbf{e})})$ *meets:* $\boldsymbol{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e})) \cap \mathbb{V}_{\mathbf{e}}\mathbb{E} = \{0\}$ *.*

Definition 1.4.15 (Linear connector) *A linear connector* $\mathbf{K}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{E})$ *is a linear homomorphism also from the bundle* $T\mathbf{p} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TM})$ *to the bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *:*

$$\begin{array}{ccc} \mathbb{TE} & \xrightarrow{\mathbf{K}_{\mathbb{E}}} & \mathbb{E} \\ {\scriptstyle T\mathbf{p}}\Big\downarrow & & \Big\downarrow{\scriptstyle \mathbf{p}} \\ \mathbb{TM} & \xrightarrow{\boldsymbol{\tau}_{\mathbb{M}}} & \mathbb{M} \end{array} \quad \Longleftrightarrow \quad \mathbf{p} \circ \mathbf{K}_{\mathbb{E}} = \boldsymbol{\tau}_{\mathbb{M}} \circ T\mathbf{p} \in \mathrm{C}^0(\mathbb{TE}\,;\mathbb{M})\,.$$

Lemma 1.4.14 (Connections and connectors in a vector bundle) *In a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *, a connection is well-defined by a connector. A connector* $\mathbf{K}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{E})$ *induces a horizontal lift* $\mathbf{H} \in \mathrm{C}^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{TM}\,;\mathbb{HE})$ *:*

$$\mathbf{H}(\mathbf{e}\,,\mathbf{v}) := (T\mathbf{p}(\mathbf{e}))^{-1} \cdot \mathbf{v} \in \mathbb{H}_{\mathbf{e}}\mathbb{E}\,, \quad \forall\, \mathbf{e} \in \mathbb{E}, \quad \mathbf{v} \in \mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M}\,,$$

where $T\mathbf{p}(\mathbf{e})^{-1} \in BL\,(\mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M}\,;\boldsymbol{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e})))$ *. In turn an horizontal lift* $\mathbf{H} \in \mathrm{C}^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{TM}\,;\mathbb{HE})$ *induces a horizontal projector* $P_{\mathrm{H}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TE})$ *:*

$$P_{\mathrm{H}} := \mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p})\,.$$

A vertical projector $P_{\mathrm{V}} = \mathbf{id}_{\,\mathbb{TE}} - P_{\mathrm{H}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TE})$ *induces a connector* $\mathbf{K}_{\mathbb{E}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{E})$ *by:*

$$\begin{cases} \mathbf{K}_{\mathbb{E}} \cdot \mathbf{X} := \mathbf{vd}_{\,(\mathbb{E},\mathbf{p},\mathbb{M})}(\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})\,, P_{\mathrm{V}}(\mathbf{X})) \iff \\ P_{\mathrm{V}}(\mathbf{X}) = \mathbf{Vl}_{\,(\mathbb{E},\mathbf{p},\mathbb{M})}(\boldsymbol{\tau}_{\mathbb{E}}(\mathbf{X})\,, \mathbf{K}_{\mathbb{E}} \cdot \mathbf{X})\,, \quad \forall\, \mathbf{X} \in \mathbb{TE}\,. \end{cases}$$

with $\mathbf{vd}_{\,(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \in BL\,(\mathbb{V}_{\mathbf{e}}\mathbb{E}\,;\mathbb{E}_{\mathbf{p}(\mathbf{e})})$ *vertical drill at* $\mathbf{e} \in \mathbb{E}$ *.*

Proof. For any $\mathbf{e} \in \mathbb{E}$ we have that $\dim \mathbf{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e})) = \dim \mathbb{M} = m$ for reason of dimensions. Indeed $\dim \mathbb{E} = n = f + m$, with $f = \dim \mathbb{E}_{\mathbf{p}(\mathbf{e})}$, and $\dim \mathbb{T}_{\mathbf{e}}\mathbb{E} = \dim \mathbf{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e})) + \dim \mathbb{V}_{\mathbf{e}}\mathbb{E} = n$, with $\dim \mathbb{V}_{\mathbf{e}}\mathbb{E} = f$. Hence the linear subspace $\mathbf{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e}))$ is supplementary to $\mathbb{V}_{\mathbf{e}}\mathbb{E}$ and therefore the tangent map $T\mathbf{p}(\mathbf{e}) \in BL\,(\mathbb{T}_{\mathbf{e}}\mathbb{E}\,;\mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M})$ is invertible when restricted to $\mathbf{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e}))$. The horizontal lift $\mathbf{H} \in \mathrm{C}^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{TM}\,;\mathbb{HE})$ is defined as:

$$\mathbf{H}(\mathbf{e}\,,\mathbf{v}) := (T\mathbf{p}(\mathbf{e}))^{-1} \cdot \mathbf{v} \in \mathbb{H}_{\mathbf{e}}\mathbb{E}\,, \quad \forall\, \mathbf{e} \in \mathbb{E}, \quad \mathbf{v} \in \mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M}\,,$$

where $T\mathbf{p}(\mathbf{e})^{-1} \in BL\,(\mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M}\,;\mathbf{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e})))$ so that $(\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p}) \circ \mathbf{H} = \mathbf{id}_{\,\mathbb{E} \times_{\mathbb{M}} \mathbb{TM}}$. Vice versa, by Lemma 1.4.6, given a horizontal lift $\mathbf{H} \in \mathrm{C}^1(\mathbb{E} \times_{\mathbb{M}} \mathbb{TM}\,;\mathbb{HE})$, the induced horizontal projector $P_{\mathrm{H}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TE})$ is given by $P_{\mathrm{H}} = \mathbf{H} \circ (\boldsymbol{\tau}_{\mathbb{E}}\,, T\mathbf{p})$ and the associated the connector is given by $\mathbf{K}_{\mathbb{E}} = \mathbf{vd}_{\,\mathbb{E}} \circ P_{\mathrm{V}} = \mathbf{vd}_{\,\mathbb{E}} \circ (\mathbf{id}_{\,\mathbb{TE}} - P_{\mathrm{H}})$. Fiber linearity is easily checked and the property $\mathbf{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e})) \cap \mathbb{V}_{\mathbf{e}}\mathbb{E} = \{0\}$, being $\mathbf{K}_{\mathbb{E}}(\mathbf{e}) = \mathbf{vd}_{\,\mathbb{E}} \circ P_{\mathrm{V}}$, follows from $\mathbf{X} \in \mathbf{ker}(\mathbf{K}_{\mathbb{E}}(\mathbf{e})) \implies P_{\mathrm{V}}(\mathbf{X}) = \mathbf{0} \in \mathbb{T}_{\mathbf{e}}\mathbb{E}$. ■

A connection on a vector bundle $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ is called a KOSZUL *connection* [84]. The distinguishing feature with respect to a connection on a general fibre bundle is that, by the identification $\mathbb{VE} \simeq \mathbb{E}$, the covariant derivative $\nabla_{\mathbf{v}}\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{VE})$ of a section $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ along a vector field $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ may be considered as a section $\nabla_{\mathbf{v}}\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ of the vector bundle and the covariant derivative $\nabla_{\mathbf{v}} \in C^1(\mathbb{E}\,;\mathbb{E})$ as an operator on $\mathbb{E}$. This amounts to compare the map $\nabla_{\mathbf{v}}\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{VE})$ with the map

$$\mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \circ \nabla_{\mathbf{v}}\mathbf{s} = \mathbf{K}_{\mathbb{E}} \circ T\mathbf{s} = \mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \circ P_{\mathrm{V}} \circ T\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})\,,$$

and to hide the vertical drill $\mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \in C^1(\mathbb{TE}\,;\mathbb{E})$. For short, when the vector bundle is the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}\,;\mathbb{M})$, the manifold $\mathbb{M}$ itself is said to have a connection.

Theorem 1.4.6 (Covariant derivative as a point derivative) *Let a linear connection be defined in a vector bundle* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$. *Then the associated covariant derivative meets the properties of a point derivation:*

$$i) \quad \nabla_{(\alpha\,\mathbf{u}+\beta\,\mathbf{v})}\mathbf{s} = \alpha\,\nabla_{\mathbf{u}}\mathbf{s} + \beta\,\nabla_{\mathbf{v}}\mathbf{s}\,,$$

$$ii) \quad \begin{cases} \nabla_{\mathbf{v}}(\mathbf{s}_1 + \mathbf{s}_2) = \nabla_{\mathbf{v}}\mathbf{s}_1 + \nabla_{\mathbf{v}}\mathbf{s}_2\,, \\ \nabla_{\mathbf{v}}(f\,\mathbf{s}) = (\nabla_{\mathbf{v}}f)\,\mathbf{s} + f\,(\nabla_{\mathbf{v}}\mathbf{s})\,, \end{cases}$$

where $\alpha,\beta \in \Re$, $\mathbf{s},\mathbf{s}_1,\mathbf{s}_2 \in C^1(\mathbb{M}\,;\mathbb{E})$ *are sections,* $\mathbf{u},\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ *are tangent vector fields and* $f \in C^1(\mathbb{M},\Re)$ *is a scalar field with derivative* $\nabla_{\mathbf{v}}f$.

Proof. The point values of the covariant derivative $\nabla_{\mathbf{v}}\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ at $\mathbf{x} \in \mathbb{M}$ are vectors of a linear space, since $\nabla_{\mathbf{v}_{\mathbf{x}}}\mathbf{s} \in \mathbb{E}_{\mathbf{x}}$. Then property $i)$ stems from the tensoriality of the natural derivative. Properties $ii)$ follow from the linearity of the differential and the assumed linearity of the connection. ■

We remark that while property $i)$ and $ii)_1$ are shared also by the natural derivative and by the horizontal lift, the property $ii)_2$, which requires that the derivative of a section be still a section, is characteristic of a covariant derivative on a vector bundle.

Theorem 1.4.7 *In a vector bundle* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ *the iterated and the second covariant derivative according to a given connection are meaningful. Hence, for any section* $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$, *the curvature form may be written as*

$$\begin{aligned} \text{CURV}(\mathbf{s})(\mathbf{u},\mathbf{v}) &= (\nabla_{\mathbf{u}} \circ \nabla_{\mathbf{v}} - \nabla_{\mathbf{v}} \circ \nabla_{\mathbf{u}} - \nabla_{[\mathbf{u},\mathbf{v}]})\,\mathbf{s} \\ &= (\nabla^2_{\mathbf{uv}} - \nabla^2_{\mathbf{vu}} + \nabla_{\text{TORS}(\mathbf{u},\mathbf{v})})\,\mathbf{s}\,, \end{aligned}$$

in terms of the second covariant derivative $\nabla^2_{\mathbf{uv}} := \nabla_{\mathbf{u}} \circ \nabla_{\mathbf{v}} - \nabla_{\nabla_{\mathbf{u}}\mathbf{v}}$ *and of the torsion* $\text{TORS}(\mathbf{u},\mathbf{v}) := \nabla_{\mathbf{u}}\mathbf{v} - \nabla_{\mathbf{v}}\mathbf{u} - \nabla_{[\mathbf{u},\mathbf{v}]}$ *which are both tensor fields.*

Remark 1.4.1 *We underline that, in the context of general fibre bundles, in the formula:* $\text{CURV}(\mathbf{s})(\mathbf{u},\mathbf{v}) = [\nabla_{\mathbf{u}}, \nabla_{\mathbf{v}}](\mathbf{s}) - \nabla_{[\mathbf{u},\mathbf{v}]}(\mathbf{s})$*, provided by Theorem 1.4.5, the term* $[\nabla_{\mathbf{u}}, \nabla_{\mathbf{v}}](\mathbf{s})$ *cannot be written as* $(\nabla_{\mathbf{u}} \circ \nabla_{\mathbf{v}} - \nabla_{\mathbf{v}} \circ \nabla_{\mathbf{u}})\,\mathbf{s}$ *since the compositions* $\nabla_{\mathbf{u}} \circ \nabla_{\mathbf{v}}$ *and* $\nabla_{\mathbf{v}} \circ \nabla_{\mathbf{u}}$*, being* $\nabla_{\mathbf{u}}, \nabla_{\mathbf{v}} \in C^1(\mathbb{E}\,;\mathbb{VE})$*, are not defined, unless the identification* $\mathbb{VE} \simeq \mathbb{E}$ *can be made.*

Let us consider the dual vector bundles $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ and $\mathbf{p}^* \in C^1(\mathbb{E}^*\,;\mathbb{M})$.

Definition 1.4.16 *The* ***covariant derivative*** $\nabla_{\mathbf{u}}\mathbf{v}^* \in C^1(\mathbb{M}\,;\mathbb{E}^*)$ *of a section* $\mathbf{v}^* \in C^1(\mathbb{M}\,;\mathbb{E}^*)$ *of the dual vector bundle* $\mathbf{p}^* \in C^1(\mathbb{E}^*\,;\mathbb{M})$ *is defined by a formal application of* Leibniz *rule*

$$\langle \nabla_{\mathbf{u}}\mathbf{v}^*, \mathbf{v} \rangle := d_{\mathbf{u}} \langle \mathbf{v}^*, \mathbf{v} \rangle - \langle \mathbf{v}^*, \nabla_{\mathbf{u}}\mathbf{v} \rangle\,, \quad \forall\, \mathbf{v}, \mathbf{u} \in C^1(\mathbb{M}\,;\mathbb{TM})\,.$$

The result is a tensor field because, although each one of the addends at the r.h.s. depends on the extension of the vector $\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ to a vector field $\mathbf{v} \in C^1(U(\mathbf{x})\,;\mathbb{TM})$, the l.h.s. does not depend on such an extension, as may be shown by applying the tensoriality criterion of Lemma 1.2.1. The covariant derivative of a $(1,1)$ tensor field $\mathbf{T} \in C^1(\mathbb{M}\,; BL(\mathbb{TM}, \mathbb{T}^*\mathbb{M}\,;\Re))$ is also defined by a formal application of Leibniz rule:

$$(\nabla_{\mathbf{u}}\,\mathbf{T})(\mathbf{v}, \mathbf{v}^*) := d_{\mathbf{u}}(\mathbf{T}(\mathbf{v}, \mathbf{v}^*)) - \mathbf{T}(\nabla_{\mathbf{u}}\mathbf{v}, \mathbf{v}^*) - \mathbf{T}(\mathbf{v}, \nabla_{\mathbf{u}}\mathbf{v}^*)\,.$$

The result is a tensor field, as may be shown by Lemma 1.2.1.

When the model space is finite dimensional, the Christoffel *symbols* corresponding to a set of coordinate vector fields $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ are defined by

$$\boxed{\nabla_{\mathbf{e}_i}\,\mathbf{e}_j := \Gamma^k_{ij}\,\mathbf{e}_k\,.}$$

The next proposition provides the expression of the covariant derivative in coordinates.

Proposition 1.4.3 (Components of the covariant derivative) *The expression of the covariant derivative* $\nabla_{\mathbf{u}}\mathbf{v}$ *in terms of components, defined by a local chart* $\boldsymbol{\varphi} : \mathcal{U} \subseteq \mathbb{M} \mapsto \Re^n$*, is*

$$\nabla_{\mathbf{u}}\mathbf{v} = Y^j\,(X^i_{/j} + \Gamma^i_{jk} X^k)\,\mathbf{e}_i\,,$$

where the Einstein *notation has been adopted.*

Figure 1.19: Elwin Bruno Christoffel (1829 - 1900)

Proof. By posing $\mathbf{v} = X^A\,\mathbf{e}_A\,, \quad \mathbf{u} = Y^B\,\mathbf{e}_B\,,$ we have:

$$\begin{aligned}\nabla_{\mathbf{u}}\,\mathbf{v} &= \nabla_{(Y^j\,\mathbf{e}_j)}\,(X^k\,\mathbf{e}_k) = Y^j\,\nabla_{\mathbf{e}_j}\,(X^k\,\mathbf{e}_k) = \\ &= Y^j\,\big[(\partial_{\mathbf{e}_j}\,X^k)\,\mathbf{e}_k + (\nabla_{\mathbf{e}_j}\,\mathbf{e}_k)\,X^k\big] = \\ &= Y^j\,(X^i_{/j} + \Gamma^i_{jk}X^k)\,\mathbf{e}_i\,,\end{aligned}$$

and then the result.

1.4.8 Second covariant derivative

Given a vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ and a cross section $\mathbf{s} \in \mathrm{C}^2(\mathbb{M}\,;\mathbb{E})$, the iterated covariant derivative along the tangent vector fields $\mathbf{v},\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ is the covariant derivative along $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ of the covariant derivative along $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$. By a formal application of LEIBNIZ rule we get the expression

$$\nabla_{\mathbf{v}}\,\nabla_{\mathbf{u}}\,\mathbf{s} = \nabla^2_{\mathbf{vu}}\,\mathbf{s} + \nabla_{(\nabla_{\mathbf{v}}\mathbf{u})}\,\mathbf{s}\,,$$

which provides the way to introduce the *second covariant derivative* as the vector valued tensor field $\nabla^2\mathbf{s} \in \mathrm{C}(\mathbb{M}\,; BL\,(\mathbb{TM}^2\,;\mathbb{E}))$ defined by:

$$\nabla^2_{\mathbf{vu}}\,\mathbf{s} := \nabla_{\mathbf{v}}\,\nabla_{\mathbf{u}}\,\mathbf{s} - \nabla_{(\nabla_{\mathbf{v}}\mathbf{u})}\,\mathbf{s}\,.$$

Although the evaluation of the two terms on the r.h.s involves the derivatives of the vector field $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ in a neighborhood of the point of evaluation of the second covariant derivative, the l.h.s is in fact independent of the values of the field $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ at points other than the evaluation point. The second covariant derivative $\nabla^2_{\mathbf{vu}}\,\mathbf{s}$ is in fact tensorial in $\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ as can be shown by a direct application of the tensoriality criterion of Lemma 1.2.1. The tensoriality in $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ is apparent.

1.4.9 Curvature in a vector bundle

In a vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ the iterated covariant derivative along the vector fields $\mathbf{v}, \mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ makes sense and is given by

$$\nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\mathbf{s} := \nabla_{\mathbf{v}} \circ \mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \circ \nabla_{\mathbf{u}} \circ \mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})\,.$$

Moreover we have the Lie-bracket formula

$$[\nabla_{\mathbf{v}}\,,\nabla_{\mathbf{u}}] = \nabla_{\mathbf{v}} \circ \mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \circ \nabla_{\mathbf{u}} - \nabla_{\mathbf{u}} \circ \mathbf{vd}_{(\mathbb{E},\mathbf{p},\mathbb{M})} \circ \nabla_{\mathbf{v}} = \nabla_{\mathbf{v}}\nabla_{\mathbf{u}} - \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\,,$$

where on the l.h.s. $\nabla_{\mathbf{v}}, \nabla_{\mathbf{u}} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{VE})$ are vector fields, while on the r.h.s. $\nabla_{\mathbf{v}}, \nabla_{\mathbf{u}} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{E})$ are linear operators.

We may conclude that:

Proposition 1.4.4 *In a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *the curvature of a connection is given by*

$$\mathrm{CURV}(\mathbf{v}, \mathbf{u}) = \nabla_{\mathbf{v}}\nabla_{\mathbf{u}} - \nabla_{\mathbf{u}}\nabla_{\mathbf{v}} - \nabla_{[\mathbf{v}\,,\mathbf{u}]} \in \Lambda^2(\mathbb{M}\,; BL\,(\mathbb{E}\,;\mathbb{E}))\,.$$

and is a differential two-form taking values in the tensor bundle $BL\,(\mathbb{E}\,;\mathbb{E})$ *.*

Tensoriality of the curvature may also be got, and is usually made, by a direct application of the criterion in Lemma 1.2.1. This is the content of the next two lemmas.

Lemma 1.4.15 (1st tensoriality lemma) *In a vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *endowed with a connection, for any fixed section* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *, the curvature field*

$$\begin{aligned}\mathrm{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u}) &:= [\nabla_{\mathbf{v}}\,,\nabla_{\mathbf{u}}](\mathbf{s}) - \nabla_{[\mathbf{v}\,,\mathbf{u}]}(\mathbf{s}) \\ &= \nabla_{\mathbf{v}}\nabla_{\mathbf{u}}(\mathbf{s}) - \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}(\mathbf{s}) - \nabla_{[\mathbf{v}\,,\mathbf{u}]}(\mathbf{s}) \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{VE})\,,\end{aligned}$$

is vertical-valued in $\mathbb{TE}$ *and tensorial in the vector fields* $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *, i.e. the value of the field* $\mathrm{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u}) \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{VE})$ *at a point* $\mathbf{x} \in \mathbb{M}$ *depends only on the point values* $\mathbf{v}(\mathbf{x}), \mathbf{u}(\mathbf{x}) \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ *.*

Proof. By Proposition 1.4.1 the vertical bundle is integrable, FROBENIUS' condition implies that $[\nabla_{\mathbf{v}}, \nabla_{\mathbf{u}}, \in] C^1(\mathbb{E}; \mathbb{VE})$ and hence $\text{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u}) \in C^1(\mathbb{E}; \mathbb{VE})$.

The proof is based on the LIE-bracket formula and, according to Lemma 1.2.1, consists in verifying the property $\text{CURV}(f\mathbf{v}, g\mathbf{u}) = fg\, \text{CURV}(\mathbf{v}, \mathbf{u})$ for any $f, g \in C^1(\mathbb{M}; \Re)$. A simple computation yields that

$$\begin{aligned}
[\nabla_{\mathbf{v}}, \nabla_{g\mathbf{u}}] \circ \mathbf{s} &= (\nabla_{\mathbf{v}} \nabla_{g\mathbf{u}} - \nabla_{g\mathbf{u}} \nabla_{\mathbf{v}}) \circ \mathbf{s} \\
&= g\, \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{s} + \mathcal{L}_{\mathbf{v}} g\, \nabla_{\mathbf{u}} \mathbf{s} - f\, \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} \mathbf{s} \\
&= g\, [\nabla_{\mathbf{v}}, \nabla_{\mathbf{u}}] \circ \mathbf{s} + \mathcal{L}_{\mathbf{v}} g\, \nabla_{\mathbf{u}} \mathbf{s}, \\
\nabla_{[\mathbf{v}, g\mathbf{u}]} \circ \mathbf{s} &= \nabla_{g[\mathbf{v}, \mathbf{u}] + (\mathcal{L}_{\mathbf{v}} g)\mathbf{u}} \circ \mathbf{s} \\
&= g\, \nabla_{[\mathbf{v}, \mathbf{u}]} \mathbf{s} + \mathcal{L}_{\mathbf{v}} g\, \nabla_{\mathbf{u}} \mathbf{s}.
\end{aligned}$$

Then $\text{CURV}(\mathbf{s})(\mathbf{v}, g\mathbf{u}) = g\, \text{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u})$. An analogous computation shows that $\text{CURV}(\mathbf{s})(f\mathbf{v}, \mathbf{u}) = f\, \text{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u})$. ■

Lemma 1.4.16 (2nd tensoriality lemma) *The curvature of a connection on a vector bundle* $\mathbf{p} \in C^1(\mathbb{E}; \mathbb{M})$ *is a tensorial function of the section* $\mathbf{s} \in C^1(\mathbb{M}; \mathbb{E})$. *Indeed for any* $f \in C^1(\mathbb{M}; \Re)$ *it is:*

$$\text{CURV}(f\, \mathbf{s})(\mathbf{v}, \mathbf{u}) = f\, \text{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u}).$$

Proof. We have that

$$\begin{aligned}
\nabla_{\mathbf{v}} \nabla_{\mathbf{u}} (f\, \mathbf{s}) &= \nabla_{\mathbf{v}} (\nabla_{\mathbf{u}} f)\, \mathbf{s} + \nabla_{\mathbf{v}} (f\, \nabla_{\mathbf{u}} \mathbf{s}) \\
&= (\nabla_{\mathbf{v}} \nabla_{\mathbf{u}} f)\, \mathbf{s} + (\nabla_{\mathbf{u}} f)\, (\nabla_{\mathbf{v}} \mathbf{s}) + (\nabla_{\mathbf{v}} f)\, \nabla_{\mathbf{u}} \mathbf{s} + f\, \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{s},
\end{aligned}$$

and that

$$\nabla_{[\mathbf{v}, \mathbf{u}]} (f\, \mathbf{s}) = \big(\nabla_{\mathbf{v}} \nabla_{\mathbf{u}} f - \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} f\big)\, \mathbf{s} + f\, \nabla_{[\mathbf{v}, \mathbf{u}]} \mathbf{s}.$$

Then a simple computation and the tensoriality criterion of Lemma 1.2.1 yield the result. ■

In section 1.7.2 on page 170, it will be shown that the sum of curvature and cocurvature is equal to onehalf the FN bracket of the connection by itself:

$$\tfrac{1}{2}[P_{\mathrm{V}}, P_{\mathrm{V}}] = \mathbf{R} + \mathbf{R}^{\mathbf{c}}.$$

In the case of a connection on a fibre bundle, the cocurvature vanishes due to the integrability of the vertical bundle and the formula above gives: $\tfrac{1}{2}[P_{\mathrm{V}}, P_{\mathrm{V}}] = \mathbf{R}$. The *graded* JACOBI *identity* for the FN bracket implies that $[P_{\mathrm{V}}, [P_{\mathrm{V}}, P_{\mathrm{V}}], =]0$ which yields the BIANCHI identity for the curvature [80]: $[P_{\mathrm{V}}, \mathbf{R}] = 0$.

1.4.10 Parallel transport

Let $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ be a fibre bundle with a connection and $\mathbf{v} \in C^0(\mathbb{M}\,;\mathbb{TM})$ a vector field in the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}\,;\mathbb{M})$. According to the definition given in Proposition 1.3.15 on page 82, the *push* of a section $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ along a pair of $\mathbf{p}$-related vector fields $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ and $\mathbf{X} \in C^1(\mathbb{E}\,;\mathbb{TE})$ is given by: $\mathbf{Fl}^{(\mathbf{X},\mathbf{v})}_\lambda\!\uparrow\!\mathbf{s} := \mathbf{Fl}^{\mathbf{X}}_\lambda \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{-\lambda}$. Since a vector field $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ and its horizontal lift $\mathbf{H}_{\mathbf{v}} \in C^1(\mathbb{E}\,;\mathbb{TE})$ are $\mathbf{p}$-related, we may introduce the following concepts due to Gregorio Ricci-Curbastro and Tullio Levi-Civita.

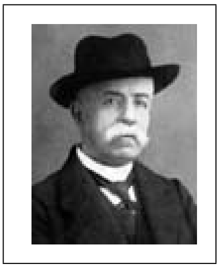

Figure 1.20: Gregorio Ricci-Curbastro (1853 - 1925)

Definition 1.4.17 (Parallel transport) *Let $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ be a fibre bundle with a connection. The parallel transport $\mathbf{Fl}^{\mathbf{v}}_\lambda\!\Uparrow\!\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ of a section $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ along the flow $\mathbf{Fl}^{\mathbf{v}}_\lambda \in C^1(\mathbb{M}\,;\mathbb{M})$ is defined by:*

$$\mathbf{Fl}^{\mathbf{v}}_\lambda\!\Uparrow\!\mathbf{s} := \mathbf{Fl}^{\mathbf{H}_\mathbf{v}}_\lambda \circ \mathbf{s} = (\mathbf{Fl}^{(\mathbf{H}_\mathbf{v},\mathbf{v})}_\lambda\!\uparrow\!\mathbf{s}) \circ \mathbf{Fl}^{\mathbf{v}}_\lambda\,,$$

so that $\mathbf{p} \circ \mathbf{Fl}^{\mathbf{v}}_\lambda\!\Uparrow\!\mathbf{s} = \mathbf{p} \circ \mathbf{Fl}^{\mathbf{H}_\mathbf{v}}_\lambda \circ \mathbf{s} = \mathbf{Fl}^{\mathbf{v}}_\lambda \circ \mathbf{p} \circ \mathbf{s} = \mathbf{Fl}^{\mathbf{v}}_\lambda$. We set $\mathbf{Fl}^{\mathbf{v}}_\lambda\!\Downarrow := \mathbf{Fl}^{\mathbf{v}}_{-\lambda}\!\Uparrow$.

A section $\mathbf{s} \in C^1(\mathbb{M}\,;\mathbb{E})$ of the fibre bundle $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ is *parallel transported* along the flow $\mathbf{Fl}^{\mathbf{v}}_\lambda \in C^1(\mathbb{M}\,;\mathbb{M})$ if

$$\mathbf{Fl}^{\mathbf{v}}_\lambda\!\Uparrow\!\mathbf{s} = \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_\lambda \iff \mathbf{Fl}^{(\mathbf{H}_\mathbf{v},\mathbf{v})}_\lambda\!\uparrow\!\mathbf{s} = \mathbf{s}\,, \quad \forall\, \lambda \in I\,.$$

The parallel transport enjoyes the same characteristic properties of a push:

$$\mathbf{Fl}^{\mathbf{u}}_0\!\Uparrow = \mathbf{id}_{\mathbb{E}}\,, \quad \mathbf{Fl}^{\mathbf{u}}_{\lambda+\mu}\!\Uparrow = \mathbf{Fl}^{\mathbf{u}}_\mu\!\Uparrow \mathbf{Fl}^{\mathbf{u}}_\lambda\!\Uparrow = \mathbf{Fl}^{\mathbf{u}}_\lambda\!\Uparrow \mathbf{Fl}^{\mathbf{u}}_\mu\!\Uparrow\,,$$

and $\mathbf{Fl}^{\mathbf{u}}_\lambda\!\Uparrow \mathbf{Fl}^{\mathbf{u}}_{-\lambda}\!\Uparrow = \mathbf{Fl}^{\mathbf{u}}_{-\lambda}\!\Uparrow \mathbf{Fl}^{\mathbf{u}}_\lambda\!\Uparrow = \mathbf{id}_{\mathbb{E}}$.

The horizontal lift $\mathbf{H_v}$ is defined pointwise in $\mathbb{M}$, and hence the parallel transport along a curve in $\mathbb{M}$ of a section defined only on that curve is meaningful and so is for the covariant derivative.

From the definition of parallel transport and Lemma 1.4.10 we infer that the covariant derivative and the horizontal lift are given by:

$$\begin{aligned} \nabla_{\mathbf{v}}\mathbf{s} &= \partial_{\lambda=0}\,\mathbf{Fl}_{\lambda}^{(\mathbf{H_v},\mathbf{v})}\!\downarrow\!\mathbf{s} = \partial_{\lambda=0}\,\mathbf{Fl}_{\lambda}^{\mathbf{v}}\Downarrow\mathbf{s}\circ\mathbf{Fl}_{\lambda}^{\mathbf{v}}\,, \\ \mathbf{H_v}\mathbf{s} &= \partial_{\lambda=0}\,\mathbf{Fl}_{\lambda}^{\mathbf{H_v}}\circ\mathbf{s} = \partial_{\lambda=0}\,\mathbf{Fl}_{\lambda}^{\mathbf{v}}\Uparrow\mathbf{s}\,. \end{aligned}$$

1.4.11 Parallel transport in a vector bundle

The covariant derivative of a parallel transported section $\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ of a fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$, vanishes identically:

$$\nabla_{\mathbf{v}}\mathbf{s} = \partial_{\lambda=0}\,\mathbf{Fl}_{\lambda}^{(\mathbf{H_v},\mathbf{v})}\!\downarrow\!\mathbf{s} = \partial_{\lambda=0}\,\mathbf{s} = 0\,.$$

In a vector bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$ the converse implication holds too.

Lemma 1.4.17 (Covariant derivative and parallel transport) *The covariant derivative* $\nabla_{\mathbf{v}}\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{VE})$, *of a section* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ *of the vector bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E};\mathbb{M})$ *along a vector field* $\mathbf{v} \in \mathrm{C}^0(\mathbb{M};\mathbb{TM})$, *vanishes along the path* $\mathbf{Fl}_{\lambda}^{\mathbf{v}}(\mathbf{x}) \in \mathrm{C}^1(I;\mathbb{M})$ *if and only if the section* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ *is parallel transported along that path.*

Proof. In a vector bundle the covariant derivative may be considered as a section $\nabla_{\mathbf{v}}\mathbf{s} = \mathcal{L}_{(\mathbf{H_v},\mathbf{v})}\mathbf{s} \in \mathrm{C}^1(\mathbb{M};\mathbb{E})$ and, from Proposition 1.3.15 on page 82, we get:

$$\partial_{\mu=\lambda}\,(\mathbf{Fl}_{\mu}^{(\mathbf{H_v},\mathbf{v})}\!\downarrow\!\mathbf{s}) = \mathbf{Fl}_{\lambda}^{(\mathbf{H_v},\mathbf{v})}\!\downarrow\!(\mathcal{L}_{(\mathbf{H_v},\mathbf{v})}\mathbf{s})\,,$$

Then $\nabla_{\mathbf{v}}\mathbf{s} = \mathcal{L}_{(\mathbf{H_v},\mathbf{v})}\mathbf{s} = 0$ implies that the pull back $\mathbf{Fl}_{\lambda}^{(\mathbf{H_v},\mathbf{v})}\!\downarrow\!\mathbf{s}$ is independent of $\lambda \in I$ and hence that $\mathbf{Fl}_{\lambda}^{(\mathbf{H_v},\mathbf{v})}\!\downarrow\!\mathbf{s} = \mathbf{s}$. ■

If the parallel transport of cross sections is independent of the curve chosen to join two points, we say that the connection defines a *distant parallelism.*

Proposition 1.4.5 *The curvature tensor field vanishes identically if the connection is defined by a distant parallelism.*

Proof. By the tensoriality property stated in Theorem 1.4.5, the curvature at a point $\mathbf{x} \in \mathbb{M}$ depends only on the point values $\mathbf{v}(\mathbf{x}), \mathbf{u}(\mathbf{x}), \mathbf{s}(\mathbf{x}) \in \mathbb{TM}$. To compute the point value $\text{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u})(\mathbf{x})$ according to the formula

$$\text{CURV}(\mathbf{s})(\mathbf{v}, \mathbf{u})(\mathbf{x}) := (\nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{s} - \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} \mathbf{s} - \nabla_{[\mathbf{v}, \mathbf{u}]} \mathbf{s})(\mathbf{x}),$$

we may extend the argument $\mathbf{s}(\mathbf{x})$ to a vector field $\mathbf{s} \in C^1(\mathbb{M}; \mathbb{TM})$ defined by a distant parallel transport. Hence the covariant derivative $\nabla \mathbf{s}$ vanishes along any curve and the curvature at $\mathbf{x} \in \mathbb{M}$ vanishes too. ■

1.4.12 Fiber and base derivative

Let us consider a fibre bundle $\mathbf{p} \in C^1(\mathbb{E}; \mathbb{M})$, and a manifold $\mathbb{F}$.

Definition 1.4.18 *The* ***fibre tangent map*** $T_{\text{F}}\mathbf{f} \in C^1(\mathbb{VE}; \mathbb{TF})$ *of a morphism* $\mathbf{f} \in C^1(\mathbb{E}; \mathbb{F})$ *is the restriction of the tangent map* $T\mathbf{f} \in C^1(\mathbb{TE}; \mathbb{TF})$ *to the vertical bundle of* $\mathbb{E}$*. It gives the rate of variation of* $\mathbf{f}$ *when the base point of its argument is held fixed in* $\mathbb{M}$*.*

For a real-valued functional $f \in C^1(\mathbb{E}; \Re)$, setting $\mathbb{T}_{\mathbf{f(b)}}\Re \simeq \Re$, it is

$$T_{\text{F}} f(\mathbf{b}) \in BL(\mathbb{V}_{\mathbf{b}}\mathbb{E}; \mathbb{T}_{\mathbf{f(b)}}\Re) = BL(\mathbb{V}_{\mathbf{b}}\mathbb{E}; \Re) = \mathbb{V}^*_{\mathbf{b}}\mathbb{E}.$$

Then $T_{\text{F}} f \in C^1(\mathbb{E}; \mathbb{V}^*\mathbb{E})$ is a section of the bundle $\boldsymbol{\tau}^*_{\mathbb{E}} \in C^1(\mathbb{T}^*\mathbb{B}; \mathbb{E})$, i.e. $\boldsymbol{\tau}^*_{\mathbb{E}} \circ T_{\text{F}} f = \mathbf{id}_{\mathbb{E}}$. In a vector bundle $\mathbf{p} \in C^1(\mathbb{E}; \mathbb{M})$ we have the identification $\mathbb{V}_{\mathbf{e}}\mathbb{E} \simeq \mathbb{E}_{\mathbf{e}}$ and hence $\mathbb{VE} \simeq \mathbb{E}$. For any morphism $\mathbf{f} \in C^1(\mathbb{E}; \mathbb{F})$, $T_{\text{F}}\mathbf{f} \in C^1(\mathbb{E}; \mathbb{TF})$, and for a real-valued functional $f \in C^1(\mathbb{E}; \Re)$:

$$T_{\text{F}} f(\mathbf{e}) \in BL(\mathbb{V}_{\mathbf{e}}\mathbb{E}; \mathbb{T}_{\mathbf{f(e)}}\Re) = BL(\mathbb{E}_{\mathbf{e}}; \Re).$$

Then $T_{\text{F}} f \in C^1(\mathbb{E}; \mathbb{E}^*)$.

Definition 1.4.19 *The* ***fibre derivative*** *at* $\mathbf{e} \in \mathbb{E}$ *of a morphism* $\mathbf{f} \in C^1(\mathbb{E}; \mathbb{F})$ *from a vector bundle* $\mathbf{p} \in C^1(\mathbb{E}; \mathbb{M})$ *to a manifold* $\mathbb{F}$ *is the linear map* $d_{\text{F}}\mathbf{f}(\mathbf{e}) \in C^1(\mathbb{E}_{\mathbf{p(e)}}; \mathbb{T}_{\mathbf{f(e)}}\mathbb{F})$ *defined by*

$$d_{\text{F}}\mathbf{f}(\mathbf{e}) \cdot \boldsymbol{\eta} = T\mathbf{f}(\mathbf{e}) \cdot \mathbf{Vl}_{(\mathbb{E}, \mathbf{p}, \mathbb{M})}(\mathbf{e}) \cdot \boldsymbol{\eta}, \quad \forall \boldsymbol{\eta} \in \mathbb{E}_{\mathbf{p(e)}}.$$

For a morphism $\mathbf{f} \in C^1(\mathbb{E}; F)$ with values in a BANACH space F, we have that $d_{\text{F}}\mathbf{f}(\mathbf{v}) \in BL(\mathbb{E}_{\mathbf{p(v)}}; \mathbb{T}_{\mathbf{f(v)}}F) \simeq BL(\mathbb{E}_{\mathbf{p(v)}}; F)$, and in a tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in C^1(\mathbb{TM}; \mathbb{M})$ and a morphism $\mathbf{f} \in C^1(\mathbb{TM}; F)$, we have that $d_{\text{F}}\mathbf{f}(\mathbf{v}) \in BL(\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}; \mathbb{T}_{\mathbf{f(v)}}F) \simeq BL(\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}; F)$,

Definition 1.4.20 *The **fibre derivative** of a functional* $f \in \mathrm{C}^1(\mathbb{TM}\,;\Re)$ *at* $\mathbf{v} \in \mathbb{TM}$ *, is defined by*

$$\langle d_{\mathrm{F}} f(\mathbf{v}), \mathbf{w} \rangle = \langle df(\mathbf{v}), \mathbf{Vl}_{(\mathbb{TM}, \boldsymbol{\tau}_{\mathbb{M}}, \mathbb{M})}(\mathbf{v}) \cdot \mathbf{w} \rangle \,, \quad \forall\, \mathbf{w} \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M} \,,$$

which may be written also $d_{\mathrm{F}} f = \mathbf{Vl}^*_{\mathbb{TM}} \cdot df$ *, or explicitly as*

$$d_{\mathrm{F}} f(\mathbf{v}) = \mathbf{Vl}^*_{(\mathbb{TM}, \boldsymbol{\tau}_{\mathbb{M}}, \mathbb{M})}(\mathbf{v}) \cdot df(\mathbf{v}) \,.$$

We have that $df \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^*\mathbb{TM})$ with $df(\mathbf{v}) \in \mathbb{T}^*_{\mathbf{v}}\mathbb{TM}$ and that $\mathbf{Vl}^*_{\mathbb{TM}}(\mathbf{v}) \in BL(\mathbb{T}^*_{\mathbf{v}}\mathbb{TM}\,;\mathbb{T}^*_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M})$. Then $d_{\mathrm{F}} f \in BL(\mathbb{TM}\,;\mathbb{T}^*\mathbb{M})$ with $d_{\mathrm{F}} f(\mathbf{v}) \in \mathbb{T}^*_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}$.

By the identification $\mathbb{VTM} \simeq \mathbb{TM}$, the **fibre derivative** $d_{\mathrm{F}} f(\mathbf{v}) \in \mathbb{T}^*_{\mathbf{v}}\mathbb{TM}$ of a functional $f \in \mathrm{C}^1(\mathbb{TM}\,;\Re)$ may also be defined by the formula

$$d_{\mathrm{F}} f(\mathbf{v}) \cdot \mathbf{w} := \partial_{\lambda=0}\, f(\mathbf{v} + \lambda\, \mathbf{w}) \,.$$

Let $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ be a fibre bundle with a connection ∇ and $\mathbb{F}$ a manifold.

Definition 1.4.21 *The **fibre-covariant derivative** of a morphism* $\mathbf{f} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{F})$ *at a section* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *of the fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$*, is the map* $d_{\mathrm{F}}\mathbf{f}(\mathbf{s}(\mathbf{x})) \in BL(\mathbb{T}_{\mathbf{s}(\mathbf{x})}\mathbb{E}_{\mathbf{x}}\,;\mathbb{T}_{\mathbf{f}(\mathbf{s}(\mathbf{x}))}\mathbb{F})$ *defined by*

$$\begin{aligned} d_{\mathrm{F}}\mathbf{f}(\mathbf{s}(\mathbf{x})) \cdot \nabla_{\mathbf{v}_{\mathbf{x}}}\mathbf{s} &:= (T_{\mathrm{F}}\mathbf{f} \circ T\mathbf{s}) \cdot \mathbf{v}_{\mathbf{x}} = (T\mathbf{f} \circ P_{\mathrm{V}} \circ T\mathbf{s}) \cdot \mathbf{v}_{\mathbf{x}} \\ &= T\mathbf{f} \cdot \nabla_{\mathbf{v}_{\mathbf{x}}}\mathbf{s} \in \mathbb{T}_{\mathbf{f}(\mathbf{s}(\mathbf{x}))}\mathbb{F} \,, \end{aligned}$$

for any tangent vector $\mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ *.*

Definition 1.4.22 *The **horizontal tangent map*** $T_{\mathrm{B}}\mathbf{f} \in \mathrm{C}^1(\mathbb{HE}\,;\mathbb{TF})$ *of a morphism* $\mathbf{f} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{F})$ *is the composition of the tangent map* $T\mathbf{f} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TF})$ *with the projection* $P_{\mathrm{H}} \in \mathrm{C}^1(\mathbb{TE}\,;\mathbb{TE})$ *on the horizontal bundle of* $\mathbb{E}$ *.*

Definition 1.4.23 *The **base derivative** of a morphism* $\mathbf{f} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{F})$ *at a point* $\mathbf{s}_{\mathbf{x}} \in \mathbb{E}_{\mathbf{x}}$ *in the fibre bundle* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$*, is the map* $d_{\mathrm{B}}\mathbf{f}(\mathbf{s}_{\mathbf{x}}) \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}_{\mathbf{f}(\mathbf{s}_{\mathbf{x}})}\mathbb{F})$ *is defined by*

$$\begin{aligned} d_{\mathrm{B}}\mathbf{f}(\mathbf{s}_{\mathbf{x}}) \cdot \mathbf{v}_{\mathbf{x}} &:= (T_{\mathrm{B}}\mathbf{f} \circ T\mathbf{s}) \cdot \mathbf{v}_{\mathbf{x}} = (T\mathbf{f} \circ P_{\mathrm{H}} \circ T\mathbf{s}) \cdot \mathbf{v}_{\mathbf{x}} \\ &= (T\mathbf{f} \circ \mathbf{H}_{\mathbf{v}_{\mathbf{x}}})(\mathbf{s}_{\mathbf{x}}) \in \mathbb{T}_{\mathbf{f}(\mathbf{s}_{\mathbf{x}})}\mathbb{F} \,, \end{aligned}$$

for any vector $\mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ *. Here* $\mathbf{s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{E})$ *is any section of* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ *such that* $\mathbf{s}(\mathbf{x}) = \mathbf{s}_{\mathbf{x}} \in \mathbb{E}_{\mathbf{x}}$ *.*

Any vector $\mathbf{X}(\mathbf{e}) \in \mathbb{T}_{\mathbf{e}}\mathbb{E}$, tangent to a vector bundle $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ endowed with a connection ∇, may be uniquely split into a vertical and a horizontal component. Moreover, vertical and horizontal vectors in $\mathbb{T}_{\mathbf{e}}\mathbb{E}$ are uniquely determined respectively as vertical lifting at $\mathbf{e} \in \mathbb{E}$ of a vector in the linear fibre $\mathbb{E}_{\mathbf{p}(\mathbf{e})}$ and horizontal lifting at $\mathbf{e} \in \mathbb{E}$ of a tangent vector in $\mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M}$. Then we may state the following result.

Lemma 1.4.18 (Decomposition of the tangent map) *Let* $\mathbf{f} \in C^1(\mathbb{E};\mathbb{F})$ *be a morphism from a vector bundle* $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ *into a manifold* $\mathbb{F}$. *Then for any* $\mathbf{X}(\mathbf{e}) \in \mathbb{T}_{\mathbf{e}}\mathbb{E}$ *we have the unique decomposition*

$$T\mathbf{f}(\mathbf{e}) \cdot \mathbf{X}(\mathbf{e}) = d_{\mathrm{F}}\mathbf{f}(\mathbf{e}) \cdot \mathbf{v}_{\mathbf{X}} + d_{\mathrm{B}}\mathbf{f}(\mathbf{e}) \cdot \mathbf{h}_{\mathbf{X}}\,, \quad \mathbf{e} \in \mathbb{E}\,,$$

where $\mathbf{h}_{\mathbf{X}} := T\mathbf{p}(\mathbf{e}) \cdot \mathbf{X}(\mathbf{e}) \in \mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M}$ *and* $\mathbf{v}_{\mathbf{X}} \in \mathbb{E}_{\mathbf{p}(\mathbf{e})}$ *is defined by*

$$\mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \cdot \mathbf{v}_{\mathbf{X}} = \mathbf{X}(\mathbf{e}) - \mathbf{H}(\mathbf{e}) \cdot \mathbf{h}_{\mathbf{X}} \in \mathbb{V}_{\mathbf{e}}\mathbb{E}\,,$$

and $\mathbf{h}_{\mathbf{X}} \in \mathbb{T}_{\mathbf{p}(\mathbf{e})}\mathbb{M}$ *so that:*

$$\mathbf{X}(\mathbf{e}) = \mathbf{Vl}_{(\mathbb{E},\mathbf{p},\mathbb{M})}(\mathbf{e}) \cdot \mathbf{v}_{\mathbf{X}} + \mathbf{H}(\mathbf{e}) \cdot \mathbf{h}_{\mathbf{X}} \in \mathbb{T}_{\mathbf{e}}\mathbb{E}\,.$$

Theorem 1.4.8 (Fiber-covariant and base derivative) *Let us consider a fibre bundle* $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ *with a connection* ∇. *Then the tangent map of the composition* $\mathbf{f} \circ \mathbf{s} \in C^1(\mathbb{M};\mathbb{F})$ *of a morphism* $\mathbf{f} \in C^1(\mathbb{E};\mathbb{F})$ *with a section* $\mathbf{s} \in C^1(\mathbb{M};\mathbb{E})$ *may be uniquely split according to the formula*

$$T(\mathbf{f} \circ \mathbf{s}) = T\mathbf{f} \circ T\mathbf{s} = d_{\mathrm{F}}\mathbf{f}(\mathbf{s}) \cdot \nabla\mathbf{s} + d_{\mathrm{B}}\mathbf{f}(\mathbf{s})\,,$$

as sum of the ***fibre-covariant derivative*** *and the* ***base derivative*** *whose expressions in terms of the parallel transport* $\mathbf{Fl}^{\mathbf{v}}_{\lambda}\Uparrow \in C^1(\mathbb{E};\mathbb{E})$ *along the flow associated with the vector field* $\mathbf{v} \in C^1(\mathbb{M};\mathbb{TM})$ *are given by*

$$\begin{aligned} d_{\mathrm{F}}\mathbf{f}(\mathbf{s}(\mathbf{x})) \cdot \nabla_{\mathbf{v}(\mathbf{x})}\mathbf{s} &:= \partial_{\lambda=0}\,(\mathbf{f} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}\Downarrow \mathbf{s} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda})(\mathbf{x})\,,\\ d_{\mathrm{B}}\mathbf{f}(\mathbf{s}(\mathbf{x})) \cdot \mathbf{v}(\mathbf{x}) &:= \partial_{\lambda=0}\,(\mathbf{f} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}\Uparrow \mathbf{s})(\mathbf{x})\,. \end{aligned}$$

Proof. By the definitions and the chain rule:

$$\begin{aligned} d_{\mathrm{F}}\mathbf{f}(\mathbf{s}(\mathbf{x}))\cdot\nabla_{\mathbf{v}(\mathbf{x})}\mathbf{s} &= (T\mathbf{f}\circ\nabla_{\mathbf{v}(\mathbf{x})}\mathbf{s})(\mathbf{x}) \\ &= (T\mathbf{f}\circ P_{\mathrm{V}}\circ T\mathbf{s}\circ\mathbf{v})(\mathbf{x}) \\ &= T_{\mathbf{s}(\mathbf{x})}\mathbf{f}\cdot\partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{v}}_{\lambda}\Downarrow\mathbf{s})(\mathbf{Fl}^{\mathbf{v}}_{\lambda}(\mathbf{x})) \\ &= \partial_{\lambda=0}\,(\mathbf{f}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\Downarrow\mathbf{s}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda})(\mathbf{x})\,, \\ d_{\mathrm{B}}\mathbf{f}(\mathbf{s}(\mathbf{x}))\cdot\mathbf{v}(\mathbf{x}) &= (T\mathbf{f}\circ\mathbf{H}_{\mathbf{v}(\mathbf{x})}\mathbf{s})(\mathbf{x}) \\ &= (T\mathbf{f}\circ P_{\mathrm{H}}\circ T\mathbf{s}\circ\mathbf{v})(\mathbf{x}) \\ &= (T\mathbf{f}\circ\partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{v}}_{\lambda}\Uparrow\mathbf{s})(\mathbf{x}) \\ &= \partial_{\lambda=0}\,(\mathbf{f}\circ\mathbf{Fl}^{\mathbf{v}}_{\lambda}\Uparrow\mathbf{s})(\mathbf{x})\,, \end{aligned}$$

so that

$$T(\mathbf{f}\circ\mathbf{s})\cdot\mathbf{v}(\mathbf{x}) = d_{\mathrm{F}}\mathbf{f}(\mathbf{s}(\mathbf{x}))\cdot\nabla_{\mathbf{v}(\mathbf{x})}\mathbf{s} + d_{\mathrm{B}}\mathbf{f}(\mathbf{s}(\mathbf{x}))\cdot\mathbf{v}(\mathbf{x})\,,$$

and we get the result. ∎

1.4.13 Legendre transform

Let us consider a manifold $\mathbb{M}$ with a connection.

Definition 1.4.24 *The* **Legendre *transform*** *is the fibre-linear correspondence between the dual bundles,* $\boldsymbol{\tau}\in \mathrm{C}^1(\mathbb{TM};\mathbb{M})$ *and* $\boldsymbol{\tau}^*\in \mathrm{C}^1(\mathbb{T}^*\mathbb{M};\mathbb{M})$, *induced by the fibre derivative* $d_{\mathrm{F}}L\in \mathrm{C}^1(\mathbb{VTM};\mathbb{T}^*\mathbb{M}) = \mathrm{C}^1(\mathbb{TM};\mathbb{T}^*\mathbb{M})$ *of a functional* $L\in \mathrm{C}^1(\mathbb{TM};\Re)$ *on the tangent bundle* $\boldsymbol{\tau}\in \mathrm{C}^1(\mathbb{TM};\mathbb{M})$, *defined by:*

$$d_{\mathrm{F}}L(\mathbf{v}) := \mathbf{Vl}^*_{\mathbb{TM}}(\mathbf{v})\cdot dL(\mathbf{v})\,.$$

More in general we may assume that the functional $\mathrm{L}\in \mathrm{C}^1(\mathbb{TM};\Re)$ is fibre-wise strictly convex, i.e. when evaluated holding the base point fixed, it fulfils the inequality:

$$\mathrm{L}(\alpha_1\,\mathbf{v}_1+\alpha_2\,\mathbf{v}_2)\le \alpha_1\,\mathrm{L}(\mathbf{v}_1)+\alpha_2\,\mathrm{L}(\mathbf{v}_2)\,,$$

for all $\alpha_1,\alpha_2\in\Re$ such that $\alpha_1+\alpha_2=1$ and $0<\alpha_1,\alpha_2<1$ and for all $\mathbf{v}_1,\mathbf{v}_2\in\mathbb{TM}$ such that $\boldsymbol{\tau}(\mathbf{v}_1)=\boldsymbol{\tau}(\mathbf{v}_2)$, with equality iff $\mathbf{v}_1=\mathbf{v}_2$. Then the fibre derivative of $\mathrm{L}\in \mathrm{C}^1(\mathbb{TM};\Re)$ has a fibre-wise strictly monotone graph, i.e.

$$\langle d_{\mathrm{F}}\mathrm{L}(\mathbf{v}_2)-d_{\mathrm{F}}\mathrm{L}(\mathbf{v}_1),\mathbf{v}_2-\mathbf{v}_1\rangle\ge 0\,,$$

for $\mathbf{v}_1,\mathbf{v}_2$ in $\mathbb{TM}$ such that $\boldsymbol{\tau}(\mathbf{v}_1)=\boldsymbol{\tau}(\mathbf{v}_2)$, with equality only if $\mathbf{v}_1=\mathbf{v}_2$.

Figure 1.21: Adrien-Marie Legendre (1752 - 1833)

Moreover $d_{\mathrm{F}}\mathrm{L} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^*\mathbb{M})$ admits fibre-wise a strictly monotone inverse $(d_{\mathrm{F}}\mathrm{L})^{-1} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{TM})$ which is in turn the fibre derivative $d_{\mathrm{F}}H \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{TM})$ of a functional $H \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\Re)$ on the dual vector bundle:

$$(d_{\mathrm{F}}\mathrm{L})^{-1} = d_{\mathrm{F}}H\,.$$

By adjusting the integration constant, the two functionals are related by the *conjugacy relation*:

$$\mathrm{L}(\mathbf{v}) + H(d_{\mathrm{F}}\mathrm{L}(\mathbf{v})) = \langle\, d_{\mathrm{F}}\mathrm{L}(\mathbf{v}), \mathbf{v}\,\rangle\,,$$

equivalent to $\mathrm{L}(d_{\mathrm{F}}H(\mathbf{v}^*)) + H(\mathbf{v}^*) = \langle\, \mathbf{v}^*, d_{\mathrm{F}}H(\mathbf{v}^*)\,\rangle\,.$

Defining the energy functional $E := H \circ d_{\mathrm{F}}\mathrm{L}$ and recalling the definitions of the fibre derivative $d_{\mathrm{F}}\mathrm{L} := \mathbf{Vl}^*_{(\mathbb{TM}\,,\boldsymbol{\tau}_{\mathbb{M}}\,,\mathbb{M})} \circ d\mathrm{L}$ and of the LIOUVILLE vector field $\mathbf{C_V l}_{(\mathbb{TM}\,,\boldsymbol{\tau}_{\mathbb{M}}\,,\mathbb{M})}(\mathbf{v}) := \mathbf{Vl}_{(\mathbb{TM}\,,\boldsymbol{\tau}_{\mathbb{M}}\,,\mathbb{M})}(\mathbf{v}) \cdot \mathbf{v}$, the conjugacy relation may also be written as

$$\mathrm{L}(\mathbf{v}) + E(\mathbf{v}) = \langle\, d_{\mathrm{F}}\mathrm{L}(\mathbf{v}), \mathbf{v}\,\rangle = d\mathrm{L}(\mathbf{v}) \cdot \mathbf{C}_{\mathbb{TM}}(\mathbf{v})\,.$$

The strict monotonicity of the fibre derivative implies the strict convexity of the potentials and hence the following fibre-wise inequality is fulfilled:

$$\mathrm{L}(\mathbf{v}) + H(\mathbf{v}^*) \geq \langle\, \mathbf{v}^*, \mathbf{v}\,\rangle\,, \quad \forall\, \{\mathbf{v}\,,\mathbf{v}^*\} \in \mathbb{TM} \times \mathbb{T}^*\mathbb{M} \,:\, \boldsymbol{\tau}^*(\mathbf{v}^*) = \boldsymbol{\tau}(\mathbf{v})\,,$$

with equality if and only if the pair $\{\mathbf{v}\,,\mathbf{v}^*\} \in \mathbb{TM} \times \mathbb{T}^*\mathbb{M}$ is in the graph of the LEGENDRE transform, that is:

$$\mathbf{v}^* = d_{\mathrm{F}}\mathrm{L}(\mathbf{v})\,, \quad \mathbf{v} = d_{\mathrm{F}}H(\mathbf{v}^*)\,, \quad \text{with} \quad \boldsymbol{\tau}^*(\mathbf{v}^*) = \boldsymbol{\tau}(\mathbf{v})\,.$$

The conjugacy relation may then be written as

$$\begin{aligned} H(\mathbf{v}^*) &= \sup_{\mathbf{v}\in\mathbb{E}_{\boldsymbol{\tau}^*(\mathbf{v}^*)}} \{\langle \mathbf{v}^*, \mathbf{v}\rangle - \mathrm{L}(\mathbf{v})\}\,, \\ \mathrm{L}(\mathbf{v}) &= \sup_{\mathbf{v}^*\in\mathbb{E}^*_{\boldsymbol{\tau}(\mathbf{v})}} \{\langle \mathbf{v}^*, \mathbf{v}\rangle - H(\mathbf{v}^*)\}\,. \end{aligned}$$

The Fenchel transform yields an analogous result under the weaker assumption that the functionals are convex and subdifferentiable.

This more general situation arises naturally in the analysis of problems of calculus of variations involving the stationarity of a lenght and will be illustrated in section 2.4.7 with reference to the Hamilton-Jacobi equation in dynamics.

The Legendre transform plays an important role in physics. In dynamics, Lagrange and Hamilton functionals, and, in thermodynamics, the *internal energy*, Helmholtz *potential*, Gibbs *potential* and the *enthalpy*, are related one another by a Legendre transform.

1.4.14 Whitney sum of dual vector bundles

Figure 1.22: Hassler Whitney (1907 - 1989)

Definition 1.4.25 (Whitney sum) *Let* $\mathbf{p} \in C^1(\mathbb{E}\,;\mathbb{M})$ *and* $\mathbf{p}^* \in C^1(\mathbb{E}^*\,;\mathbb{M})$ *be dual vector bundles. Their* WHITNEY *sum is the vector bundle* $\mathbb{E}\oplus\mathbb{E}^*$ *whose fibres are the cartesian products of the corresponding dual fibres:*

$$(\mathbb{E}\oplus\mathbb{E}^*)_{\mathbf{x}} = \mathbb{E}_{\mathbf{x}} \times \mathbb{E}^*_{\mathbf{x}}\,, \quad \forall\,\mathbf{x} \in \mathbb{M}\,.$$

On a WHITNEY sum we define the evaluation map $\text{EVAL} \in C^1(\mathbb{E}\oplus\mathbb{E}^*\,;\Re)$ by

$$\text{EVAL}(\mathbf{v}\,,\mathbf{v}^*) := \langle\,\mathbf{v}^*,\mathbf{v}\,\rangle\,, \quad \forall\,(\mathbf{v}\,,\mathbf{v}^*) \in \mathbb{E}\oplus\mathbb{E}^*\,.$$

The *fibre derivative* $d_{\text{F}}\text{EVAL} \in C^0(\mathbb{E}\oplus\mathbb{E}^*\,;\mathbb{E}^*\oplus\mathbb{E})$ is given by

$$\begin{aligned} d_{\text{F}}\text{EVAL}(\mathbf{v}\,,\mathbf{v}^*)\cdot(\mathbf{w}\,,\mathbf{w}^*) &:= \lim_{\lambda\to 0}\frac{1}{\lambda}\big[\langle\,\mathbf{v}^*+\lambda\mathbf{w}^*,\mathbf{v}+\lambda\mathbf{w}\,\rangle - \langle\,\mathbf{v}^*,\mathbf{v}\,\rangle\big] \\ &= \langle\,\mathbf{v}^*,\mathbf{w}\,\rangle + \langle\,\mathbf{w}^*,\mathbf{v}\,\rangle\,, \end{aligned}$$

and can be identified with a symmetric tensor $d_{\text{F}}\text{EVAL} \in C^1(\mathbb{E}\oplus\mathbb{E}^*,\mathbb{E}\oplus\mathbb{E}^*\,;\Re)$.

From the property

$$d_{\text{F}}\text{EVAL}(\mathbf{v}\,,\mathbf{v}^*)\cdot(\mathbf{v}\,,\mathbf{v}^*) = 2\,\text{EVAL}(\mathbf{v}\,,\mathbf{v}^*)\,.$$

we infer that, by EULER's theorem, the evaluation map $\text{EVAL} \in C^1(\mathbb{E}\oplus\mathbb{E}^*\,;\Re)$ is homogeneous of order 2 and hence, being indefinitely derivable, quadratic. Assuming reflexivity, that is $\mathbb{E}^{**} = \mathbb{E}$, the evaluation map is weakly nondegenerate since

$$\begin{aligned} \langle\,\mathbf{v}^*,\mathbf{w}\,\rangle = 0\,, &\quad \forall\,\mathbf{w}\ \in \mathbb{E}\ \implies \mathbf{v}^* = 0\,, \\ \langle\,\mathbf{w}^*,\mathbf{v}\,\rangle = 0\,, &\quad \forall\,\mathbf{w}^* \in \mathbb{E}^* \implies \mathbf{v} = 0\,. \end{aligned}$$

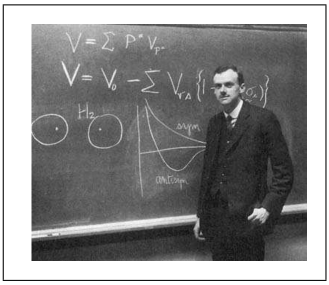

Figure 1.23: Paul Adrien Maurice Dirac (1902 - 1984)

Definition 1.4.26 (Dirac's structure) *A* DIRAC*'s structure is a vector subbundle* $\mathbf{D}_{\mathbb{E}} \subseteq \mathbb{E} \oplus \mathbb{E}^*$ *such that* $\mathbf{D} = \mathbf{D}^{\perp}$ *that is* $\mathbf{D}_{\mathbf{x}} = \mathbf{D}_{\mathbf{x}}^{\perp}$ *for every* $\mathbf{x} \in \mathbb{M}$ *where orthogonality* $\perp$ *is intended with respect to the pairing induced by the fibre derivative of the evaluation map* $d_{\mathrm{F}}\mathrm{EVAL} \in \mathrm{C}^1(\mathbb{E} \oplus \mathbb{E}^*, \mathbb{E} \oplus \mathbb{E}^* ; \Re)$.

It follows that $(\mathbf{v}\,, \mathbf{v}^*) \in \mathbf{D}$ implies that $\mathrm{EVAL}(\mathbf{v}\,, \mathbf{v}^*) = \langle \mathbf{v}, \mathbf{v}^* \rangle = 0$.

Definition 1.4.27 *Let* $\mathbf{p} \in \mathrm{C}^1(\mathbb{E} ; \mathbb{M})$ *and* $\mathbf{p}^* \in \mathrm{C}^1(\mathbb{E}^* ; \mathbb{M})$ *be dual vector bundles, and* $(\mathbf{v}\,, \mathbf{v}^*) \in \mathrm{C}^1(\mathbb{M} ; \mathbb{E})$ *and* $\mathbf{v}^* \in \mathrm{C}^1(\mathbb{M} ; \mathbb{E} \oplus \mathbb{E}^*)$ *be a section of their* WHITNEY *sum. The parallel transport of the covector field* $\mathbf{v}^* \in \mathrm{C}^1(\mathbb{M} ; \mathbb{E}^*)$ *along a flow* $\mathbf{Fl}^{\mathbf{u}}_{\lambda} \in \mathrm{C}^1(\mathbb{M} ; \mathbb{M})$ *with velocity* $\mathbf{u} \in \mathrm{C}^1(\mathbb{M} ; \mathbb{TM})$ *is defined by invariance:*

$$\mathrm{EVAL}(\mathbf{Fl}^{\mathbf{u}}_{\lambda}\Uparrow \mathbf{v}^*\,, \mathbf{Fl}^{\mathbf{u}}_{\lambda}\Uparrow \mathbf{v}) = \mathrm{EVAL}(\mathbf{v}^*\,, \mathbf{v})\,.$$

Proposition 1.4.6 *The **covariant derivative*** $\nabla_{\mathbf{u}}\mathbf{v}^* \in \mathrm{C}^1(\mathbb{M} ; \mathbb{E}^*)$ *is defined in terms of parallel transport, as*

$$\nabla_{\mathbf{u}}\mathbf{v}^*(\mathbf{x}) := \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{u}}_{\lambda}\Downarrow \mathbf{v}^*(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x}))\,.$$

Proof. The result follows from the relation

$$\begin{aligned}
\partial_{\lambda=0}\,\langle \mathbf{v}^*(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x})), \mathbf{v}(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x}))\rangle &= \partial_{\lambda=0}\,\langle \mathbf{Fl}^{\mathbf{u}}_{\lambda}\Downarrow \mathbf{v}^*(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x})), \mathbf{Fl}^{\mathbf{u}}_{\lambda}\Downarrow \mathbf{v}(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x}))\rangle \\
&= \langle \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{u}}_{\lambda}\Downarrow \mathbf{v}^*(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x})), \mathbf{v}(\mathbf{x})\rangle \\
&+ \langle \mathbf{v}^*(\mathbf{x}), \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{u}}_{\lambda}\Downarrow \mathbf{v}(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x}))\rangle \\
&= \langle \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{u}}_{\lambda}\Downarrow \mathbf{v}^*(\mathbf{Fl}^{\mathbf{u}}_{\lambda}(\mathbf{x})), \mathbf{v}(\mathbf{x})\rangle \\
&+ \langle \mathbf{v}^*(\mathbf{x}), \nabla_{\mathbf{u}}\mathbf{v}(\mathbf{x})\rangle\,,
\end{aligned}$$

which may be written as $d_{\mathbf{u}}\langle \mathbf{v}^*, \mathbf{v}\rangle = \langle \nabla_{\mathbf{u}}\mathbf{v}^*, \mathbf{v}\rangle + \langle \mathbf{v}^*, \nabla_{\mathbf{u}}\mathbf{v}\rangle$. ∎

1.4.15 Torsion tensor

Let us now consider a connection on the tangent bundle $\boldsymbol{\tau} \in C^1(\mathbb{TM};\mathbb{M})$. The lack of symmetry of the second covariant derivative of a scalar field $f \in C^2(\mathbb{M};\Re)$ is measured by

$$\begin{aligned}(\nabla d)_{\mathbf{vu}} f - (\nabla d)_{\mathbf{uv}} f &= d_{\mathbf{v}}\, d_{\mathbf{u}} f - d_{\mathbf{u}}\, d_{\mathbf{v}} f - d_{(\nabla_{\mathbf{v}} \mathbf{u})} f + d_{(\nabla_{\mathbf{u}} \mathbf{v})} f \\ &= d_{(\nabla_{\mathbf{u}} \mathbf{v})} f - d_{(\nabla_{\mathbf{v}} \mathbf{u})} f - d_{[\mathbf{u}\,,\mathbf{v}]} f \\ &= (\nabla_{\mathbf{u}} \mathbf{v} - \nabla_{\mathbf{v}} \mathbf{u} - [\mathbf{u}\,,\mathbf{v}])\, f\,.\end{aligned}$$

Definition 1.4.28 *The* ***torsion*** *of a connection* ∇ *is the vector-valued two-form* $\text{TORS} \in \Lambda^2(\mathbb{M};\mathbb{TM})$ *defined by:*

$$\text{TORS}(\mathbf{u},\mathbf{v}) := \nabla_{\mathbf{u}} \mathbf{v} - \nabla_{\mathbf{v}} \mathbf{u} - [\mathbf{u}\,,\mathbf{v}]\,,$$

for any pair of tangent vector fields $\mathbf{v},\mathbf{u} \in C^1(\mathbb{M};\mathbb{TM})$.

The torsion fibrewise linear, skew-symmetric in its arguments and tensorial as be shown by a direct application of Lemma 1.2.1 and relying on the property of the Lie derivative stated in Proposition 1.3.11 on page 73, formula *iii*). We may then write

$$(\nabla d)_{\mathbf{uv}} f - (\nabla d)_{\mathbf{vu}} f = d_{\text{TORS}(\mathbf{v},\mathbf{u})} f\,.$$

The torsion operator can be equivalently characterized as the $(1,2)$ tensor $\text{TORS} \in BL(\mathbb{TM}^2, \mathbb{T}^*\mathbb{M};\Re)$ defined by

$$\text{TORS}(\mathbf{u},\mathbf{v},\boldsymbol{\alpha}) := \langle \boldsymbol{\alpha}, \text{TORS}(\mathbf{u},\mathbf{v}) \rangle\,.$$

The vanishing of the torsion of a connection states that the second covariant derivative of any scalar field is symmetric.

Remark 1.4.2 *Let* F *be a* Banach *space and* $v \in C^1(\mathbb{M};F)$ *a vector-valued field on a manifold* $\mathbb{M}$ *modeled on a* Banach *space* E. *The directional derivative* $(d_{\mathbf{v}} v)(\mathbf{x})$ *depends linearly on* $\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ *and the second covariant derivative is given by*

$$(\nabla d)_{\mathbf{vu}}\, v := d_{\mathbf{v}}\, d_{\mathbf{u}}\, v - d_{(\nabla_{\mathbf{v}} \mathbf{u})}\, v\,.$$

We have that $d_{\mathbf{v}}\, d_{\mathbf{u}}\, v - d_{\mathbf{u}}\, d_{\mathbf{v}}\, v = d_{[\mathbf{v}\,,\mathbf{u}]} v$ *and hence the implication*

$$d_{\text{TORS}(\mathbf{v},\mathbf{u})}\, v = 0 \implies (\nabla d)_{\mathbf{vu}}\, v = (\nabla d)_{\mathbf{uv}}\, v\,.$$

Proposition 1.4.7 (Symmetry of Christoffel symbols) *The torsion of the connection vanishes if and only if the* CHRISTOFFEL *symbols corresponding to any system of coordinate vector fields* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ *are symmetric with respect to the lower pair of indices, that is*

$$\Gamma^k_{ij} = \Gamma^k_{ji} \,.$$

Proof. By proposition 1.3.10 we have that $[\mathbf{e}_i \,, \mathbf{e}_j] = 0$ and hence

$$\textsc{tors}(\mathbf{e}_i, \mathbf{e}_j) = \nabla_{\mathbf{e}_i} \mathbf{e}_j - \nabla_{\mathbf{e}_j} \mathbf{e}_i - [\mathbf{e}_i \,, \mathbf{e}_j] = 0 \iff \nabla_{\mathbf{e}_i} \mathbf{e}_j = \nabla_{\mathbf{e}_j} \mathbf{e}_i \,.$$

The lack of symmetry of the second covariant derivatives of a tensor field $\mathbf{T}$ is measured by

$$\nabla^2_{\mathbf{vu}} \mathbf{T} - \nabla^2_{\mathbf{uv}} \mathbf{T} = \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{T} - \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} \mathbf{T} - \nabla_{(\nabla_{\mathbf{v}} \mathbf{u})} \mathbf{T} + \nabla_{(\nabla_{\mathbf{u}} \mathbf{v})} \mathbf{T} \,.$$

which may be written

$$\nabla^2_{\mathbf{vu}} \mathbf{T} - \nabla^2_{\mathbf{uv}} \mathbf{T} = \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{T} - \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} \mathbf{T} - \nabla_{[\mathbf{v} \,, \mathbf{u}]} \mathbf{T} - \nabla_{\textsc{tors}(\mathbf{v}, \mathbf{u})} \mathbf{T} \,.$$

While the torsion of a connection provides the lack of symmetry of the second covariant derivative of scalar fields, the lack of symmetry of the second covariant derivative of a section $\mathbf{s} \in \mathrm{C}^2(\mathbb{M}\,; \mathbb{TM})$ of the tangent bundle $\boldsymbol{\tau} \in \mathrm{C}^1(\mathbb{TM}\,; \mathbb{M})$ along the tangent vector fields $\mathbf{v}, \mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TM})$ is measured by the curvature tensor, when the torsion vanishes. Indeed we have that:

$$\begin{aligned} \nabla^2_{\mathbf{vu}} \mathbf{s} - \nabla^2_{\mathbf{uv}} \mathbf{s} &= \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{s} - \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} \mathbf{s} - \nabla_{(\nabla_{\mathbf{v}} \mathbf{u})} \mathbf{s} + \nabla_{(\nabla_{\mathbf{u}} \mathbf{v})} \mathbf{s} \\ &= \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{s} - \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} \mathbf{s} - \nabla_{[\mathbf{v} \,, \mathbf{u}]} \mathbf{s} - \nabla_{\textsc{tors}(\mathbf{v}, \mathbf{u})} \mathbf{s} \,. \end{aligned}$$

The *curvature* of the connection ∇ on the vector bundle $\boldsymbol{\tau} \in \mathrm{C}^1(\mathbb{TM}\,; \mathbb{M})$ is defined by

$$\textsc{curv}(\mathbf{s})(\mathbf{v}, \mathbf{u}) := \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} \mathbf{s} - \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} \mathbf{s} - \nabla_{[\mathbf{v} \,, \mathbf{u}]} \mathbf{s} \,,$$

Accordingly, the lack of symmetry of the second covariant derivative of the cross section $\mathbf{s} \in \mathrm{C}^2(\mathbb{M}\,; \mathbb{TM})$ may be written as:

$$\nabla^2_{\mathbf{vu}} \mathbf{s} - \nabla^2_{\mathbf{uv}} \mathbf{s} = \textsc{curv}(\mathbf{s})(\mathbf{v}, \mathbf{u}) - \nabla_{\textsc{tors}(\mathbf{v}, \mathbf{u})} \mathbf{s} \,.$$

1.4.16 Formulas for curvature and torsion forms

From Theorem 37.15 of [80] we infer the following results.

Lemma 1.4.19 (A formula for the curvature) *Let a connection on a vector bundle* $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ *be assigned by a connector* $\mathbf{K}_{\mathbb{E}} \in C^1(\mathbb{TE};\mathbb{E})$ *. Then, for any section* $\mathbf{s} \in C^1(\mathbb{M};\mathbb{VE})$ *and any pair of vector fields* $\mathbf{u},\mathbf{v} \in C^1(\mathbb{M};\mathbb{TM})$ *, the curvature two-form* $\text{CURV}(\mathbf{s}) \in \Lambda^2(\mathbb{M};\mathbb{E})$ *, given by:*

$$\text{CURV}(\mathbf{s})(\mathbf{u},\mathbf{v}) := \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\mathbf{s} - \nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\mathbf{s} - \nabla_{[\mathbf{u},\mathbf{v}]}\mathbf{s}\,,$$

is equivalently expressed by:

$$\begin{aligned}\text{CURV}(\mathbf{s})(\mathbf{u},\mathbf{v}) &= (\mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ \mathbf{k}_{\mathbb{TE}} - \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}}) \circ T^2\mathbf{s} \circ T\mathbf{u} \circ \mathbf{v} \\ &= \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ (\partial_{\mu=0}\,\partial_{\lambda=0} - \partial_{\lambda=0}\,\partial_{\mu=0}) \circ \mathbf{s} \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}\,.\end{aligned}$$

Proof. The iterated covariant derivatives may be writtes as:

$$\begin{aligned}\nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\mathbf{s} &= \mathbf{K}_{\mathbb{E}} \circ T(\mathbf{K}_{\mathbb{E}} \circ T\mathbf{s} \circ \mathbf{v}) \circ \mathbf{u} \\ &= \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{v} \circ \mathbf{u}\,.\end{aligned}$$

Recalling from Lemma 1.2.10, page 57, the definition: $\mathbf{Vl}_{\mathbb{E}} = \partial_{t=0}\,\mathbf{mult}^{\,t}_{\mathbb{E}}$, by the fibre linearity of the connector $\mathbf{K}_{\mathbb{E}} \in C^1(\mathbb{TE};\mathbb{E})$ we have that

$$\begin{aligned}\mathbf{K}_{\mathbb{E}} \circ \mathbf{mult}^{\,t}_{\mathbb{TE}} &= \mathbf{mult}^{\,t}_{\mathbb{E}} \circ \mathbf{K}_{\mathbb{E}}\,, \\ T\mathbf{K}_{\mathbb{E}} \circ \mathbf{Vl}_{\mathbb{TE}} &= \mathbf{Vl}_{\mathbb{E}} \circ \mathbf{K}_{\mathbb{E}}\,.\end{aligned}$$

and, again by Lemma 1.2.10, that: $\mathbf{Vl}_{\mathbb{TE}} \circ T\mathbf{s} = T^2\mathbf{s} \circ \mathbf{Vl}_{\mathbb{TM}}$. Then we may write:

$$\begin{aligned}\mathbf{Vl}_{\mathbb{E}} \circ \nabla_{[\mathbf{u},\mathbf{v}]}\mathbf{s} &= \mathbf{Vl}_{\mathbb{E}} \circ \mathbf{K}_{\mathbb{E}} \circ T\mathbf{s} \circ [\mathbf{u},\mathbf{v}] \\ &= T\mathbf{K}_{\mathbb{E}} \circ \mathbf{Vl}_{\mathbb{TE}} \circ T\mathbf{s} \circ [\mathbf{u},\mathbf{v}] \\ &= T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ \mathbf{Vl}_{\mathbb{TM}} \circ [\mathbf{u},\mathbf{v}] \\ &= T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ (T\mathbf{v} \circ \mathbf{u} - \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{u} \circ \mathbf{v}) \\ &= P_{\mathrm{V}} \circ T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{v} \circ \mathbf{u} - P_{\mathrm{V}} \circ T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{u} \circ \mathbf{v} \\ &= P_{\mathrm{V}} \circ T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{v} \circ \mathbf{u} - P_{\mathrm{V}} \circ T\mathbf{K}_{\mathbb{E}} \circ \mathbf{k}_{\mathbb{T}^3\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{u} \circ \mathbf{v}\,.\end{aligned}$$

Now, being $\mathbf{K}_{\mathbb{E}} = \mathbf{vd}_{\mathbb{E}} \circ P_{\mathrm{V}}$, we get

$$\begin{aligned}\textsc{curv}(\mathbf{s})(\mathbf{u},\mathbf{v}) &= \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{v} \circ \mathbf{u} - \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{u} \circ \mathbf{v} \\ &\quad - \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{v} \circ \mathbf{u} + \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ \mathbf{k}_{\mathbb{T}^3\mathbb{E}} \circ T^2\mathbf{s} \circ T\mathbf{u} \circ \mathbf{v} \\ &= (\mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}} \circ \mathbf{k}_{\mathbb{T}^3\mathbb{E}} - \mathbf{K}_{\mathbb{E}} \circ T\mathbf{K}_{\mathbb{E}}) \circ T^2\mathbf{s} \circ T\mathbf{u} \circ \mathbf{v} \,.\end{aligned}$$

The second formula in the statement follows from Lemma 1.2.5. ■

Lemma 1.4.20 (A formula for the torsion) *Let a connection in the tangent bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ *be given by a connector* $\mathbf{K}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM})$ *. Then, for any pair of vector fields* $\mathbf{u},\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *, the torsion two-form* $\textsc{tors} \in \Lambda^2(\mathbb{M}\,;\mathbb{TM})$ *, defined by:*

$$\textsc{tors}(\mathbf{u},\mathbf{v}) := \nabla_{\mathbf{u}}\mathbf{v} - \nabla_{\mathbf{v}}\mathbf{u} - [\mathbf{u}\,,\mathbf{v}]\,,$$

is equivalently expressed by:

$$\begin{aligned}\textsc{tors}(\mathbf{u},\mathbf{v}) &= (\mathbf{K}_{\mathbb{TM}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} - \mathbf{K}_{\mathbb{TM}}) \circ T\mathbf{u} \circ \mathbf{v} \\ &= \mathbf{K}_{\mathbb{TM}} \circ (\partial_{\mu=0}\ \partial_{\lambda=0} - \partial_{\lambda=0}\ \partial_{\mu=0}) \circ \mathbf{Fl}^{\mathbf{u}}_{\mu} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \,.\end{aligned}$$

Proof. The result follows from the evaluations:

$$\begin{aligned}\nabla_{\mathbf{u}}\mathbf{v} &= \mathbf{K}_{\mathbb{TM}} \circ T\mathbf{v} \circ \mathbf{u}\,, \\ \nabla_{\mathbf{v}}\mathbf{u} &= \mathbf{K}_{\mathbb{TM}} \circ T\mathbf{u} \circ \mathbf{v}\,, \\ [\mathbf{u}\,,\mathbf{v}] &= \mathbf{K}_{\mathbb{TM}} \circ (T\mathbf{v} \circ \mathbf{u} - \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{u} \circ \mathbf{v}) \\ &= \mathbf{K}_{\mathbb{TM}} \circ T\mathbf{v} \circ \mathbf{u} - \mathbf{K}_{\mathbb{TM}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ T\mathbf{u} \circ \mathbf{v}\,.\end{aligned}$$

The second formula in the statement is inferred from the definition of the flip $\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{T}^2\mathbb{M})$ by observing that $T\mathbf{u} \circ \mathbf{v} = \partial_{\lambda=0}\ \partial_{\mu=0}\, \mathbf{Fl}^{\mathbf{u}}_{\mu} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda}$. ■

Lemma 1.4.21 (Symmetric connections are torsion free) *The following equivalences hold:*

$$\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{H} = \mathbf{H} \circ \mathbf{flip}_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}} \iff \mathbf{K}_{\mathbb{TM}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}} = \mathbf{K}_{\mathbb{TM}} \iff \textsc{tors} = 0\,.$$

Proof. The connector associated with a given horizontal lift is defined by the property (see Lemma 1.4.14): $\mathbf{K}_{\mathbb{TM}}(\mathbf{X}) = \mathbf{vd}_{\mathbb{TM}}(\mathbf{X} - \mathbf{H}(\boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}), T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}))$. The first equivalence follows from the relation:

$$\begin{aligned}(\mathbf{K}_{\mathbb{TM}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}})(\mathbf{X}) &= \mathbf{vd}_{\mathbb{TM}}(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}) -_{T\boldsymbol{\tau}_{\mathbb{M}}} \mathbf{H}(\boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X})), T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X})))\\
&= \mathbf{vd}_{\mathbb{TM}}(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}) -_{T\boldsymbol{\tau}_{\mathbb{M}}} \mathbf{H}(T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}, \boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X})))\\
&= \mathbf{vd}_{\mathbb{TM}}(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}) -_{T\boldsymbol{\tau}_{\mathbb{M}}} (\mathbf{H} \circ \mathbf{flip}_{\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}})(\boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}), T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}))\\
&= \mathbf{vd}_{\mathbb{TM}}(\mathbf{k}_{\mathbb{T}^2\mathbb{M}}(\mathbf{X}) -_{T\boldsymbol{\tau}_{\mathbb{M}}} (\mathbf{k}_{\mathbb{T}^2\mathbb{M}} \circ \mathbf{H})(\boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}), T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}))\\
&= (\mathbf{vd}_{\mathbb{TM}} \circ \mathbf{k}_{\mathbb{T}^2\mathbb{M}})(\mathbf{X} - \mathbf{H}(\boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}), T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X}))\\
&= \mathbf{vd}_{\mathbb{TM}}(\mathbf{X} - \mathbf{H}(\boldsymbol{\tau}_{\mathbb{TM}}(\mathbf{X}), T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{X})) = \mathbf{K}_{\mathbb{TM}}(\mathbf{X}).\end{aligned}$$

The second equivalence is a simple consequence of Lemma 1.4.20. ■

1.4.17 Pushed connections

Let us consider a diffeomorphism $\boldsymbol{\varphi} \in C^1(\mathbb{M}; \mathbb{N})$ between two differentiable manifolds $\mathbb{M}$ and $\mathbb{N}$. A connection ∇ on $\mathbb{M}$ induces a pushed connection $\boldsymbol{\varphi}\!\uparrow\!\nabla$ on $\mathbb{N}$ defined by

$$(\boldsymbol{\varphi}\!\uparrow\!\nabla)_{(\boldsymbol{\varphi}\uparrow\mathbf{u})}\, \boldsymbol{\varphi}\!\uparrow\!\mathbf{v} := \boldsymbol{\varphi}\!\uparrow\!(\nabla_{\mathbf{u}}\mathbf{v}).$$

The parallel transports associated with the connections ∇ and $\boldsymbol{\varphi}\!\uparrow\!\nabla$ are related by

$$\mathbf{Fl}_{\lambda}^{\boldsymbol{\varphi}\uparrow\mathbf{u}}\!\Uparrow(\boldsymbol{\varphi}\!\uparrow\!\mathbf{v}) := \boldsymbol{\varphi}\!\uparrow\!(\mathbf{Fl}_{\lambda}^{\mathbf{u}}\!\Uparrow\mathbf{v}).$$

Indeed we have that

$$\begin{aligned}(\boldsymbol{\varphi}\!\uparrow\!\nabla)_{(\boldsymbol{\varphi}\uparrow\mathbf{u})}\, \boldsymbol{\varphi}\!\uparrow\!\mathbf{v} &= \partial_{\lambda=0}\, \mathbf{Fl}_{\lambda}^{\boldsymbol{\varphi}\uparrow\mathbf{u}}\!\Uparrow(\boldsymbol{\varphi}\!\uparrow\!\mathbf{v}) \circ \mathbf{Fl}_{\lambda}^{\boldsymbol{\varphi}\uparrow\mathbf{u}} = \partial_{\lambda=0}\, \boldsymbol{\varphi}\!\uparrow\!(\mathbf{Fl}_{\lambda}^{\boldsymbol{\varphi}\uparrow\mathbf{u}}\!\Uparrow\mathbf{v}) \circ \boldsymbol{\varphi}\!\uparrow\!(\mathbf{Fl}_{\lambda}^{\mathbf{u}})\\
&= \partial_{\lambda=0}\, \boldsymbol{\varphi}\!\uparrow\!(\mathbf{Fl}_{\lambda}^{\boldsymbol{\varphi}\uparrow\mathbf{u}}\!\Uparrow\mathbf{v} \circ \mathbf{Fl}_{\lambda}^{\boldsymbol{\varphi}\uparrow\mathbf{u}}) = \boldsymbol{\varphi}\!\uparrow\!(\nabla_{\mathbf{u}}\mathbf{v}).\end{aligned}$$

According to the previous definition, the connection is natural with respect to the push and such is the LIE derivative by Proposition 1.3.4. Then we infer that

$$\mathrm{TORS}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi}\!\uparrow\!\mathrm{TORS}, \qquad \mathrm{CURV}^{\boldsymbol{\varphi}} = \boldsymbol{\varphi}\!\uparrow\!\mathrm{CURV},$$

and also that

$$(\boldsymbol{\varphi}\!\uparrow\!\nabla)^2_{(\boldsymbol{\varphi}\uparrow\mathbf{v}\,\boldsymbol{\varphi}\uparrow\mathbf{u})}\, \boldsymbol{\varphi}\!\uparrow\!\mathbf{s} - (\boldsymbol{\varphi}\!\uparrow\!\nabla)^2_{(\boldsymbol{\varphi}\uparrow\mathbf{u}\,\boldsymbol{\varphi}\uparrow\mathbf{v})}\, \boldsymbol{\varphi}\!\uparrow\!\mathbf{s} = \boldsymbol{\varphi}\!\uparrow\!(\nabla^2_{\mathbf{vu}}\,\mathbf{s} - \nabla^2_{\mathbf{uv}}\,\mathbf{s}).$$

Remark 1.4.3 *The definition of a pushed connection has a nice interpretation when dealing with differentiation in nD euclidean spaces $\{\mathbb{S}\,,\text{CAN}\}$ equipped with the canonical metric, in terms of curvilinear coordinates.*

Indeed let $\boldsymbol{\varphi} \in \mathrm{C}^1(\Re^n\,;\mathbb{S})$ be the diffeomorphism induced by a curvilinear coordinate system and $\mathbf{v} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$ a vector field. Then $\boldsymbol{\varphi}\!\downarrow\!\mathbf{v} \in \mathrm{C}^1(\Re^n\,;\Re^n)$ is the numerical vector field of the components associated to $\mathbf{v} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$. Denoting by d the usual derivative in $\mathbb{S}$ and by $\nabla = \boldsymbol{\varphi}\!\downarrow\! d$ the pushed connection in $\Re^n$, we have that

$$\nabla_{\boldsymbol{\varphi}\downarrow\mathbf{h}}\, \boldsymbol{\varphi}\!\downarrow\!\mathbf{v} = \boldsymbol{\varphi}\!\downarrow\!(d_{\mathbf{h}}\mathbf{v})\,.$$

This formula tells us that the numerical vector field of the components of the directional derivative $d_{\mathbf{h}}\mathbf{v}$ is the covariant derivative of the numerical vector field of the components of $\mathbf{v}$, performed according to the connection $\nabla = \boldsymbol{\varphi}\!\downarrow\! d$ in $\Re^n$. The explicit expression of this covariant derivative is provided by a direct computation, as in Proposition 1.4.3:

$$d_{\mathbf{h}}\mathbf{v} = d_{\mathbf{h}}(v^\alpha \mathbf{e}_\alpha) = (d_{h^\beta \mathbf{e}_\beta} v^\alpha)\,\mathbf{e}_\alpha + v^\alpha\,(d_{h^\beta \mathbf{e}_\beta}\mathbf{e}_\alpha) = h^\beta (d_{\mathbf{e}_\beta} v^\alpha + v^\gamma\,\Gamma^\alpha_{\beta\gamma})\,\mathbf{e}_\alpha\,.$$

Analogous reasoning and computations can be performed for tensor fields to get the formula for the directional derivative in terms of the covariant derivative of the matrix of the components, according to the connection $\nabla = \boldsymbol{\varphi}\!\downarrow\! d$ in $\Re^n$:

$$d_{\mathbf{h}}\mathbf{T} = d_{\mathbf{h}}(\mathbf{T}^{\alpha\beta}\mathbf{e}_\alpha \otimes \mathbf{e}_\beta) = h^\gamma (d_{\mathbf{e}_\gamma} T^{\alpha\beta} + T^{\alpha\tau}\,\Gamma^\beta_{\tau\gamma} + T^{\tau\beta}\,\Gamma^\alpha_{\tau\gamma})\,(\mathbf{e}_\alpha \otimes \mathbf{e}_\beta)\,.$$

1.4.18 Accelerations and pushes

On the basis of the previous results we can establish a general formula which relates the acceleration fields corresponding to a pair of flows related each other thru a diffeomorphism between manifolds.

Let $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ be a time dependent diffeomorphism between the manifolds $\mathbb{M}$ and $\mathbb{N}$. We then consider the associated flow

$$\boldsymbol{\varphi}_{t,s} := \boldsymbol{\varphi}_t \circ \boldsymbol{\varphi}_s^{-1} : \mathbb{N} \mapsto \mathbb{N}\,,$$

and denote by $\mathbf{v}_t \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{TN})$ the relevant velocity vector field, that is the field such that $\partial_{\tau=t}\,\boldsymbol{\varphi}_{\tau,s} = \mathbf{v}_t \circ \boldsymbol{\varphi}_{t,s}\,, \quad \boldsymbol{\varphi}_{s,s}(\mathbf{x}) = \mathbf{x}\,, \quad \forall\, \mathbf{x} \in \mathbb{N}$. Now let $\mathbf{u}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ and $\mathbf{w}_t \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{TN})$ be time dependent vector fields and $\mathbf{Fl}^{\mathbf{u}}_{t,s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ and $\mathbf{Fl}^{\mathbf{w}}_{t,s} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{N})$ be the relevant flows related by the push:

$$\boldsymbol{\varphi}_t \circ \mathbf{Fl}^{\mathbf{u}}_{t,s} = \mathbf{Fl}^{\mathbf{w}}_{t,s} \circ \boldsymbol{\varphi}_s\,.$$

The accelerations along the flows associated with the velocity fields

$$\mathbf{v}_t \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{TN})\,, \quad \mathbf{u}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})\,, \quad \mathbf{w}_t \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{TN})\,,$$

are given by the corresponding *material time derivatives*:

$$\begin{aligned}
\dot{\mathbf{v}}_t &:= \partial_{\tau=t}\,\mathbf{v}_\tau + (\boldsymbol{\varphi}_t{\uparrow}\nabla)_{\mathbf{v}_t}\,\mathbf{v}_t\,,\\
\dot{\mathbf{u}}_t &:= \partial_{\tau=t}\,\mathbf{u}_\tau + \nabla_{\mathbf{u}_t}\,\mathbf{u}_t\,,\\
\dot{\mathbf{w}}_t &:= \partial_{\tau=t}\,\mathbf{w}_\tau + (\boldsymbol{\varphi}_t{\uparrow}\nabla)_{\mathbf{w}_t}\,\mathbf{w}_t\,.
\end{aligned}$$

The covariant derivative $\nabla_{\mathbf{u}_t}\,\mathbf{u}_t$ is performed according to a connection ∇ on the manifold $\mathbb{M}$, while the covariant derivatives $(\boldsymbol{\varphi}_t{\uparrow}\nabla)_{\mathbf{v}_t}\,\mathbf{v}_t$ and $(\boldsymbol{\varphi}_t{\uparrow}\nabla)_{\mathbf{w}_t}\,\mathbf{w}_t$ are performed according to the pushed connection on the manifold $\mathbb{N}$.

Figure 1.24: Gaspard-Gustave de Coriolis (1792 - 1843)

We have then the following result.

Proposition 1.4.8 *If the connection ∇ is torsion-free, the acceleration of the transformed flow $\mathbf{Fl}^{\mathbf{w}}_{t,s} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{N})$ is given by*

$$\boxed{\dot{\mathbf{w}}_t = \boldsymbol{\varphi}_t{\uparrow}\dot{\mathbf{u}}_t + \dot{\mathbf{v}}_t + 2\,(\boldsymbol{\varphi}_t{\uparrow}\nabla)_{(\boldsymbol{\varphi}_t\uparrow\mathbf{u}_t)}\mathbf{v}_t}$$

where $\dot{\mathbf{u}}_t$ is the acceleration along the flow $\mathbf{Fl}^{\mathbf{w}}_{t,s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$, the term $\dot{\mathbf{v}}_t$ is the drag-acceleration due to the pushing flow $\boldsymbol{\varphi}_{t,s} \in \mathrm{C}^1(\mathbb{N}\,;\mathbb{N})$, and the term $2\,(\boldsymbol{\varphi}_t{\uparrow}\nabla)_{(\boldsymbol{\varphi}_t\uparrow\mathbf{u}_t)}\mathbf{v}_t$ is the Coriolis *acceleration.*

Proof. We make recourse to some previous results. Proposition 1.3.13 gives the relation between the velocity fields

$$\mathbf{w}_t = \mathbf{v}_t + \boldsymbol{\varphi}_t\!\uparrow\!\mathbf{u}_t \,.$$

From the expression of the acceleration along the flow $\mathbf{Fl}^{\mathbf{u}}_{t,s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, by performing the push along $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ we have that

$$\boldsymbol{\varphi}_t\!\uparrow\!\dot{\mathbf{u}}_t = \boldsymbol{\varphi}_t\!\uparrow\!\partial_{\tau=t}\,\mathbf{u}_\tau + \boldsymbol{\varphi}_t\!\uparrow\!(\nabla_{\mathbf{u}_t}\,\mathbf{u}_t)\,.$$

By definition of the pushed connection given in section 1.4.17, we can write

$$\boldsymbol{\varphi}_t\!\uparrow\!(\nabla_{\mathbf{u}_t}\,\mathbf{u}_t) = (\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\boldsymbol{\varphi}_t\uparrow\mathbf{u}_t)}\,\boldsymbol{\varphi}_t\!\uparrow\!\mathbf{u}_t = (\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\mathbf{w}_t-\mathbf{v}_t)}(\mathbf{w}_t-\mathbf{v}_t)\,.$$

Moreover we have that

$$\boldsymbol{\varphi}_t\!\uparrow = \boldsymbol{\varphi}_{t,s}\!\uparrow\!\boldsymbol{\varphi}_s\!\uparrow = \boldsymbol{\varphi}_{s,t}\!\downarrow\!\boldsymbol{\varphi}_s\!\uparrow\,,$$

and then

$$\begin{aligned}\boldsymbol{\varphi}_t\!\uparrow\!(\partial_{s=t}\,\mathbf{u}_s) &= \partial_{s=t}\,(\boldsymbol{\varphi}_{s,t}\!\downarrow\!\boldsymbol{\varphi}_s\!\uparrow\!\mathbf{u}_s)\\ &= \mathcal{L}_{t,\mathbf{v}_t}(\boldsymbol{\varphi}_t\!\uparrow\!\mathbf{u}_t) = [\mathbf{v}_t, \boldsymbol{\varphi}_t\!\uparrow\!\mathbf{u}_t\,,+]\partial_{s=t}\,(\boldsymbol{\varphi}_s\!\uparrow\!\mathbf{u}_s)\\ &= [\mathbf{v}_t, \mathbf{w}_t-\mathbf{v}_t\,,+]\partial_{s=t}\,(\mathbf{w}_s-\mathbf{v}_s)\,.\end{aligned}$$

The symmetry of the connection ∇ ensures that the pushed connection $\boldsymbol{\varphi}_t\!\uparrow\!\nabla$ is torsion-free too, so that

$$\begin{aligned}&\textsc{tors}(\mathbf{v}_t, \mathbf{w}_t-\mathbf{v}_t) = 0 \iff\\ &[\mathbf{v}_t, \mathbf{w}_t-\mathbf{v}_t\,,=](\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{\mathbf{v}_t}(\mathbf{w}_t-\mathbf{v}_t) - (\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\mathbf{w}_t-\mathbf{v}_t)}\mathbf{v}_t\,.\end{aligned}$$

Finally we have that

$$\begin{aligned}\boldsymbol{\varphi}_t\!\uparrow\!\dot{\mathbf{u}}_t = (\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{\mathbf{v}_t}(\mathbf{w}_t-\mathbf{v}_t) &-(\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\mathbf{w}_t-\mathbf{v}_t)}\mathbf{v}_t + \partial_{s=t}\,\mathbf{w}_s\\ &+(\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\mathbf{w}_t-\mathbf{v}_t)}(\mathbf{w}_t-\mathbf{v}_t) - \partial_{s=t}\,\mathbf{v}_s\,.\end{aligned}$$

By the properties of the covariant derivative, we can group as follows

$$\begin{aligned}\boldsymbol{\varphi}_t\!\uparrow\!\dot{\mathbf{u}}_t &= -2\,(\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\mathbf{w}_t-\mathbf{v}_t)}\mathbf{v}_t + (\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\mathbf{w}_t-\mathbf{v}_t)}\mathbf{w}_t\\ &\quad+(\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{\mathbf{v}_t}(\mathbf{w}_t-\mathbf{v}_t) + \partial_{s=t}\,\mathbf{w}_s - \partial_{s=t}\,\mathbf{v}_s\\ &= -2\,(\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{(\mathbf{w}_t-\mathbf{v}_t)}\mathbf{v}_t + (\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{\mathbf{w}_t}\mathbf{w}_t + \partial_{s=t}\,\mathbf{w}_s - (\boldsymbol{\varphi}_t\!\uparrow\!\nabla)_{\mathbf{v}_t}\mathbf{v}_t - \partial_{s=t}\,\mathbf{v}_s\,,\end{aligned}$$

which is the formula to be proven. ■

1.5 Lie groups and algebras

A *group* $\mathbf{G}$ is a set endowed with an operation $\mu : \mathbf{G} \times \mathbf{G} \mapsto \mathbf{G}$ called *group multiplication*, which is associative and admits a *unit* or *neutral* element $\mathbf{e} \in \mathbf{G}$ with the property that $\mu(\mathbf{e}, \mathbf{g}) = \mu(\mathbf{g}, \mathbf{e}) = \mathbf{g}$ for all $\mathbf{g} \in \mathbf{G}$, and a bijective map $\nu : \mathbf{G} \mapsto \mathbf{G}$, the *reversion*, such that

$$\mu(\nu(\mathbf{g}), \mathbf{g}) = \mu(\mathbf{g}, \nu(\mathbf{g})) = \mathbf{e}\,.$$

The unit element is unique since, if $\mathbf{e}_1, \mathbf{e}_2 \in \mathbf{G}$ are unit elements, it follows that: $\mathbf{e}_1 = \mu(\mathbf{e}_1, \mathbf{e}_2) = \mathbf{e}_2$. Moreover we have that $\nu(\mathbf{e}) = \mu(\nu(\mathbf{e}), \mathbf{e}) = \mathbf{e}$. It is customary to write simply $\mathbf{g}_1.\mathbf{g}_2$ for $\mu(\mathbf{g}_1, \mathbf{g}_2)$.

By the associativity of the group multiplication, the *reversion* map $\nu : \mathbf{G} \mapsto \mathbf{G}$ is uniquely defined and we have that $\nu \circ \nu = \mathbf{id}_{\mathbf{G}}$ and $\nu(\mathbf{a}.\mathbf{b}) = \nu(\mathbf{b}).\nu(\mathbf{a})$. Indeed, if $\overline{\nu} : \mathbf{G} \mapsto \mathbf{G}$ is another *reversion* map, we have that

$$\overline{\nu}(\mathbf{g}).\mathbf{g} = \mathbf{e} = \nu(\mathbf{g}).\mathbf{g} \implies (\overline{\nu}(\mathbf{g}).\mathbf{g}).\nu(\mathbf{g}) = (\nu(\mathbf{g}).\mathbf{g}).\nu(\mathbf{g}) \implies \overline{\nu}(\mathbf{g}) = \nu(\mathbf{g})\,.$$

Moreover

$$\begin{gathered} \mathbf{e} = \nu(\mathbf{a}.\mathbf{b}).(\mathbf{a}.\mathbf{b}) = (\nu(\mathbf{a}.\mathbf{b}).\mathbf{a}).\mathbf{b} \implies \\ \nu(\mathbf{b}) = \nu(\mathbf{a}.\mathbf{b}).\mathbf{a} \implies \nu(\mathbf{b}).\nu(\mathbf{a}) = \nu(\mathbf{a}.\mathbf{b})\,, \end{gathered}$$

and by the uniqueness of the reversion map we have also that

$$(\nu \circ \nu)(\mathbf{g}).\nu(\mathbf{g}) = \mathbf{e} = \nu(\mathbf{g}).\mathbf{g} = \mathbf{g}.\nu(\mathbf{g}) \implies (\nu \circ \nu)(\mathbf{g}) = \mathbf{g}\,,$$

which may be written as $\nu = \nu^{-1}$. It is customary to write $\mathbf{g}^{-1}$ for $\nu(\mathbf{g})$, so that $(\mathbf{g}^{-1})^{-1} = \mathbf{g}$.

If the group multiplication $\mu : \mathbf{G} \times \mathbf{G} \mapsto \mathbf{G}$ is commutative, the group $\mathbf{G}$ is said to be commutative.

- The *left* and *right* translations are the diffeomorphic maps $\lambda_{\mathbf{a}} \in \mathrm{C}^1(\mathbf{G}\,;\mathbf{G})$ and $\rho_{\mathbf{a}} \in \mathrm{C}^1(\mathbf{G}\,;\mathbf{G})$ defined by

$$\begin{cases} \lambda_{\mathbf{a}}\mathbf{g} = \mathbf{a}.\mathbf{g}\,, \\ \rho_{\mathbf{a}}\mathbf{g} = \mathbf{g}.\mathbf{a}\,, \end{cases} \qquad \mathbf{g} \in \mathbf{G}\,, \quad \forall\, \mathbf{a} \in \mathbf{G}\,.$$

Then $\lambda_{\mathbf{a}} \circ \lambda_{\mathbf{b}} = \lambda_{\mathbf{ab}}$ and $\rho_{\mathbf{a}} \circ \rho_{\mathbf{b}} = \rho_{\mathbf{ba}}$ so that $\lambda_{\mathbf{a}^{-1}} = \lambda_{\mathbf{a}}^{-1}$ and $\rho_{\mathbf{a}^{-1}} = \rho_{\mathbf{a}}^{-1}$. Moreover $\rho_{\mathbf{b}} \circ \lambda_{\mathbf{a}} = \lambda_{\mathbf{a}} \circ \rho_{\mathbf{b}}$.

1.5.1 Lie groups

- A LIE group $\mathbf{G}$ is a differentiable manifold endowed with a differentiable group operation $\mu \in \mathrm{C}^1(\mathbf{G}^2\,;\mathbf{G})$.

By the chain rule, the tangent map $T\mu \in \mathrm{C}^1(\mathbb{T}\mathbf{G}^2\,;\mathbb{T}\mathbf{G})$ is given by

$$T_{\{\mathbf{a}\,,\mathbf{b}\}}\mu \cdot \{\mathbf{X_a}\,,\mathbf{Y_b}\} = T_{\mathbf{a}}\rho_{\mathbf{b}} \cdot \mathbf{X_a} + T_{\mathbf{b}}\lambda_{\mathbf{a}} \cdot \mathbf{Y_b}\,,$$

where $\mathbf{a},\mathbf{b} \in \mathbf{G}$ and $\mathbf{X_a} \in \mathbb{T}_{\mathbf{a}}\mathbf{G}$, $\mathbf{Y_b} \in \mathbb{T}_{\mathbf{b}}\mathbf{G}$. The reversion map is also differentiable as stated by the next proposition.

Proposition 1.5.1 *The tangent map* $T\nu \in \mathrm{C}^1(\mathbb{T}\mathbf{G}\,;\mathbb{T}\mathbf{G})$ *of the reversion map* $\nu : \mathbf{G} \mapsto \mathbf{G}$ *is given by*

$$T_{\mathbf{a}}\nu = -(T_{\nu(\mathbf{a})}\rho_{\mathbf{a}})^{-1} \cdot T_{\mathbf{a}}\lambda_{\nu(\mathbf{a})}\,.$$

Proof. Let $\mathbf{g} \in \mathrm{C}^1(I\,;\mathbf{G})$ be a curve such that $\mathbf{g}(0) = \mathbf{a}$ and $\partial_{t=0}\,\mathbf{g}(t) = \mathbf{X_a}$. Then, differentiating $\mu(\nu(\mathbf{g}(t)),\mathbf{g}(t)) = \mathbf{e}$ we get

$$0 = \partial_{t=0}\,\mu(\nu(\mathbf{g}(t)),\mathbf{g}(t)) = T_{\nu(\mathbf{a})}\rho_{\mathbf{a}} \cdot T_{\mathbf{a}}\nu \cdot \mathbf{X_a} + T_{\mathbf{a}}\lambda_{\nu(\mathbf{a})} \cdot \mathbf{X_a}\,,$$

and hence the result. In particular we have that $T_{\mathbf{e}}\nu = -\mathbf{id}_{\mathbb{T}_{\mathbf{e}}\mathbf{G}}$. ■

1.5.2 Left and right invariant vector fields

- A vector field $\mathbf{v} \in \mathrm{C}^1(\mathbf{G}\,;\mathbb{T}\mathbf{G})$ on $\mathbf{G}$ is *left invariant* if $\lambda_{\mathbf{g}}{\uparrow}\mathbf{v} = \mathbf{v}$ for all $\mathbf{g} \in \mathbf{G}$, with the push given by $\lambda_{\mathbf{g}}{\uparrow}\mathbf{v} = T\lambda_{\mathbf{g}} \circ \mathbf{v} \circ \lambda_{\mathbf{g}^{-1}}$.
- A vector field $\mathbf{v} \in \mathrm{C}^1(\mathbf{G}\,;\mathbb{T}\mathbf{G})$ on $\mathbf{G}$ is *right invariant* if $\rho_{\mathbf{g}}{\uparrow}\mathbf{v} = \mathbf{v}$ for all $\mathbf{g} \in \mathbf{G}$, with the push given by $\rho_{\mathbf{g}}{\uparrow}\mathbf{v} = T\rho_{\mathbf{g}} \circ \mathbf{v} \circ \rho_{\mathbf{g}^{-1}}$.

For a left invariant vector field $\mathbf{v}$ we have that

$$\mathbf{v}(\mathbf{g}) = (\lambda_{\mathbf{a}}{\uparrow}\mathbf{v})(\mathbf{g}) = (T\lambda_{\mathbf{a}} \circ \mathbf{v} \circ \lambda_{\mathbf{a}^{-1}})(\mathbf{g})\,.$$

Hence, setting $\mathbf{b} = \mathbf{a}^{-1}.\mathbf{g}$, so that $\mathbf{g} = \mathbf{a}.\mathbf{b}$, we get

$$\mathbf{v}(\mathbf{a}.\mathbf{b}) = (\lambda_{\mathbf{a}}{\uparrow}\mathbf{v})(\mathbf{a}.\mathbf{b}) = (T\lambda_{\mathbf{a}} \circ \mathbf{v} \circ \lambda_{\mathbf{a}^{-1}})(\mathbf{a}.\mathbf{b}) = (T\lambda_{\mathbf{a}} \circ \mathbf{v})(\mathbf{b}) = T_{\mathbf{b}}\lambda_{\mathbf{a}} \cdot \mathbf{v}(\mathbf{b})\,.$$

A left invariant vector field $\mathbf{v}$ is thus uniquely determined by its value at the unit of the group.

More precisely, between the tangent space $\mathbb{T}_{\mathbf{e}}\mathbf{G}$ and the subgroup of left invariant vector fields on $\mathbf{G}$, there is a linear isomorphism defined by

$$\mathrm{L}_{\mathbf{X}}(\mathbf{a}) := T_{\mathbf{e}}\lambda_{\mathbf{a}} \cdot \mathbf{X}\,, \quad \forall\, \mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,.$$

Clearly we have that $\mathrm{L}_{\mathbf{X}}(\mathbf{e}) = \mathbf{X}$ and hence $\mathrm{L}_{\mathbf{X}}(\mathbf{a}) = T_{\mathbf{e}}\lambda_{\mathbf{a}} \cdot \mathrm{L}_{\mathbf{X}}(\mathbf{e})$. The property of being *left invariant* may be equivalently expressed by requiring that the vector field is λ-related to itself

$$T\lambda \cdot \mathrm{L}_{\mathbf{X}} = \mathrm{L}_{\mathbf{X}} \circ \lambda\,.$$

Since the Lie bracket $[\mathbf{u}\,,\mathbf{v}]$ of two vector fields on $\mathbf{G}$ is natural with respect to the push, the space of left invariant vector fields is a subalgebra of the Lie bracket algebra on $\mathbf{G}$:

$$\left.\begin{aligned} &\lambda_{\mathbf{g}}\!\uparrow\!\mathbf{u} = \mathbf{u}\,, \\ &\lambda_{\mathbf{g}}\!\uparrow\!\mathbf{v} = \mathbf{v}\,, \end{aligned}\right\} \Longrightarrow [\mathbf{u}\,,\mathbf{v}] = [\lambda_{\mathbf{g}}\!\uparrow\!\mathbf{u}\,,\lambda_{\mathbf{g}}\!\uparrow\!\mathbf{v}] = \lambda_{\mathbf{g}}\!\uparrow\![\mathbf{u}\,,\mathbf{v}]\,.$$

In turn this subalgebra defines the Lie algebra $\mathrm{LIE}\,(\mathbf{G})$ in the linear space $\mathbb{T}_{\mathbf{e}}\mathbf{G}$ by setting

$$[\mathbf{X}\,,\mathbf{Y}] := [\mathrm{L}_{\mathbf{X}}\,,\mathrm{L}_{\mathbf{Y}}](\mathbf{e})\,, \qquad \forall\, \mathbf{X},\mathbf{Y} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,.$$

Then

$$\mathrm{L}_{[\mathbf{X}\,,\mathbf{Y}]} = \mathrm{L}_{[\mathrm{L}_{\mathbf{X}}\,,\mathrm{L}_{\mathbf{Y}}](\mathbf{e})} = [\mathrm{L}_{\mathbf{X}}\,,\mathrm{L}_{\mathbf{Y}}]\,.$$

Analogous results hold for right invariant vector fields, which are generated by vectors of the tangent space $\mathbb{T}_{\mathbf{e}}\mathbf{G}$ according to the linear isomorphism

$$\mathrm{R}_{\mathbf{X}}(\mathbf{a}) := T_{\mathbf{e}}\rho_{\mathbf{a}} \cdot \mathbf{X}\,, \quad \forall\, \mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,.$$

The property of being *right invariant* may be equivalently expressed by requiring that the vector field is ρ-related to itself

$$T\rho \cdot \mathrm{R}_{\mathbf{X}} = \mathrm{R}_{\mathbf{X}} \circ \rho\,.$$

If $\mathbf{v}$ is a left (right) invariant vector field on $\mathbf{G}$, then the vector fields $\nu\!\downarrow\!\mathbf{v}$ and $\nu\!\uparrow\!\mathbf{v}$ are right (left) invariant. Indeed from the relations

$$\begin{aligned} &\rho_{\mathbf{a}} \circ \nu = \nu \circ \lambda_{\mathbf{a}^{-1}} \iff \mathbf{g}^{-1}.\mathbf{a} = (\mathbf{a}^{-1}.\mathbf{g})^{-1}\,, \\ &\nu \circ \rho_{\mathbf{a}} = \lambda_{\mathbf{a}^{-1}} \circ \nu \iff (\mathbf{a}.\mathbf{g})^{-1} = \mathbf{g}^{-1}.\mathbf{a}^{-1}\,, \end{aligned}$$

we get that

$$\rho_{\mathbf{a}}\!\uparrow\!(\nu\!\uparrow\!\mathbf{v}) = (\rho_{\mathbf{a}} \circ \nu)\!\uparrow\!\mathbf{v} = (\nu \circ \lambda_{\mathbf{a}^{-1}})\!\uparrow\!\mathbf{v} = \nu\!\uparrow\!(\lambda_{\mathbf{a}^{-1}}\!\uparrow \cdot\, \mathbf{v}) = \nu\!\uparrow\!\mathbf{v}\,,$$
$$\rho_{\mathbf{a}}\!\downarrow\!(\nu\!\downarrow\!\mathbf{v}) = (\nu \circ \rho_{\mathbf{a}})\!\downarrow\!\mathbf{v} = (\lambda_{\mathbf{a}^{-1}} \circ \nu)\!\downarrow\!\mathbf{v} = \nu\!\downarrow\!(\lambda_{\mathbf{a}^{-1}}\!\downarrow \cdot\, \mathbf{v}) = \nu\!\downarrow\!\mathbf{v}\,.$$

Recalling that $T_{\mathbf{e}}\nu = -\mathbf{id}_{\mathbb{T}_{\mathbf{e}}\mathbf{G}} \in BL(\mathbb{T}_{\mathbf{e}}\mathbf{G}\,;\mathbb{T}_{\mathbf{e}}\mathbf{G})$ is a linear isomorphism, the formula $\nu\!\downarrow\!\mathbf{v} := T\nu \circ \mathbf{v} \circ \nu$, tells us that

$$(\nu\!\downarrow\!\mathbf{v})(\mathbf{e}) = T_{\mathbf{e}}\nu \cdot \mathbf{v}(\mathbf{e}) = -\mathbf{v}(\mathbf{e})\,.$$

To any $\mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$ there correspond a left invariant vector field $\mathrm{L}_{\mathbf{X}} \in \mathrm{C}^1(\mathbf{G}\,;\mathbb{T}\mathbf{G})$ and a right invariant vector field $\mathrm{R}_{-\mathbf{X}} := \nu\!\downarrow\!\mathrm{L}_{\mathbf{X}} \in \mathrm{C}^1(\mathbf{G}\,;\mathbb{T}\mathbf{G})$ so that

$$\mathrm{R}_{[\mathbf{Y}\,,\mathbf{X}]} = \mathrm{R}_{-[\mathbf{X}\,,\mathbf{Y}]} = \nu\!\downarrow\!\mathrm{L}_{[\mathbf{X}\,,\mathbf{Y}]} = \nu\!\downarrow\![\mathrm{L}_{\mathbf{X}}\,,\mathrm{L}_{\mathbf{Y}}] = [\nu\!\downarrow\!\mathrm{L}_{\mathbf{X}}\,,\nu\!\downarrow\!\mathrm{L}_{\mathbf{Y}}] = [\mathrm{R}_{-\mathbf{X}}\,,\mathrm{R}_{-\mathbf{Y}}]\,,$$

and hence, by the bilinearity of the bracket $[\mathbf{X}\,,\mathbf{Y}] \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$:

$$[\mathrm{R}_{\mathbf{X}}\,,\mathrm{R}_{\mathbf{Y}}] = \mathrm{R}_{[-\mathbf{Y}\,,-\mathbf{X}]} = \mathrm{R}_{[\mathbf{Y}\,,\mathbf{X}]}\,.$$

We summarize with the following statement.

Proposition 1.5.2 *The* LIE *bracket of left and right invariant vector fields fulfill the properties:*

$$[\mathrm{L}_{\mathbf{X}}\,,\mathrm{L}_{\mathbf{Y}}] = \mathrm{L}_{[\mathbf{X}\,,\mathbf{Y}]}$$
$$[\mathrm{R}_{\mathbf{X}}\,,\mathrm{R}_{\mathbf{Y}}] = \mathrm{R}_{[\mathbf{Y}\,,\mathbf{X}]}\,.$$

Proposition 1.5.3 *The* LIE *bracket between a left and a right invariant vector field vanishes identically:*

$$[\mathrm{L}_{\mathbf{X}}\,,\mathrm{R}_{\mathbf{Y}}] = 0\,,$$

so that the flows of left and right invariant vector fields commute.

Proof. From the chain expression of the tangent map $T\mu \in \mathrm{C}^1(\mathbb{T}\mathbf{G}^2\,;\mathbb{T}\mathbf{G})$ provided on page 128, we get:

$$T_{\{\mathbf{a}\,,\mathbf{b}\}}\mu \cdot \{0_{\mathbf{a}}\,,\mathrm{L}_{\mathbf{X}}(\mathbf{b})\} = T_{\mathbf{a}}\rho_{\mathbf{b}} \cdot 0_{\mathbf{a}} + T_{\mathbf{b}}\lambda_{\mathbf{a}} \cdot \mathrm{L}_{\mathbf{X}}(\mathbf{b})\,,$$

and by the left invariance it is

$$T_{\mathbf{b}}\lambda_{\mathbf{a}} \cdot \mathrm{L}_{\mathbf{X}}(\mathbf{b}) = \mathrm{L}_{\mathbf{X}}(\mathbf{a}.\mathbf{b}) = (\mathrm{L}_{\mathbf{X}} \circ \mu)(\mathbf{a},\mathbf{b})\,,$$

so that we have the μ relatedness:

$$T\mu \cdot \{0\,,\mathrm{L}_{\mathbf{X}}\} = \mathrm{L}_{\mathbf{X}} \circ \mu\,.$$

In the same way we may prove that $T\mu \cdot \{\mathrm{R}_{\mathbf{Y}}\,, 0\} = \mathrm{R}_{\mathbf{Y}} \circ \mu\,.$ Hence, by Lemma 1.3.4:

$$T\mu \cdot [\{0\,, \mathrm{L}_{\mathbf{X}}\}\,, \{\mathrm{R}_{\mathbf{Y}}\,, 0\}] = [\mathrm{L}_{\mathbf{X}}\,, \mathrm{R}_{\mathbf{Y}}] \circ \mu\,,$$

and the result follows since $[\{0\,, \mathrm{L}_{\mathbf{X}}\}\,, \{\mathrm{R}_{\mathbf{Y}}\,, 0\}] = 0\,.$ ∎

Proposition 1.5.4 *Let* $\varphi \in \mathrm{C}^1(\mathbf{G}\,; \overline{\mathbf{G}})$ *be a homomorphism of* LIE *groups, so that* $\varphi \circ \lambda_{\mathbf{x}} = \lambda_{\varphi(\mathbf{x})} \circ \varphi$ *and* $\overline{\mathbf{e}} = \varphi(\mathbf{e})\,.$ *Then* $T_{\mathbf{e}}\varphi \in \mathrm{C}^1(\mathbb{T}_{\mathbf{e}}\mathbf{G}\,; \mathbb{T}_{\overline{\mathbf{e}}}\overline{\mathbf{G}})$ *is a homomorphism of* LIE *algebras:*

$$T_{\mathbf{e}}\varphi \cdot [\mathbf{X}\,, \mathbf{Y}] = [T_{\mathbf{e}}\varphi \cdot \mathbf{X} T_{\mathbf{e}}\varphi \cdot \mathbf{Y}\,,\,], \quad \forall\, \mathbf{X}, \mathbf{Y} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,.$$

Proof. We have that

$$\begin{aligned} T_{\mathbf{x}}\varphi \cdot \mathrm{L}_{\mathbf{X}}(\mathbf{x}) &= T_{\mathbf{x}}\varphi \cdot T_{\mathbf{e}}\lambda_{\mathbf{x}} \cdot \mathbf{X} = T_{\mathbf{e}}(\varphi \circ \lambda_{\mathbf{x}}) \cdot \mathbf{X} \\ &= T_{\mathbf{e}}(\lambda_{\varphi(\mathbf{x})} \circ \varphi) \cdot \mathbf{X} = T_{\mathbf{e}}\lambda_{\varphi(\mathbf{x})} \cdot T_{\mathbf{e}}\varphi \cdot \mathbf{X} \\ &= \mathrm{L}_{T_{\mathbf{e}}\varphi \cdot \mathbf{X}}(\varphi(\mathbf{x}))\,. \end{aligned}$$

Then $\mathrm{L}_{\mathbf{X}}$ is φ-related to $\mathrm{L}_{T_{\mathbf{e}}\varphi \cdot \mathbf{X}}$, i.e.:

$$T_{\mathbf{x}}\varphi \cdot \mathrm{L}_{\mathbf{X}} = \mathrm{L}_{T_{\mathbf{e}}\varphi \cdot \mathbf{X}} \circ \varphi\,.$$

From Lemma 1.3.4 it follows that the bracket $[\mathrm{L}_{\mathbf{X}}\,, \mathrm{L}_{\mathbf{Y}}]$ is φ-related to the bracket $[\mathrm{L}_{T_{\mathbf{e}}\varphi \cdot \mathbf{X}}\,, \mathrm{L}_{T_{\mathbf{e}}\varphi \cdot \mathbf{Y}}]\,.$ Hence

$$T\varphi \circ \mathrm{L}_{[\mathbf{X}\,, \mathbf{Y}]} = \mathrm{L}_{[T_{\mathbf{e}}\varphi \cdot \mathbf{X}\,, T_{\mathbf{e}}\varphi \cdot \mathbf{Y}]} \circ \varphi\,,$$

and the result follows by evaluating at $\mathbf{e}\,.$ ∎

1.5.3 One parameter subgroups

Given a LIE group $\mathbf{G}$ and the associated LIE algebra LIE $(\mathbf{G})$ a *one parameter subgroup* $\mathbf{c} \in \mathrm{C}^1(\{\Re\,, +\}\,; \mathbf{G})$ is a LIE group homomorphism from the addition group $\{\Re\,, +\}$ and the LIE group $\mathbf{G}\,.$ In other terms, a *one parameter subgroup* is a smooth curve in $\mathbf{G}$ with $\mathbf{c}(0) = \mathbf{e}$ and $\mathbf{c}(s + t) = \mathbf{c}(s).\mathbf{c}(t)\,.$

Proposition 1.5.5 *If* $\mathbf{c} \in \mathrm{C}^1(\Re\,;\mathbf{G})$ *is a curve with* $\mathbf{c}(0) = \mathbf{e}$*, setting* $\mathbf{X} = \partial_{t=0}\, \mathbf{c}(t) \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$*, the following assertions are equivalent, [80]:*

$$
\begin{aligned}
&1) \quad \mathbf{c} \quad \textit{is a one parameter subgroup}\,,\\
&2) \quad \mathbf{c}(t) = \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}}(\mathbf{e})\,,\\
&3) \quad \mathbf{c}(t) = \mathbf{Fl}_t^{\mathrm{R}_\mathbf{X}}(\mathbf{e})\,,\\
&4) \quad \mathbf{g}.\mathbf{c}(t) = \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}}(\mathbf{g}) \iff \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}} = \rho_{\mathbf{c}(t)}\,,\\
&5) \quad \mathbf{c}(t).\mathbf{g} = \mathbf{Fl}_t^{\mathrm{R}_\mathbf{X}}(\mathbf{g}) \iff \mathbf{Fl}_t^{\mathrm{R}_\mathbf{X}} = \lambda_{\mathbf{c}(t)}\,.
\end{aligned}
$$

Proof. Indeed the velocity of the flows $\rho_{\mathbf{c}(t)} \in \mathrm{C}^1(\mathbf{G}\,;\mathbf{G})$ and $\lambda_{\mathbf{c}(t)} \in \mathrm{C}^1(\mathbf{G}\,;\mathbf{G})$ are given by

$$
\begin{aligned}
\partial_{t=0}\, \rho_{\mathbf{c}(t)}(\mathbf{g}) &= \partial_{t=0}\, \mu(\mathbf{g}, \mathbf{c}(t)) = T_{\mathbf{e}}\lambda_{\mathbf{g}} \cdot \mathbf{X} = \mathrm{L}_{\mathbf{X}}(\mathbf{g})\,,\\
\partial_{t=0}\, \lambda_{\mathbf{c}(t)}(\mathbf{g}) &= \partial_{t=0}\, \mu(\mathbf{c}(t), \mathbf{g}) = T_{\mathbf{e}}\rho_{\mathbf{g}} \cdot \mathbf{X} = \mathrm{R}_{\mathbf{X}}(\mathbf{g})\,,
\end{aligned}
$$

so that $\rho_{\mathbf{c}(t)} = \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}}$ and $\lambda_{\mathbf{c}(t)} = \mathbf{Fl}_t^{\mathrm{R}_\mathbf{X}}$. ■

1.5.4 Exponential mapping

The *exponential mapping* $\mathrm{EXP} \in \mathrm{C}^1(\mathrm{LIE}\,(\mathbf{G})\,;\mathbf{G})$ of a LIE group $\mathbf{G}$ is the map defined by

$$
\mathrm{EXP}\,(\mathbf{X}) := \mathbf{Fl}_1^{\mathrm{L}_\mathbf{X}}(\mathbf{e}) = \mathbf{Fl}_1^{\mathrm{R}_\mathbf{X}}(\mathbf{e}) = \boldsymbol{\alpha}_{\mathbf{X}}(1)\,, \quad \forall\, \mathbf{X} \in \mathrm{LIE}\,(\mathbf{G})\,,
$$

where $\boldsymbol{\alpha}_{\mathbf{X}} \in \mathrm{C}^1(\{\Re\,,+\}\,;\mathbf{G})$ is the one parameter subgroup of $\mathbf{G}$ such that $\partial_{t=0}\, \boldsymbol{\alpha}(t) = \mathbf{X}$.

Proposition 1.5.6 *The exponential map enjoyces the following properties:*

$$
\begin{aligned}
&1) \quad \mathrm{EXP}\,(t\mathbf{X}) = \mathbf{Fl}_1^{\mathrm{L}_{t\mathbf{X}}}(\mathbf{e}) = \mathbf{Fl}_1^{t\mathrm{L}_\mathbf{X}}(\mathbf{e}) = \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}}(\mathbf{e}) = \boldsymbol{\alpha}_{\mathbf{X}}(t)\,,\\
&2) \quad \mathbf{g}.\mathrm{EXP}\,(t\mathbf{X}) = \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}}(\mathbf{g}) \iff \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}} = \rho_{\mathrm{EXP}\,(t\mathbf{X})}\,,\\
&3) \quad \mathrm{EXP}\,(t\mathbf{X}).\mathbf{g} = \mathbf{Fl}_t^{\mathrm{R}_\mathbf{X}}(\mathbf{g}) \iff \mathbf{Fl}_t^{\mathrm{R}_\mathbf{X}} = \lambda_{\mathrm{EXP}\,(t\mathbf{X})}\\
&4) \quad \mathrm{EXP}\,(0) = \mathbf{e}\,,\\
&5) \quad T\mathrm{EXP}\,(0) = \mathbf{id}_{\,\mathrm{LIE}\,(\mathbf{G})}\,.
\end{aligned}
$$

Proof. The first three properties are a consequence of Proposition 1.5.5. It is clear that $\text{EXP}\,(0) = \mathbf{e}$. Moreover, we have that

$$T\text{EXP}\,(0) \cdot \mathbf{X} = \partial_{t=0}\, \text{EXP}\,(t\mathbf{X}) = \partial_{t=0}\, \mathbf{Fl}_t^{\mathrm{L}_\mathbf{X}}(\mathbf{e}) = \mathrm{L}_\mathbf{X}(\mathbf{e}) = \mathbf{X}\,,$$

so that $T\text{EXP}\,(0) = \mathbf{id}_{\text{LIE}\,(\mathbf{G})}$. ■

1.5.5 Adjoint representation

- A *representation* of a LIE group $\mathbf{G}$ on a linear space V is an homomorphism of LIE groups $\rho \in \mathrm{C}^1(\mathbf{G}\,; \mathrm{GL}(V))$, where $\mathrm{GL}(V)$ is the general linear group of linear invertible maps on V.

According to Proposition 1.5.4, the tangent map $T_\mathbf{e}\rho \in \mathrm{C}^1(\text{LIE}\,(\mathbf{G})\,; BL\,(V\,;V))$ is a homomorphism of LIE algebras, [80]. An injective representation of a LIE group $\mathbf{G}$ is said to be *faithful*.

A representation of a LIE group $\mathbf{G}$ may be provided by defining first, for every $\mathbf{a} \in \mathbf{G}$, the *conjugation* $\text{CONJ}_\mathbf{a} \in \mathrm{C}^2(\mathbf{G}\,;\mathbf{G})$ by

$$\text{CONJ}_\mathbf{a}(\mathbf{g}) = \mathbf{a}.\mathbf{g}.\mathbf{a}^{-1}\,,$$

that is

$$\text{CONJ}_\mathbf{a} := \lambda_\mathbf{a} \circ \rho_{\mathbf{a}^{-1}} = \rho_{\mathbf{a}^{-1}} \circ \lambda_\mathbf{a}\,.$$

The conjugations fulfill the properties

$$\text{CONJ}_{\mathbf{a}.\mathbf{b}} = \text{CONJ}_\mathbf{a} \circ \text{CONJ}_\mathbf{b}\,,$$

$$\text{CONJ}_\mathbf{a}(\mathbf{x}.\mathbf{y}) = \text{CONJ}_\mathbf{a}(\mathbf{x}).\text{CONJ}_\mathbf{a}(\mathbf{y})\,.$$

- The *adjoint representation* of a LIE group $\mathbf{G}$ on the linear space $\text{LIE}\,(\mathbf{G})$ is the map $\text{ADJ} \in \mathrm{C}^1(\mathbf{G}\,; BL\,(\text{LIE}\,(\mathbf{G})\,; \text{LIE}\,(\mathbf{G})))$ defined as the tangent map to the conjugation at the unit of $\mathbf{G}$:

$$\text{ADJ} := T_\mathbf{e}\text{CONJ}\,.$$

Since conjugations are LIE groups automorphisms, by Proposition 1.5.4 the adjoint representation is a LIE algebra homomorphism:

$$\text{ADJ}_{\mathbf{a}.\mathbf{b}} = T_\mathbf{e}\text{CONJ}_{\mathbf{a}.\mathbf{b}} = T_\mathbf{e}\text{CONJ}_\mathbf{a} \circ T_\mathbf{e}\text{CONJ}_\mathbf{b} = \text{ADJ}_\mathbf{a} \circ \text{ADJ}_\mathbf{b}\,,$$

$$\text{ADJ}_\mathbf{a} \cdot [\mathbf{X}\,,\mathbf{Y}] = [\text{ADJ}_\mathbf{a} \cdot \mathbf{X}\,, \text{ADJ}_\mathbf{a} \cdot \mathbf{Y}]\,, \quad \forall\, \mathbf{a} \in \mathbf{G}\,, \quad \forall\, \mathbf{X}, \mathbf{Y} \in \text{LIE}\,(\mathbf{G})\,.$$

A simple calculation shows that

$$\begin{aligned}\mathrm{ADJ}_{\mathbf{a}} := T_{\mathbf{e}}\mathrm{CONJ}_{\mathbf{a}} = T_{\mathbf{e}}(\lambda_{\mathbf{a}} \circ \rho_{\mathbf{a}^{-1}}) = T_{\mathbf{a}^{-1}}\lambda_{\mathbf{a}} \cdot T_{\mathbf{e}}\rho_{\mathbf{a}^{-1}} \\ = T_{\mathbf{e}}(\rho_{\mathbf{a}^{-1}} \circ \lambda_{\mathbf{a}}) = T_{\mathbf{a}}\rho_{\mathbf{a}^{-1}} \cdot T_{\mathbf{e}}\lambda_{\mathbf{a}} \,.\end{aligned}$$

Let $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbf{G})$ be a path with $\mathbf{c}(0) = \mathbf{e}$ and velocity $\partial_{t=0}\,\mathbf{c}(t) = \mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$. Then, by Proposition 1.5.5, the velocity of the flow $\rho_{\mathbf{c}(t)} \in \mathrm{C}^1(\mathbf{G}\,;\mathbf{G})$ is $\partial_{t=0}\,\rho_{\mathbf{c}(t)} = \mathrm{L}_{\mathbf{X}}$ and we have that:

Proposition 1.5.7 *The adjoint representation of the* LIE *group* $\mathbf{G}$ *is given by*

$$\mathrm{ADJ}_{\mathbf{c}(t)} \cdot \mathbf{Y} = (T\mathbf{Fl}^{\mathrm{L}_{\mathbf{X}}}_{-t} \circ \mathrm{L}_{\mathbf{Y}} \circ \mathbf{Fl}^{\mathrm{L}_{\mathbf{X}}}_{t})(\mathbf{e}) = (\mathbf{Fl}^{\mathrm{L}_{\mathbf{X}}}_{t} \!\downarrow\! \mathrm{L}_{\mathbf{Y}})(\mathbf{e})\,, \quad \forall\, \mathbf{Y} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,.$$

Proof. Recalling the formula $\mathrm{L}_{\mathbf{Y}}(\mathbf{c}(t)) = T_{\mathbf{e}}\lambda_{\mathbf{c}(t)} \cdot \mathbf{Y}$, we have, for all $\mathbf{Y} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$:

$$\begin{aligned}\mathrm{ADJ}_{\mathbf{c}(t)} \cdot \mathbf{Y} &= T_{\mathbf{c}(t)}\rho_{\mathbf{c}(t)^{-1}} \cdot T_{\mathbf{e}}\lambda_{\mathbf{c}(t)} \cdot \mathbf{Y} \\ &= T_{\mathbf{c}(t)}\rho_{\mathbf{c}^{-1}(t)} \cdot \mathrm{L}_{\mathbf{Y}}(\mathbf{c}(t)) \\ &= (T\rho_{\mathbf{c}^{-1}(t)} \circ \mathrm{L}_{\mathbf{Y}} \circ \rho_{\mathbf{c}(t)})(\mathbf{e}) \\ &= (T\mathbf{Fl}^{\mathrm{L}_{\mathbf{X}}}_{-t} \circ \mathrm{L}_{\mathbf{Y}} \circ \mathbf{Fl}^{\mathrm{L}_{\mathbf{X}}}_{t})(\mathbf{e}) \\ &= (\mathbf{Fl}^{\mathrm{L}_{\mathbf{X}}}_{t} \!\downarrow\! \mathrm{L}_{\mathbf{Y}})(\mathbf{e})\,,\end{aligned}$$

and the result is proven. ∎

Proposition 1.5.8 *The adjoint representation of the* LIE *group* $\mathbf{G}$ *meets the property:*

$$\mathrm{L}_{\mathbf{X}}(\mathbf{a}) = \mathrm{R}_{\mathrm{ADJ}_{\mathbf{a}}\mathbf{X}}(\mathbf{a})\,, \qquad \mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,, \quad \mathbf{a} \in \mathbf{G}\,.$$

Proof. A simple computation:

$$\mathrm{L}_{\mathbf{X}}(\mathbf{a}) := T_{\mathbf{e}}\lambda_{\mathbf{a}} \cdot \mathbf{X} = T_{\mathbf{e}}\rho_{\mathbf{a}} \cdot T_{\mathbf{e}}(\rho_{\mathbf{a}^{-1}} + \lambda_{\mathbf{a}}) \cdot \mathbf{X} = \mathrm{R}_{\mathrm{ADJ}_{\mathbf{a}}\mathbf{X}}(\mathbf{a})\,, \quad \forall\, \mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,,$$

yields the result. ∎

- The *adjoint representation* of the LIE algebra LIE $(\mathbf{G})$ is the tangent map $\mathrm{ADJ} = T_{\mathbf{e}}\mathrm{ADJ} \in BL\,(\mathrm{LIE}\,(\mathbf{G})\,; BL\,(\mathrm{LIE}\,(\mathbf{G})\,; \mathrm{LIE}\,(\mathbf{G})))$.

Proposition 1.5.9 *The adjoint representation* $\mathrm{ADJ} = T_{\mathbf{e}}\mathrm{ADJ}$ *of the* LIE *algebra* LIE $(\mathbf{G})$ *is characterized by the property:*

$$\mathrm{ADJ}(\mathbf{X}) \cdot \mathbf{Y} = [\mathbf{X}\,, \mathbf{Y}]\,, \qquad \mathbf{X}, \mathbf{Y} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}\,,$$

which may be also written as $\mathrm{ADJ}(\mathbf{X}) = [\mathbf{X}\,, \bullet]$.

Proof. From Propositions 1.5.5 and 1.5.7 we infer that:

$$\begin{aligned}\mathrm{ADJ}(\mathbf{X}) \cdot \mathbf{Y} = T_{\mathbf{e}}\mathrm{ADJ}(\mathbf{X}) \cdot \mathbf{Y} &= \partial_{t=0}\, \mathrm{ADJ}_{\mathbf{c}(t)} \cdot \mathbf{Y} \\ &= \partial_{t=0}\, (\mathbf{Fl}_t^{\mathrm{L}_{\mathbf{X}}} {\downarrow} \mathrm{L}_{\mathbf{Y}})(\mathbf{e}) = [\mathrm{L}_{\mathbf{X}} \,, \mathrm{L}_{\mathbf{Y}}](\mathbf{e}) = [\mathbf{X} \,, \mathbf{Y}]\,,\end{aligned}$$

and the result is proven. ■

1.5.6 Maurer-Cartan form

The *canonical form* or MAURER-CARTAN *form* on a LIE group $\mathbf{G}$ is the differential one-form $\boldsymbol{\omega} \in \mathrm{C}^1(\mathbf{G}\,; \mathrm{LIE}\,(\mathbf{G}))$ with values in the LIE algebra $\mathrm{LIE}\,(\mathbf{G})$, defined by

$$\langle \boldsymbol{\omega}, \mathbf{v} \rangle(\mathbf{g}) := (T_{\mathbf{e}}\lambda_{\mathbf{g}})^{-1} \cdot \mathbf{v}(\mathbf{g}) = T_{\mathbf{e}}\lambda_{\mathbf{g}^{-1}} \cdot \mathbf{v}(\mathbf{g})\,, \quad \mathbf{g} \in \mathbf{G}\,, \quad \mathbf{v} \in \mathrm{C}^1(\mathbf{G}\,; \mathbb{T}\mathbf{G})\,.$$

The MAURER-CARTAN form is then pointwise defined by the rule which associates with the tangent vector $\mathbf{v}(\mathbf{g}) \in \mathbb{T}_{\mathbf{g}}\mathbf{G}$, the vector $\mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$ which is its generator by left invariance, i.e. such that:

$$\mathbf{v}(\mathbf{g}) = \mathrm{L}_{\mathbf{X}}(\mathbf{g}) = T_{\mathbf{e}}\lambda_{\mathbf{g}} \cdot \mathbf{X}\,.$$

Let us recall that the image of a one form $\boldsymbol{\omega} \in \mathrm{C}^1(\mathbf{G}\,; \mathrm{LIE}\,(\mathbf{G}))$ under the left translation $\lambda_{\mathbf{a}} \in \mathrm{C}^1(\mathbf{G}\,; \mathbf{G})$ is defined by

$$\langle \lambda_{\mathbf{a}}{\uparrow}\boldsymbol{\omega}, \lambda_{\mathbf{a}}{\uparrow}\mathbf{v} \rangle := \lambda_{\mathbf{a}}{\uparrow}\langle \boldsymbol{\omega}, \mathbf{v} \rangle\,,$$

or equivalently by

$$\lambda_{\mathbf{a}}{\downarrow}\langle \lambda_{\mathbf{a}}{\uparrow}\boldsymbol{\omega}, \mathbf{v} \rangle := \langle \boldsymbol{\omega}, \lambda_{\mathbf{a}}{\downarrow}\mathbf{v} \rangle\,.$$

Proposition 1.5.10 *The* MAURER-CARTAN *form is left invariant, i.e.:*

$$\lambda_{\mathbf{a}}{\uparrow}\boldsymbol{\omega} = \boldsymbol{\omega}\,,$$

and the image by a right translation is given by: $\rho_{\mathbf{a}}{\uparrow}\boldsymbol{\omega} = \mathrm{ADJ}_{\mathbf{a}} \circ \boldsymbol{\omega}\,.$

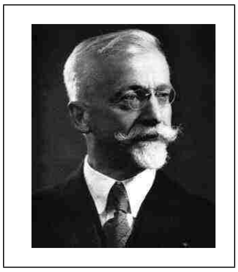

Figure 1.25: Elie Joseph Cartan (1869-1951)

Proof. To prove that $\langle \boldsymbol{\omega}, \lambda_{\mathbf{a}}\!\uparrow\!\mathbf{v} \rangle = \lambda_{\mathbf{a}}\!\uparrow\!\langle \boldsymbol{\omega}, \mathbf{v} \rangle$ we write

$$\begin{aligned}\langle \boldsymbol{\omega}, \lambda_{\mathbf{a}}\!\uparrow\!\mathbf{v} \rangle(\mathbf{g}) &= \langle \boldsymbol{\omega}(\mathbf{g}), T_{\mathbf{g}}\lambda_{\mathbf{a}} \cdot \mathbf{v}(\mathbf{a}^{-1}.\mathbf{g}) \rangle = (T_{\mathbf{e}}\lambda_{\mathbf{g}})^{-1} \cdot T_{\mathbf{g}}\lambda_{\mathbf{a}} \cdot \mathbf{v}(\mathbf{a}^{-1}.\mathbf{g}) \\ &= T_{\mathbf{e}}\lambda_{\mathbf{g}^{-1}} \cdot T_{\mathbf{g}}\lambda_{\mathbf{a}} \cdot \mathbf{v}(\mathbf{a}^{-1}.\mathbf{g}) = T_{\mathbf{e}}(\lambda_{\mathbf{g}^{-1}} \circ \lambda_{\mathbf{a}}) \cdot \mathbf{v}(\mathbf{a}^{-1}.\mathbf{g}) \\ &= T_{\mathbf{e}}\lambda_{\mathbf{g}^{-1}.\mathbf{a}} \cdot \mathbf{v}(\mathbf{a}^{-1}.\mathbf{g}) = T_{\mathbf{e}}\lambda_{(\mathbf{a}^{-1}.\mathbf{g})^{-1}} \cdot \mathbf{v}(\mathbf{a}^{-1}.\mathbf{g}) \\ &= \langle \boldsymbol{\omega}, \mathbf{v} \rangle(\mathbf{a}^{-1}.\mathbf{g}) = (\langle \boldsymbol{\omega}, \mathbf{v} \rangle \circ \lambda_{\mathbf{a}^{-1}})(\mathbf{g}) \,,\end{aligned}$$

and the former result is proven since $(\langle \boldsymbol{\omega}, \mathbf{v} \rangle \circ \lambda_{\mathbf{a}^{-1}})(\mathbf{g}) = \lambda_{\mathbf{a}}\!\uparrow\!\langle \boldsymbol{\omega}, \mathbf{v} \rangle$. The latter result is proven in a similar way. ■

1.5.7 Group actions

Let us denote by $\text{PERM}(\mathbb{M})$ the set of all the *permutations* of a set $\mathbb{M}$, i.e. of all the invertible maps from $\mathbb{M}$ onto itself.

A *left action* of a LIE group $\mathbf{G}$ on a set $\mathbb{M}$ is a group homomorphism $\ell \in \mathrm{C}^1(\mathbf{G}\,; \text{PERM}(\mathbb{M}))$.

The *left action* can be defined as a map $\ell \in \mathrm{C}^1(\mathbf{G} \times \mathbb{M}\,; \mathbb{M})$ such that, setting $\ell_{\mathbf{a}}(\mathbf{x}) = \ell^{\mathbf{x}}(\mathbf{a}) = \ell(\mathbf{a}, \mathbf{x})$ for $\mathbf{a} \in \mathbf{G}$ and $\mathbf{x} \in \mathbb{M}$, it is

$$\ell_{\mathbf{a}} \circ \ell_{\mathbf{b}} = \ell_{\mathbf{a}.\mathbf{b}} \,, \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbf{G} \,, \qquad \ell_{\mathbf{e}} = \mathbf{id}_{\mathbb{M}} \,,$$

A *right action* of a LIE group on a set $\mathbb{M}$ is a group anti-homomorphism $r \in \mathrm{C}^1(\mathbf{G}\,; \text{PERM}(\mathbb{M}))$ in which composition and multiplication are performed in reverse order.

A *right action* can be defined as a map $r \in \mathrm{C}^1(\mathbb{M} \times \mathbf{G}\,; \mathbb{M})$ such that, setting $r_{\mathbf{x}}(\mathbf{a}) = r^{\mathbf{a}}(\mathbf{x}) = r(\mathbf{x}, \mathbf{a})$ for $\mathbf{a} \in \mathbf{G}$ and $\mathbf{x} \in \mathbb{M}$, it is

$$r^{\mathbf{a}} \circ r^{\mathbf{b}} = r^{\mathbf{b}.\mathbf{a}} \,, \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbf{G} \,, \qquad r^{\mathbf{e}} = \mathbf{id}_{\mathbb{M}} \,.$$

A $\mathbf{G}$-*space* is a manifold $\mathbb{M}$ together with a left or right action of a LIE group $\mathbf{G}$ on $\mathbb{M}$. The following definitions are identical for left and right actions.

- The *orbit* through $\mathbf{x} \in \mathbb{M}$ is the subset $\ell(\mathbf{G}, \mathbf{x}) \subseteq \mathbb{M}$.
- An action is *transitive* if $\mathbb{M}$ is an orbit. This means that given $\mathbf{x}, \mathbf{y} \in \mathbb{M}$ there exists a $\mathbf{g} \in \mathbf{G}$ such that $\ell(\mathbf{g}, \mathbf{x}) = \mathbf{y}$.
- An action is *free* if $\ell_{\mathbf{a}}$ has a fixed point only if $\mathbf{a} = \mathbf{e}$. This means that if $\ell(\mathbf{a}, \mathbf{x}) = \mathbf{x}$ for some $\mathbf{x} \in \mathbb{M}$, then $\mathbf{a} = \mathbf{e}$.
- An action is *effective* if the group homomorphism $\ell \in \mathrm{C}^1(\mathbf{G}\,; \mathrm{PERM}(\mathbb{M}))$ is injective. This means that, given $\mathbf{a}, \mathbf{b} \in \mathbf{G}$ if $\ell_{\mathbf{a}} = \ell_{\mathbf{b}}$, then $\mathbf{a} = \mathbf{b}$.

Proposition 1.5.11 *An action is both transitive and free if and only if for any* $\mathbf{x}, \mathbf{b} \in \mathbb{M}$ *there is a unique* $\mathbf{g} \in \mathbf{G}$ *such that* $\ell_{\mathbf{g}}(\mathbf{x}) = \mathbf{y}$*.*

Proof. In the *only if* statement, existence is ensured by transitivity, so let us prove uniqueness. Indeed, if $\ell_{\mathbf{g}}(\mathbf{x}) = \ell_{\mathbf{h}}(\mathbf{x})$ then

$$\mathbf{x} = (\ell_{\mathbf{g}}^{-1} \circ \ell_{\mathbf{h}})(\mathbf{x}) = (\ell_{\mathbf{g}^{-1}} \circ \ell_{\mathbf{h}})(\mathbf{x}) = \ell_{\mathbf{g}^{-1}.\mathbf{h}}(\mathbf{x}) .$$

Since the action is free, this implies that $\mathbf{g}^{-1}.\mathbf{h} = \mathbf{e}$, that is $\mathbf{g} = \mathbf{h}$. The *if* statement follows by observing that the property implies transitivity. To prove that the action is free, let us assume that $\mathbf{x}$ is a fixed point of $\ell_{\mathbf{g}}$. Then $\ell_{\mathbf{g}}(\mathbf{x}) = \mathbf{x} = \ell_{\mathbf{e}}(\mathbf{x})$ and, by uniqueness, $\mathbf{g} = \mathbf{e}$. ■

Proposition 1.5.12 *A transitive action of a commutative group is free.*

Proof. Let $\ell_{\mathbf{a}}(\mathbf{x}) = \mathbf{x}$. Then, by transitivity, we may set $\mathbf{y} = \ell_{\mathbf{b}}(\mathbf{x})$ and by commutativity we get

$$\begin{aligned} \ell_{\mathbf{a}}(\mathbf{b}) = \ell_{\mathbf{a}}(\ell_{\mathbf{b}}(\mathbf{x})) &= (\ell_{\mathbf{a}} \circ \ell_{\mathbf{b}})(\mathbf{x}) = \ell_{\mathbf{a}.\mathbf{b}}(\mathbf{x}) \\ &= \ell_{\mathbf{b}.\mathbf{a}}(\mathbf{x}) = (\ell_{\mathbf{b}} \circ \ell_{\mathbf{a}})(\mathbf{x}) = \ell_{\mathbf{b}}(\mathbf{x}) = \mathbf{y} . \end{aligned}$$

By the arbitrarity of $\mathbf{y} \in \mathbb{M}$ we conclude that $\ell_{\mathbf{a}} = \mathbf{id}_{\mathbb{M}}$ and hence from the previous Proposition 1.5.11 we infer that $\mathbf{a} = \mathbf{e}$. ■

1.5.8 Killing vector fields

Let us consider a *left action* $\ell \in \mathrm{C}^1(\mathbf{G} \times \mathbb{M}\,;\mathbb{M})$ of a LIE group $\mathbf{G}$ at a point $\mathbf{x} \in \mathbb{M}$ and a *one parameter subgroup* $\mathbf{c} \in \mathrm{C}^1(\{\Re\,,+\}\,;\mathbf{G})$ of the LIE group, with velocity $\mathbf{X} := \partial_{t=0}\,\mathbf{c}(t) \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$. The composition $\ell^{\mathbf{x}} \circ \mathbf{c} \in \mathrm{C}^1(\Re\,;\mathbb{M})$ defines a one parameter transformation group on the manifold $\mathbb{M}$. The corresponding flow is given by $\ell_{\mathbf{c}(t)} = \ell_{\mathrm{EXP}\,(t\mathbf{X})} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ and the velocity field

$$\zeta_{\mathbf{X}} := \partial_{t=0}\,\ell_{\mathbf{c}(t)} = \partial_{t=0}\,\ell_{\mathrm{EXP}\,(t\mathbf{X})} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})\,,$$

is the KILLING *vector field* or the *infinitesimal generator* of the left action $\ell \in \mathrm{C}^1(\mathbf{G} \times \mathbb{M}\,;\mathbb{M})$ corresponding to the vector $\mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$.

The corresponding flow is given by $\mathbf{Fl}_t^{\zeta_{\mathbf{X}}} = \ell_{\mathrm{EXP}\,(t\mathbf{X})} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$. The chain rule shows that the KILLING vector field depends linearly on the tangent vector $\mathbf{X} := \partial_{t=0}\,\mathbf{c}(t) \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$ since:

$$\zeta_{\mathbf{X}}(\mathbf{x}) = T_{\mathbf{e}}\ell^{\mathbf{x}} \cdot \mathbf{X} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,,$$

with $T_{\mathbf{e}}\ell^{\mathbf{x}} \in BL\,(\mathbb{T}_{\mathbf{e}}\mathbf{G}\,;\mathbb{T}_{\mathbf{x}}\mathbb{M})$, equivalent to $\zeta_{\mathbf{X}}(\mathbf{x}) = T_{(\mathbf{e},\mathbf{x})}\ell \cdot \{\mathbf{X}, 0_{\mathbf{x}}\} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$. Moreover we have that the equality

$$\mathbf{Fl}_t^{\zeta_{\mathbf{X}}}(\mathbf{x}) = \ell^{\mathbf{x}}(\mathrm{EXP}\,(t\mathbf{X})) = \ell^{\mathbf{x}}(\lambda_{\mathrm{EXP}\,(t\mathbf{X})}\mathbf{e}) = (\ell^{\mathbf{x}} \circ \mathbf{Fl}_t^{\mathrm{R}_{\mathbf{X}}})(\mathbf{e})\,,$$

defines pointwise the relation which transforms the flow associated with a right invariant vector field on a LIE group into the corresponding flow associated with the KILLING vector field: $\mathbf{Fl}_t^{\zeta_{\mathbf{X}}} = \Lambda \circ \mathbf{Fl}_t^{\mathrm{R}_{\mathbf{X}}}$ with the same generator $\mathbf{X} \in \mathbb{T}_{\mathbf{e}}\mathbf{G}$. Hence, the vector bundle homomorphism $T\Lambda \in \mathrm{C}^0(\mathbb{T}\mathbf{G}\,;\mathbb{TM})$ gives:

$$\zeta_{\mathbf{X}} = \partial_{t=0}\,\mathbf{Fl}_t^{\zeta_{\mathbf{X}}}(\mathbf{x}) = \partial_{t=0}\,\Lambda \circ \mathbf{Fl}_t^{\mathrm{R}_{\mathbf{X}}} = T\Lambda \circ \mathrm{R}_{\mathbf{X}}\,.$$

Proposition 1.5.13 *The* KILLING *vector field associated with a left action of a group* $\mathbf{G}$*, meets the properties:*

$$T_{\mathbf{x}}\ell_{\mathbf{a}} \cdot \zeta_{\mathbf{X}}(\mathbf{x}) = \zeta_{\mathrm{ADJ}_{\mathbf{a}} \cdot \mathbf{X}}(\mathbf{a}.\mathbf{x})\,, \qquad [\zeta_{\mathbf{X}}\,,\zeta_{\mathbf{Y}}] = -\zeta_{[\mathbf{X}\,,\mathbf{Y}]}\,.$$

Proof. By acting with the tangent map $T\ell_{\mathbf{a}}(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}_{\ell(\mathbf{a},\mathbf{x})}\mathbb{M})$ on both sides of the defining formula $\zeta_{\mathbf{X}}(\mathbf{x}) = T\ell^{\mathbf{x}}(\mathbf{e}) \cdot \mathbf{X}$, we get

$$T_{\mathbf{x}}\ell_{\mathbf{a}} \cdot \zeta_{\mathbf{X}}(\mathbf{x}) = (T_{\mathbf{x}}\ell_{\mathbf{a}} \circ T_{\mathbf{e}}\ell^{\mathbf{x}}) \cdot \mathbf{X} = T_{\mathbf{e}}(\ell_{\mathbf{a}} \circ \ell^{\mathbf{x}}) \cdot \mathbf{X}\,.$$

But $(\ell_{\mathbf{a}} \circ \ell^{\mathbf{x}})(\mathbf{g}) = \mathbf{a}.\mathbf{g}.\mathbf{x} = (\mathbf{a}.\mathbf{g}.\mathbf{a}^{-1}).\mathbf{a}.\mathbf{x} = \mathrm{CONJ}_{\mathbf{a}}(\mathbf{g}).\mathbf{a}.\mathbf{x} = \ell_{\mathbf{a}.\mathbf{x}}(\mathrm{CONJ}_{\mathbf{a}}(\mathbf{g}))$ and hence

$$T_{\mathbf{e}}(\ell_{\mathbf{a}} \circ \ell^{\mathbf{x}}) \cdot \mathbf{X} = (T_{\mathbf{e}}\ell^{\mathbf{a}.\mathbf{x}} \circ T_{\mathbf{e}}\mathrm{CONJ}_{\mathbf{a}}) \cdot \mathbf{X} = (T_{\mathbf{e}}\ell^{\mathbf{a}.\mathbf{x}} \circ \mathrm{ADJ}_{\mathbf{a}}) \cdot \mathbf{X} = \zeta_{\mathrm{ADJ}_{\mathbf{a}} \cdot \mathbf{X}}(\mathbf{a}.\mathbf{x})\,,$$

and the first formula is proved. By Lemma 1.3.4 and Proposition 1.5.2 we have:

$$[\zeta_{\mathbf{X}}\,,\zeta_{\mathbf{Y}}] = [T\Lambda \circ \mathrm{R}_{\mathbf{X}}\,, T\Lambda \circ \mathrm{R}_{\mathbf{Y}}] = T\Lambda \circ [\mathrm{R}_{\mathbf{X}}\,, \mathrm{R}_{\mathbf{Y}}] = T\Lambda \circ \mathrm{R}_{[\mathbf{Y}\,,\mathbf{X}]} = \zeta_{[\mathbf{Y}\,,\mathbf{X}]}\,,$$

and this proves the second formula. ■

If a right action is considered instead of a left one, similar results holds, and the last result in Proposition 1.5.13 is changed into: $[\zeta_{\mathbf{X}}\,,\zeta_{\mathbf{Y}}] = \zeta_{[\mathbf{X}\,,\mathbf{Y}]}$. If the LIE group $\mathbf{G}$ acts effectively on the manifold $\mathbb{M}$, the linear space of KILLING vector fields is isomorphic to the algebra $\mathrm{LIE}\,(\mathbf{G})$.

1.6 Integration on manifolds

The integral of vector or tensor fields on a manifold makes in general no sense since the sum of the values of vector or tensor fields at different points of a manifold is not defined. Integration over a manifold is defined only for special tensor fields called volume-forms. The definition of volume-forms and of their integrals on compact manifolds is illustrated in the next subsection.

1.6.1 Exterior and differential forms

Let us give the following definition.

Definition 1.6.1 *A **exterior form**, or k-**form**, or form of **degree** k at $\mathbf{x} \in \mathbb{M}$ is a real valued k-linear skew-symmetric map $\boldsymbol{\omega}^k_{\mathbf{x}} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}^k ; \Re)$.*

The linear space of all the k-forms at $\mathbf{x} \in \mathbb{M}$ is denoted by $\Lambda^k_{\mathbf{x}}(\mathbb{M} ; \Re)$. The value of a k-form vanishes if one of its arguments is linearly dependent on the others. It follows that all k-forms with $k \geq n$ vanish identically.

Definition 1.6.2 *The **exterior** (or **wedge**) product of two forms $\boldsymbol{\omega}^k \wedge \boldsymbol{\omega}^h$ is defined by:*

$$(\boldsymbol{\omega}^k \wedge \boldsymbol{\omega}^h)(\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots \mathbf{e}_{k+h}) :=$$

$$\frac{1}{k!\, h!} \sum_{\pi} \mathrm{sign}(\pi)\, \boldsymbol{\omega}^k(\mathbf{e}_{\pi(1)}, \ldots, \mathbf{e}_{\pi(k)})\, \boldsymbol{\omega}^h(\mathbf{e}_{\pi(k+1)}, \ldots \mathbf{e}_{\pi(k+h)}) \,.$$

where the sum is taken over all permutation such that $\pi(1) < \ldots < \pi(k)$ and $\pi(k+1) < \ldots < \pi(k+h)$.

The following associative and commutation rules hold:

$$\boldsymbol{\omega}^k \wedge (\boldsymbol{\omega}^h \wedge \boldsymbol{\omega}^l) = (\boldsymbol{\omega}^k \wedge \boldsymbol{\omega}^h) \wedge \boldsymbol{\omega}^l \,, \qquad \boldsymbol{\omega}^k \wedge \boldsymbol{\omega}^h = (-1)^{kh}\, \boldsymbol{\omega}^h \wedge \boldsymbol{\omega}^k \,.$$

- The linear space $\Lambda_{\mathbf{x}}(\mathbb{M} ; \Re)$ of all real valued exterior forms at $\mathbf{x} \in \mathbb{M}$ is then a *graded commutative algebra* with respect to the exterior product, called the Grassmann *algebra*.

Let $\dim \mathbb{M} = n$ and $\{x^1, \ldots, x^n\}$ be a local coordinate system on $\mathbb{M}$. Then a basis of the linear space $\Lambda^k(\mathbb{T}_{\mathbf{x}}\mathbb{M} ; \Re)$ is provided by the family of k-th exterior products of the one-forms

$$\{dx^{i_1}, \ldots, dx^{i_k}\} \,, \qquad 1 \leq i_1 < \ldots < i_k \leq n \,,$$

which are the differentials of the coordinates.

Figure 1.26: Hermann Günter Grassmann (1809 - 1877)

The dimension of the linear space $\Lambda^k(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\Re)$ is then $n!/(k!(n-k)!)$. It follows that the dimension of $\Lambda^n_{\mathbf{x}}$ is one. The n-forms are called volume forms, and hence:

- All the volume forms $\boldsymbol{\omega}^n_{\mathbf{x}} \in \Lambda^n(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\Re)$ at $\mathbf{x} \in \mathbb{M}$ are proportional one another.

A volume form $\boldsymbol{\mu}^n_{\mathbf{x}} \in \Lambda^n_{\mathbf{x}}(\mathbb{M}\,;\Re)$ may then be choosen as *standard volume form* and all others will be proportional to it.

The value of $\boldsymbol{\mu}^n_{\mathbf{x}}$ on an n-tuple $\{\mathbf{u}_i\} \in (\mathbb{T}_{\mathbf{x}}\mathbb{M})^n$ of tangent vectors provides the standard signed-volume of the parallelepiped with edges $\{\mathbf{u}_i\}$.

- A *differential k-form* on a n-dimensional manifold $\mathbb{M}$ is a differentiable field $\boldsymbol{\omega}^k \in \Lambda^k(\mathbb{M}\,;\Re)$ of k-forms on $\mathbb{M}$. Any differential n-form $\boldsymbol{\omega}^n \in \Lambda^n(\mathbb{M}\,;\Re)$ on the n-dimensional manifold $\mathbb{M}$ is proportional to the standard differential volume form $\boldsymbol{\mu}^n \in \Lambda^n(\mathbb{M}\,;\Re)$.
- The *contraction* (or *insertion*) operator $\mathbf{i} : \Lambda^k_{\mathbf{x}}(\mathbb{M}\,;\Re) \mapsto \Lambda^{k-1}_{\mathbf{x}}(\mathbb{M}\,;\Re)$ is defined, for $\boldsymbol{\omega} \in \Lambda^k(\mathbb{M}\,;\Re)$, by the identity

$$(\mathbf{i}_{\mathbf{h}}\boldsymbol{\omega})(\mathbf{v}_1, \ldots \mathbf{v}_{(k-1)}) := \boldsymbol{\omega}(\mathbf{h}, \mathbf{v}_1, \ldots, \mathbf{v}_{(k-1)}),$$

for all $\mathbf{v}_1, \ldots, \mathbf{v}_{(k-1)} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$.

We shall also write simply $\boldsymbol{\omega}\mathbf{h}$ instead of $\mathbf{i}_{\mathbf{h}}\boldsymbol{\omega}$ when no confusion may occur.

1.6.2 Volume forms and Gram operator

Let the n-dimensional manifold $\mathbb{M}$ be endowed with a metric

$$\mathbf{g} \in \mathrm{C}^1(\mathbb{M}\,;BL\,(\mathbb{T}\mathbb{M}^2\,;\Re))$$

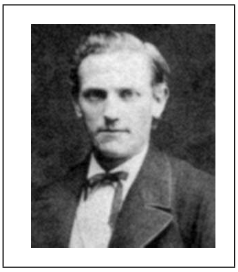

Figure 1.27: Jorgen Pedersen Gram (1850 - 1916)

which is a field of twice covariant symmetric and positive definite tensors.

- The GRAM operator $G \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}^n, \mathbb{T}_{\mathbf{x}}\mathbb{M}^n ; BL(\Re^n ; \Re^n))$ associated with the metric $\mathbf{g}$ is then defined by

$$G_{ij}(\mathbf{u}_1, \ldots, \mathbf{u}_n ; \mathbf{v}_1, \ldots, \mathbf{v}_n) := \mathbf{g}(\mathbf{u}_i, \mathbf{v}_j)\,, \qquad i, j = 1, \ldots, n\,.$$

The determinant of the matrix $G(\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{v}_1, \ldots, \mathbf{v}_n)$ is multilinear and skew-symmetric in each n-tuple $\{\mathbf{u}_i\} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}^n$ and $\{\mathbf{v}_j\} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}^n$.

It may then be written as the product of the corresponding values of a *metric-induced volume form* $\boldsymbol{\mu}_{\mathbf{g}} \in \mathrm{C}^1(\mathbb{M} ; BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}^n ; \Re))$ defined by

$$\det G(\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{v}_1, \ldots, \mathbf{v}_n) = \boldsymbol{\mu}_{\mathbf{g}}(\mathbf{u}_1, \ldots, \mathbf{u}_n)\, \boldsymbol{\mu}_{\mathbf{g}}(\mathbf{v}_1, \ldots, \mathbf{v}_n)\,.$$

Setting $\mathbf{u}_i = \mathbf{v}_i$ we get the relation

$$\det G(\mathbf{v}_1, \ldots, \mathbf{v}_n ; \mathbf{v}_1, \ldots, \mathbf{v}_n) = (\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{v}_1, \ldots, \mathbf{v}_n))^2\,.$$

It follows that the volume of the parallelepiped with edges $\{\mathbf{v}_i\}$, evaluated by the metric-induced volume form, is equal to ± 1 if the n-tuple $\{\mathbf{v}_i\} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}^n$ is orthonormal according to the metric. If the volume is positive the n-tuple is said to have a positive *orientation*.

We denote by $\partial\mathbb{M}$ the $(n-1)$-dimensional manifold which is the boundary of the n-dimensional $\mathbb{M}$.

The volume forms $\boldsymbol{\mu}_{\mathbb{M}}$ and $\boldsymbol{\mu}_{\partial\mathbb{M}}$ on the manifolds $\mathbb{M}$ e $\partial\mathbb{M}$ and the normal $\mathbf{n}_{\partial\mathbb{M}} \in \mathbb{T}\mathbb{M}$ to the manifold $\partial\mathbb{M}$ meet the relations

$$\begin{aligned}
&\mathbf{i}_{\mathbf{n}}\boldsymbol{\mu}_{\mathbb{M}} = \boldsymbol{\mu}_{\mathbb{M}}\mathbf{n}_{\partial\mathbb{M}} = \boldsymbol{\mu}_{\partial\mathbb{M}} \iff \\
&\mathbf{i}_{\mathbf{w}}\boldsymbol{\mu}_{\mathbb{M}} = \boldsymbol{\mu}_{\mathbb{M}}\mathbf{w} = \mathbf{g}\,(\mathbf{w}, \mathbf{n}_{\partial\mathbb{M}})\, \boldsymbol{\mu}_{\partial\mathbb{M}}\,, \quad \forall\, \mathbf{w} \in \mathbb{T}\mathbb{M}\,.
\end{aligned}$$

1.6.3 Integration of volume forms

The integral of the standard differential volume form $\boldsymbol{\mu}$ on an orientable compact n-dimensional manifold $\mathbb{M}$ provides the standard signed volume of the manifold. The integration may be performed in the model space E by means of local charts $\{U, \varphi\}$ which define a volume form $\varphi\!\uparrow\!\boldsymbol{\mu}$ on the set $\varphi(U) \subset E$. Accordingly, we define the integral by

$$\int_U \boldsymbol{\mu} := \int_{\varphi(U)} \varphi\!\uparrow\!\boldsymbol{\mu} = \int_{\varphi(U)} (\det d\varphi)^{-1}\, \boldsymbol{\mu}\,,$$

1.6.4 Partition of unity

Integrals over a compact manifold are then defined by means of the *partition of unity* method (see [3] chapter 7).

- An open covering $\{U_\alpha\}$, $\alpha \in \mathcal{A}$ of $\mathbb{M}$ is said to be *locally finite* if for each $\mathbf{x} \in \mathbb{M}$ there is a neighborhood $U(\mathbf{x})$ such that $U(\mathbf{x}) \cap U_\alpha = \emptyset$ except for finitely many indices $\alpha \in \mathcal{A}$.
- A C^k-*partition of unity* on $\mathbb{M}$ is a family $\{U_\alpha\,, f_\alpha\}$, $\alpha \in \mathcal{A}$ with $f_\alpha \in \mathrm{C}^k(U_\alpha\,; \Re)$ such that

$$\begin{aligned} &\{U_\alpha\},\ \alpha \in \mathcal{A} && \text{is a locally finite open covering of } \mathbb{M} \\ &f_\alpha(\mathbf{x}) \geq 0 && \text{and } f_\alpha(\mathbf{x}) = 0 \text{ outside a closed set included in } U_\alpha \\ &\sum_{\alpha\in\mathcal{A}} f_\alpha(\mathbf{x}) = 1 && \text{for all } \mathbf{x} \in \mathbb{M} \text{ (this is a finite sum)}\,. \end{aligned}$$

Compact manifolds admit a partition of unity $\{U_\alpha, \varphi_\alpha, f_\alpha\}$, $\alpha \in \mathcal{A}$ subordinated to an atlas, i.e. such that each element U_α of the partition is included in the domain of the chart φ_α.

We then define the integral over $\mathbb{M}$ by patching together the integrands:

$$\int_{\mathbb{M}} \boldsymbol{\mu} := \sum_{\alpha\in\mathcal{A}} \int_{U_\alpha} f_\alpha\, \boldsymbol{\mu} = \sum_{\alpha\in\mathcal{A}} \int_{\varphi_\alpha(U_\alpha)} \varphi_\alpha\!\uparrow\!(f_\alpha\, \boldsymbol{\mu})\,.$$

The integral is independent of the chosen atlas and of the subordinated partition of unity.

Indeed if $\{U_\beta, \varphi_\beta, g_\beta\}$, $\alpha \in \mathcal{B}$ is another subordinated partition of unity we have that

$$\sum_{\alpha\in\mathcal{A}} \int_{U_\alpha} f_\alpha\, \boldsymbol{\mu} = \sum_{\alpha\in\mathcal{A}} \int_{U_\alpha} \sum_{\beta\in\mathcal{B}} g_\beta\, f_\alpha\, \boldsymbol{\mu} = \sum_{\beta\in\mathcal{B}} \int_{U_\beta} g_\beta \sum_{\alpha\in\mathcal{A}} f_\alpha\, \boldsymbol{\mu} = \sum_{\beta\in\mathcal{B}} \int_{U_\beta} g_\beta\, \boldsymbol{\mu}\,.$$

since

$$\sum_{\alpha\in\mathcal{A}} f_\alpha = \sum_{\beta\in\mathcal{B}} g_\beta = \sum_{\beta\in\mathcal{B}} \sum_{\alpha\in\mathcal{A}} g_\beta\, f_\alpha = 1\,.$$

Any other differential volume form $\boldsymbol{\omega}^n$ on $\mathbb{M}$ provides a weighted volume of the manifold $\mathbb{M}$ since, setting

$$\boldsymbol{\omega}^n = w\, \boldsymbol{\mu}\,, \qquad w \in \mathrm{C}^1(\mathbb{M}\,; \Re)\,,$$

we have that

$$\int_{\mathbb{M}} \boldsymbol{\omega}^n = \int_{\mathbb{M}} w\, \boldsymbol{\mu}\,.$$

The scalar field $w \in \mathrm{C}^1(\mathbb{M}\,; \Re)$ is the weight function.

1.6.5 Chains

Let $\mathbb{M}$ be an oriented n-dimensional manifold with volume-form $\boldsymbol{\mu}$. Then a concording orientation on its $(n-1)$-dimensional boundary submanifold $\partial\mathbb{M}$ is defined by the volume-form $\boldsymbol{\mu}\mathbf{n}$ with $\mathbf{n}$ pointing outside $\mathbb{M}$.

- A *k-chain* is a family of oriented k-dimensional manifolds having $(k-1)$-dimensional boundary submanifolds in common. To each k-dimensional manifold of the chain we assign a positive or a negative sign so that, by taking the signed volume-form $\pm\boldsymbol{\mu}$ on each of them, the same orientation is induced on the common $(k-1)$-dimensional boundary submanifolds.

1.6.6 Stokes formula and exterior derivative

Let $\mathbb{M}$ be an n-dimensional chain and $\partial\mathbb{M}$ its boundary wich is an $(n-1)$-dimensional chain with the induced orientation.

Stokes formula states that the integral of a differential $(n-1)$-form $\boldsymbol{\omega}^{n-1}$ on the boundary $\partial\mathbb{M}$, is equal to the integral on $\mathbb{M}$ of a differential n-form called its exterior derivative $d\boldsymbol{\omega}^{n-1}$, i.e.

$$\int_{\mathbb{M}} d\boldsymbol{\omega}^{n-1} = \oint_{\partial\mathbb{M}} \boldsymbol{\omega}^{n-1}\,.$$

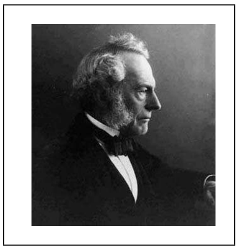

Figure 1.28: George Gabriel Stokes (1819 - 1903)

The exterior derivative is an operation on differential k-forms which is uniquely defined by the validity of STOKES formula. It is the natural extension of the fundamental theorem of calculus for functions (0-forms) to integration on compact n-dimensional chains. This celebrated formula is named the NEWTON-LEIBNIZ-GAUSS-GREEN-OSTROGRADSKY-STOKES-POINCARÉ formula by ARNOLD in [6].

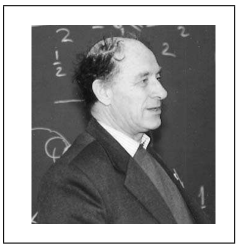

Figure 1.29: Vladimir Igorevich Arnold (1937 -)

Historical notes are reported by ERICKSEN in [46] which suggests that the classical form of STOKES theorem should be named AMPÈRE-KELVIN-HANKEL transform. The generalized version in terms of exterior derivative of forms is due to POINCARÉ.

From STOKES formula we infer the following useful result.

Proposition 1.6.1 (Exterior derivatives and pushes) *The pull back and the exterior derivative of a differential form commute, that is:*

$$d \circ \varphi\!\downarrow = \varphi\!\downarrow \circ\, d\,.$$

Proof. For any k-form $\boldsymbol{\omega}^k$ and any $(k+1)$-dimensional chain $\mathbb{N}$ we have the equality

$$\int_{\mathbb{N}} d(\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}^k) = \oint_{\partial\mathbb{N}} \boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}^k = \oint_{\boldsymbol{\varphi}(\partial\mathbb{N})} \boldsymbol{\omega}^k = \oint_{\partial\boldsymbol{\varphi}(\mathbb{N})} \boldsymbol{\omega}^k = \oint_{\boldsymbol{\varphi}(\mathbb{N})} d\boldsymbol{\omega}^k = \oint_{\mathbb{N}} \boldsymbol{\varphi}\!\downarrow\!(d\boldsymbol{\omega}^k)\,,$$

which yields the result. ■

As a direct consequence we get the following result.

Proposition 1.6.2 (Commutation of exterior and Lie derivatives) *The Lie derivative and the exterior derivative of a differential form commute:*

$$d \circ \mathcal{L}_{\mathbf{v}} = \mathcal{L}_{\mathbf{v}} \circ d\,.$$

This result may also be inferred from the homotopy formula (see section 1.6.10).

1.6.7 Cycles and boundaries, closed and exact forms

Chains and differential forms are tied by corresponding properties:

- a k-chain $\mathbb{N}^k$ is *closed* or a *cycle* if $\partial\mathbb{N}^k = 0$,
- a k-form $\boldsymbol{\omega}^k$ is *closed* or a *cocycle* if $d\boldsymbol{\omega}^k = 0$,
- a k-chain $\mathbb{N}^k$ is a *boundary* if $\mathbb{N}^k = \partial\mathbb{N}^{k-1}$,
- a k-form $\boldsymbol{\omega}^k$ is *exact* or a *coboundary* if $\boldsymbol{\omega}^k = d\boldsymbol{\omega}^{k-1}$.

Basic properties of chains and forms are the following.

- Any boundary is a cycle since: $\mathbb{N}^k = \partial\mathbb{N}^{k-1} \Longrightarrow \partial\mathbb{N}^k = \partial\partial\mathbb{N}^{k-1} = 0$,
- Any exact form is closed since: $\boldsymbol{\omega}^k = d\boldsymbol{\omega}^{k-1} \Longrightarrow d\boldsymbol{\omega}^k = dd\boldsymbol{\omega}^{k-1} = 0$.

The property $\partial\partial\mathbb{N}^{k-1} = 0$ is easily established by observing that each of its element appears twice with opposite signs. Then STOKES formula shows that

$$\partial\partial\mathbb{N}^{k+2} = 0 \implies dd\boldsymbol{\omega}^k = 0\,.$$

Indeed, if $\mathbb{N}$ is any $(k+2)$-dimensional manifold, we have that

$$\int_{\mathbb{N}^{k+2}} dd\boldsymbol{\omega}^k = \oint_{\partial\mathbb{N}^{k+2}} d\boldsymbol{\omega}^k = \oint_{\partial\partial\mathbb{N}^{k+2}} \boldsymbol{\omega}^k = 0\,.$$

On the contrary we have that, globally on a manifold:

- a cycle is not necessarily a boundary,
- a closed form (a *cocycle*) is not necessarily exact (a *coboundary*).

In this respect see the POINCARÉ lemma in secton 1.6.12 and the definition of homology and cohomology classes, DE RHAM theorem and BETTI's numbers, in section 1.6.14.

1.6.8 Transport theorem

Let $\Gamma \subset \mathbb{M}$ be a compact k-dimensional submanifold embedded in a n-dimensional manifold $\mathbb{M}$ with $k < n$ and $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M} \times I\,;\mathbb{M})$ be a flow whose velocity field is denoted by $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$.

The flow drags the submanifold $\Gamma \subset \mathbb{M}$ and the dragged submanifold $\boldsymbol{\varphi}_t(\Gamma) \subset \mathbb{M}$ traces in the interval $t \in [0,1]$ a $(k+1)$-dimensional submanifold $J_{\mathbf{v}}(\Gamma) \subseteq \mathbb{M}$ (a *flow tube*) given by

$$J_{\mathbf{v}}(\Gamma) := \bigcup_{\substack{\mathbf{x}\in\Gamma \\ t\in[0,1]}} \boldsymbol{\varphi}_t(\mathbf{x})\,.$$

The smooth transformation from $\boldsymbol{\varphi}_0(\Gamma) = \Gamma \subset \mathbb{M}$ to $\boldsymbol{\varphi}_1(\Gamma) \subset \mathbb{M}$ in the interval $t \in [0,1]$ is called an *homotopy*.

Proposition 1.6.3 (Transport theorem) *For any time-dependent differential k-form $\boldsymbol{\omega}^k_t$ on $J_{\mathbf{v}}(\Gamma)$ we have that*

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,s}(\Gamma)} \boldsymbol{\omega}^k_\tau = \int_{\boldsymbol{\varphi}_{t,s}(\Gamma)} \mathcal{L}_{t,\mathbf{v}}\, \boldsymbol{\omega}^k_t = \int_{\boldsymbol{\varphi}_{t,s}(\Gamma)} \partial_{\tau=t}\, \boldsymbol{\omega}^k_\tau + (\mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^k)_t\,.$$

Proof. Being $\boldsymbol{\varphi}_{t,s} = \boldsymbol{\varphi}_{t,\tau} \circ \boldsymbol{\varphi}_{\tau,s}$, by the formula of transformation of integrals under a diffeomorphism:

$$\int_\Gamma \boldsymbol{\omega}^k_\tau = \int_{\boldsymbol{\varphi}_{t,\tau}(\Gamma)} \boldsymbol{\varphi}_{t,\tau}\!\uparrow\!\boldsymbol{\omega}^k_\tau\,,$$

we have that

$$\int_{\boldsymbol{\varphi}_{\tau,s}(\Gamma)} \boldsymbol{\omega}^k_\tau = \int_{\boldsymbol{\varphi}_{t,s}(\Gamma)} \boldsymbol{\varphi}_{t,\tau}\!\uparrow\!\boldsymbol{\omega}^k_\tau\,.$$

Differentiating with respect to time $\tau \in I$ at $\tau = t$, the result follows from the definition of the LIE derivative. ∎

1.6.9 Fubini's theorem for differential forms

Let us consider a differential volume-form $\boldsymbol{\omega}^{k+1}$ on the $(k+1)$-dimensional manifold $J_{\mathbf{v}}(\Gamma)$.

A corresponding volume-form is then induced on each $\boldsymbol{\varphi}_t(\Gamma)$.

It is given by the contraction $\mathbf{i}_{\mathbf{n}}\boldsymbol{\omega}^{k+1} = \boldsymbol{\omega}^{k+1}\mathbf{n}$ of the volume-form $\boldsymbol{\omega}^{k+1}$ with the unit normal vector $\mathbf{n} \in \mathbb{T}_{J_{\mathbf{v}}(\Gamma)}$ to the manifold $\boldsymbol{\varphi}_t(\Gamma)$, regarded as a submanifold of the flow tube $J_{\mathbf{v}}(\Gamma)$.

We shall often write the integral of a volume-form $\boldsymbol{\alpha}$ on the manifold Γ in the contracted form $\boldsymbol{\alpha}\Gamma$, following the notation in [36].

Figure 1.30: Guido Fubini (1879 - 1943)

Fubini's theorem states that the volume of the $(k+1)$-dimensional flow tube $J_{\mathbf{v}}(\Gamma)$, evaluated according to a differential volume-form $\boldsymbol{\omega}^{k+1}$ on $J_{\mathbf{v}}(\Gamma)$, is equal to the integral, along the homotopic flow, of the corresponding flux of the velocity field $\mathbf{v}$ of the flow through the flowing k-dimensional manifold $\boldsymbol{\varphi}_t(\Gamma)$:

$$\int_{J_{\mathbf{v}}(\Gamma)} \boldsymbol{\omega}^{k+1} = \int_0^1 dt \int_{\boldsymbol{\varphi}_t(\Gamma)} (\boldsymbol{\omega}^{k+1}\mathbf{v}) \,.$$

If the flow tube $J_{\mathbf{v}}(\Gamma)$ is endowed with a rimannian metric, the velocity field may be decomposed into a normal and a parallel component according to the formula

$$\mathbf{v} = v_{\mathbf{n}}\,\mathbf{n} + \mathbf{v}^{||} \,, \quad v_{\mathbf{n}} = \mathbf{g}(\mathbf{v},\mathbf{n}) \,,$$

Hence, noting that $(\boldsymbol{\omega}^{k+1}\mathbf{v}^{||})\boldsymbol{\varphi}_t(\Gamma) = 0$, Fubini's formula takes the expression

$$\int_{J_{\mathbf{v}}(\Gamma)} \boldsymbol{\omega}^{k+1} = \int_0^1 dt \int_{\boldsymbol{\varphi}_t(\Gamma)} \mathbf{g}(\mathbf{v},\mathbf{n})\,(\boldsymbol{\omega}^{k+1}\mathbf{n}) \,.$$

In terms of time rates, FUBINI's theorem states that

- the rate of variation of the volume, evaluated according to a volume-form $\boldsymbol{\omega}^{k+1}$, of the $k+1$-dimensional flow tube $J_{\mathbf{v}}(\Gamma, t)$, traced by a k-dimensional submanifold $\boldsymbol{\varphi}_\tau(\Gamma)$, with $\tau \in [0, t]$, is equal to the flux of the velocity field $\mathbf{v}$ of the flow through the tracing manifold $\boldsymbol{\varphi}_t(\Gamma)$:

$$\partial_{\tau=t} \int_{J_{\mathbf{v}}(\Gamma,\tau)} \boldsymbol{\omega}^{k+1} = \int_{\boldsymbol{\varphi}_t(\Gamma)} (\boldsymbol{\omega}^{k+1}\mathbf{v}) = \int_{\boldsymbol{\varphi}_t(\Gamma)} \mathbf{g}(\mathbf{v}, \mathbf{n})\, (\boldsymbol{\omega}^{k+1}\mathbf{n})\,.$$

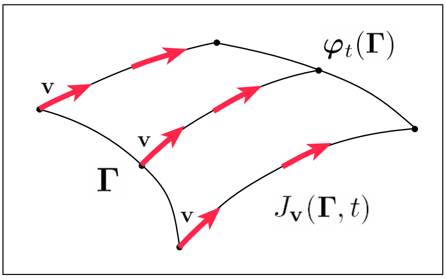

Figure 1.31: $k+1$-dimensional flow tube $J_{\mathbf{v}}(\Gamma, t)$

1.6.10 Homotopy formula

The boundary of the $(k+1)$-dimensional flow tube $J_{\mathbf{v}}(\Gamma, t)$ traced in the interval $[0, t]$ by the k-dimensional submanifold $\boldsymbol{\varphi}_\tau(\Gamma)$ $(\tau \in [0, t])$, flowing in $\mathbb{M}$ according to an orientation preserving flow $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$, is the k-chain given by the

- *geometric homotopy formula*

$$\partial(J_{\mathbf{v}}(\Gamma, t)) = \boldsymbol{\varphi}_t(\Gamma) - \Gamma - J_{\mathbf{v}}(\partial\Gamma, t)\,.$$

The signs in the formula are due to the following choice.

The orientation of the $(k+1)$-dimensional flow tube $J_{\mathbf{v}}(\Gamma, t)$ induces an orientation on its boundary $\partial(J_{\mathbf{v}}(\Gamma, t))$. Assuming on $\boldsymbol{\varphi}_t(\Gamma)$ this orientation, it follows that $\boldsymbol{\varphi}_0(\Gamma) = \Gamma$ has the opposite orientation and the same holds for $J_{\mathbf{v}}(\partial\Gamma, t)$.

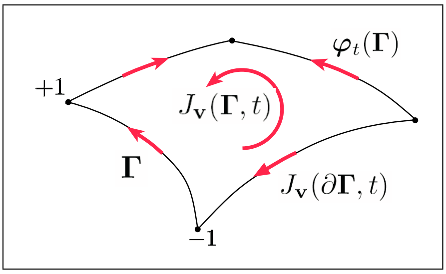

Figure 1.32: Geometric homotopy formula

For any time-independent differential form $\boldsymbol{\omega}^k$ on $J_{\mathbf{v}}(\Gamma, t)$, the geometric homotopy formula yields the additive decomposition:

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_\tau(\Gamma)} \boldsymbol{\omega}^k = \partial_{\tau=t} \int_{\partial(J_{\mathbf{v}}(\Gamma,\tau))} \boldsymbol{\omega}^k + \partial_{\tau=t} \int_{J_{\mathbf{v}}(\partial\Gamma,\tau)} \boldsymbol{\omega}^k + \partial_{\tau=t} \int_{\Gamma} \boldsymbol{\omega}^k \,,$$

where the last term on the r.h.s. vanishes, being the integral time-independent.

Applying Stokes and Fubini's formulas to the first term on the r.h.s., we get

$$\partial_{\tau=t} \int_{\partial(J_{\mathbf{v}}(\Gamma,\tau))} \boldsymbol{\omega}^k = \partial_{\tau=t} \int_{J_{\mathbf{v}}(\Gamma,\tau)} d\boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}_t(\Gamma)} (d\boldsymbol{\omega}^k)\mathbf{v} \,.$$

Hence applying Fubini's formula to the second term on the r.h.s., noting that $\boldsymbol{\varphi}_t(\partial\Gamma) = \partial(\boldsymbol{\varphi}_t(\Gamma))$, by Stokes formula we have that

$$\partial_{\tau=t} \int_{J_{\mathbf{v}}(\partial\Gamma,\tau)} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}_t(\partial\Gamma)} (\boldsymbol{\omega}^k\mathbf{v}) = \int_{\partial(\boldsymbol{\varphi}_t(\Gamma))} (\boldsymbol{\omega}^k\mathbf{v}) = \int_{\boldsymbol{\varphi}_t(\Gamma)} d(\boldsymbol{\omega}^k\mathbf{v}) \,.$$

Summing up we get the *extrusion formula*:

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_\tau(\Gamma)} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}_t(\Gamma)} (d\boldsymbol{\omega}^k)\mathbf{v} + \int_{\boldsymbol{\varphi}_t(\Gamma)} d(\boldsymbol{\omega}^k\mathbf{v}) \,.$$

On the other hand, Reynolds transport formula tells us that

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_\tau(\Gamma)} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}_t(\Gamma)} \mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^k \,.$$

Comparing the two formulas, we get

$$\int_{\boldsymbol{\varphi}_t(\Gamma)} \mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}_t(\Gamma)} (d\boldsymbol{\omega}^k)\mathbf{v} + \int_{\boldsymbol{\varphi}_t(\Gamma)} d(\boldsymbol{\omega}^k\mathbf{v}) \,.$$

By the arbitrariness of the k-dimensional submanifold $\Gamma \subset \mathbb{M}$, the extrusion formula may be localized to get the *differential homotopy formula*

$$\mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^k = (d\boldsymbol{\omega}^k)\mathbf{v} + d(\boldsymbol{\omega}^k \mathbf{v})\,,$$

also known as HENRI CARTAN *'s magic formula*, [106], [137] that provides a basic relation between the LIE and the exterior derivative of a differential form [25].

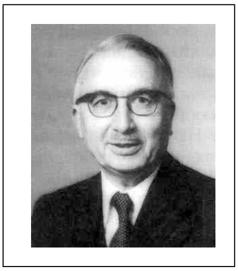

Figure 1.33: Henri Paul Cartan (1904 - 2008)

1.6.11 Palais' formula

The homotopy formula for one-forms may be readily inverted to provide PALAIS formula for the exterior derivative of one-forms. Indeed, by LEIBNIZ rule for the LIE derivative, we have that for any two vector fields $\mathbf{v}, \mathbf{w} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TM})$:

$$\begin{aligned} d\boldsymbol{\omega}^1 \cdot \mathbf{v} \cdot \mathbf{w} &= (\mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^1) \cdot \mathbf{w} - d(\boldsymbol{\omega}^1 \cdot \mathbf{v}) \cdot \mathbf{w} \\ &= d_{\mathbf{v}}\, (\boldsymbol{\omega}^1 \cdot \mathbf{w}) - d_{\mathbf{w}}\, (\boldsymbol{\omega}^1 \cdot \mathbf{v}) - \boldsymbol{\omega}^1 \cdot [\mathbf{v}\,, \mathbf{w}]\,. \end{aligned}$$

By tensoriality, the point value $(d\boldsymbol{\omega}^1 \cdot \mathbf{v} \cdot \mathbf{w})(\mathbf{x})$ at $\mathbf{x} \in \mathbb{M}$ depends only on the point values $\mathbf{v}_{\mathbf{x}}, \mathbf{w}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ and not on the knowledge of the vector fields $\mathbf{v}, \mathbf{w} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TM})$ in a neighbourhoor of $\mathbf{x} \in \mathbb{M}$.

Anyway the evaluation of the r.h.s. requires to extend these vectors to vector fields $\mathbf{v}, \mathbf{w} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TM})$, but the result is independent of the extension.

The same algebra may be applied repeatedly to deduce PALAIS formula for

Figure 1.34: Richard Palais (1931 -) with his wife and frequent co-author, Chuu-lianx Terng at the dedication of a memorial bust of Sophus Lie, at Lie's birthplace in Nordfjord, Norway.

the exterior derivative of a k-form [138]:

$$d\boldsymbol{\omega}^k(\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k) := \sum_{i=0,k} (-1)^i\, \mathbf{v}_i\, (\omega^k(\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k)_i) + \sum_{\substack{i,j=0,k \\ i<j}} (-1)^{i+j}\, (\omega^k([\mathbf{v}_i\,, \mathbf{v}_j], \mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_k)_{i,j})\,,$$

where the subscript $()_i$ means that the i-th term in the parenthesis is missing and the subscript $()_{i,j}$ means that the i-th and j-th terms are missing.

The tensoriality of the exterior derivative follows from the criterion of Lemma 1.2.1 by invoking the LEIBNIZ formula for the LIE derivative (formula *iii* of Proposition 1.3.11):

$$[\mathbf{v}_i\,, f\mathbf{v}_j] = f[\mathbf{v}_i\,, \mathbf{v}_j] + (\mathbf{v}_i f)\, \mathbf{v}_j\,.$$

Indeed we have that

$$(-1)^i \mathbf{v}_i\, (\omega^k(\dots, f\mathbf{v}_j, \dots)_i) = (-1)^i f \mathbf{v}_i\, (\omega^k(\dots, \mathbf{v}_j, \dots)_i) + (-1)^i (\mathbf{v}_i f)\, (\omega^k(\dots, \mathbf{v}_j, \dots)_i)\,,$$

$$\begin{aligned}(-1)^{i+j}\, \omega^k([\mathbf{v}_i\,, f\mathbf{v}_j], \dots)_{i,j} &= (-1)^{i+j}\, f\, \omega^k([\mathbf{v}_i\,, \mathbf{v}_j], \dots)_{i,j} \\ &\quad + (-1)^{i+j}\, (\mathbf{v}_i f)\, (\omega^k(\mathbf{v}_j, \dots)_i) \\ &= (-1)^{i+j}\, f\, \omega^k([\mathbf{v}_i\,, \mathbf{v}_j], \dots)_{i,j} \\ &\quad - (-1)^i\, (\mathbf{v}_i f)\, (\omega^k(\dots, \mathbf{v}_j, \dots)_i)\,.\end{aligned}$$

By tensoriality, the argument vectors may be extended to vector fields in an arbitrary way.

If the associated flows commute pairwise, so that $[\mathbf{v}_i\,,\mathbf{v}_j]=0$ for $i,j=0,\dots,k$, PALAIS' formula for the exterior derivative of a k-form $\boldsymbol{\omega}^k$ reduces to:

$$d\boldsymbol{\omega}^k(\mathbf{v}_0,\mathbf{v}_1,\dots,\mathbf{v}_k)=\sum_{i=0,k}(-1)^i\,\mathbf{v}_i\,(\omega^k(\mathbf{v}_0,\mathbf{v}_1,\dots,\mathbf{v}_k)_i)\,.$$

The exterior derivative of the exterior product of two differential forms is given by the formula

$$d(\boldsymbol{\alpha}^p\wedge\boldsymbol{\omega}^k)=(d\boldsymbol{\alpha}^p)\wedge\boldsymbol{\omega}^k+(-1)^p\,\boldsymbol{\alpha}^p\wedge d\boldsymbol{\omega}^k\,.$$

Hence the exterior derivative is an *anti-derivation* for the exterior algebra, i.e. a graded derivation of degree $+1$ (see section 1.7). Let $\{x^1,\dots,x^n\}$ be a local coordinate system on $\mathbb{M}$ with local basis $\{\partial x_1,\dots,\partial x_n\}$ and dual local basis $\{dx^1,\dots,dx^n\}$ so that $\langle dx^i,\partial x_j\rangle=\delta^i_j$.

A k-form $\boldsymbol{\omega}^k\in\Lambda^k_{\mathbf{x}}(\mathbb{M}\,;\Re)$ may be written as a linear combination of k-fold exterior products of the differentials of the coordinates:

$$\boldsymbol{\omega}^k=\omega_{i_1,\dots,i_k}\,dx^{i_1}\wedge\dots\wedge dx^{i_k}\,,$$

where the components are given by $\omega_{i_1,\dots,i_k}=\boldsymbol{\omega}(\partial x_{i_1},\dots,\partial x_{i_k})$ and the sum is performed over the set of indices $1\le i_1<\dots<i_k\le n$.

Accordingly, the expression of the exterior derivative $d\boldsymbol{\omega}^k$ in terms of components is given by

$$d\boldsymbol{\omega}^k=d\omega_{i_1,\dots,i_k}\wedge dx^{i_1}\wedge\dots\wedge dx^{i_k}\,.$$

An alternative formula for the exterior derivative in terms of components is deduced from PALAIS formula taking into account that the LIE bracket of any pair of coordinate vector fields vanish i.e. $[\partial_i\,,\partial_j]=0$ for $i,j=0,\dots,k$, so that:

$$d\boldsymbol{\omega}^k(\partial x_0,\partial x_1,\dots,\partial x_k)=\sum_{i=0,k}(-1)^i\,\partial_i\,(\omega^k(\partial x_0,\partial x_1,\dots,\partial x_k)_i)\,.$$

When acting on exterior forms, the contraction $\mathbf{i}_{\mathbf{v}}$ and the exterior derivative are operators with a null iterate:

$$d\circ d=0\,,$$

$$\mathbf{i}_{\mathbf{v}}\circ\mathbf{i}_{\mathbf{v}}=0\,.$$

The homotopy formula then yields the commutativity properties:

$$\begin{aligned} \mathcal{L}_{\mathbf{v}} \circ d &= d \circ \mathcal{L}_{\mathbf{v}} \,, \\ \mathcal{L}_{\mathbf{v}} \circ \mathbf{i}_{\mathbf{v}} &= \mathbf{i}_{\mathbf{v}} \circ \mathcal{L}_{\mathbf{v}} \,, \end{aligned}$$

and the equality

$$\mathcal{L}_{\mathbf{v}} \circ \mathbf{i}_{\mathbf{u}} - \mathbf{i}_{\mathbf{v}} \circ \mathcal{L}_{\mathbf{u}} = d \circ \mathbf{i}_{\mathbf{v}} \circ \mathbf{i}_{\mathbf{u}} - \mathbf{i}_{\mathbf{v}} \circ \mathbf{i}_{\mathbf{u}} \circ d \,.$$

The homotopy formula provides a simpler proof of property $v)$ of the Lie derivative, provided in Proposition 1.3.7 for general tensors, in the special case of the Lie derivative of a k-form:

$$\begin{aligned} \mathcal{L}_{(f\,\mathbf{v})}\, \boldsymbol{\omega}^k &= d(f\, \boldsymbol{\omega}^k\, \mathbf{v}) + f\,(d\boldsymbol{\omega}^k)\, \mathbf{v} = d(f\, \boldsymbol{\omega}^k\, \mathbf{v}) + f\,(\mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^k - d(\boldsymbol{\omega}^k\, \mathbf{v})) \\ &= df \wedge (\boldsymbol{\omega}^k\, \mathbf{v}) + f\, d(\boldsymbol{\omega}^k\, \mathbf{v}) + f\,(\mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^k - d(\boldsymbol{\omega}^k\, \mathbf{v})) \\ &= df \wedge (\boldsymbol{\omega}^k\, \mathbf{v}) + f\, \mathcal{L}_{\mathbf{v}}\, \boldsymbol{\omega}^k \,. \end{aligned}$$

Moreover for volume forms $\boldsymbol{\mu}$ we get a simple proof of property $vi)$ in Proposition 1.3.7:

$$\begin{aligned} \mathcal{L}_{(f\,\mathbf{v})}\, \boldsymbol{\mu} &= d(\boldsymbol{\mu}\, f\, \mathbf{v}) + (d\boldsymbol{\mu})\, f\, \mathbf{v} = d(f\, \boldsymbol{\mu}\, \mathbf{v}) \\ &= d(f\, \boldsymbol{\mu}\, \mathbf{v}) + d(f\, \boldsymbol{\mu})\, \mathbf{v} = \mathcal{L}_{\mathbf{v}}\,(f\, \boldsymbol{\mu}) \,. \end{aligned}$$

1.6.12 Poincaré lemma

Figure 1.35: Jules Henri Poincaré (1854 - 1912)

Let us now give the following definition:

- An n-dimensional manifold $\mathbb{M}$ is a *star-shaped manifold* if there exists a point $\mathbf{x}_0 \in \mathbb{M}$ and an homotopy $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$, continuous in $t \in [0,1]$, such that $\boldsymbol{\varphi}_1$ is the identity map, i.e. $\boldsymbol{\varphi}_1(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{M}$, and $\boldsymbol{\varphi}_0$ is the constant map $\boldsymbol{\varphi}_0(\mathbf{x}) = \mathbf{x}_0$ for all $\mathbf{x} \in \mathbb{M}$. This homotopy is called a *contraction* to $\mathbf{x}_0 \in \mathbb{M}$.

Denoting by $\mathbf{v} = \partial_{t=0}\,\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ the velocity of the flow, by Proposition 1.3.1, p. 63, we have that, for any differential k-form ($k \le n$) $\boldsymbol{\omega}^k \in \mathrm{C}^1(\mathbb{M}\,;\Lambda^k)$:

$$\partial_{\tau=t}\,(\boldsymbol{\varphi}_\tau\!\downarrow\!\boldsymbol{\omega}^k) = \boldsymbol{\varphi}_t\!\downarrow\!(\mathcal{L}_{\mathbf{v}}\boldsymbol{\omega}^k)\,.$$

Integrating along the flow $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ in the interval $t \in [0,1]$, we get

$$\boldsymbol{\varphi}_1\!\downarrow\!\boldsymbol{\omega}^k - \boldsymbol{\varphi}_0\!\downarrow\!\boldsymbol{\omega}^k = \int_0^1 \boldsymbol{\varphi}_t\!\downarrow\!(\mathcal{L}_{\mathbf{v}}\boldsymbol{\omega}^k)\,dt\,.$$

By the homotopy formula and the property $\boldsymbol{\varphi}_t\!\downarrow \circ\, d = d \circ \boldsymbol{\varphi}_t\!\downarrow$ (see Proposition 1.6.1), we infer that

$$\boldsymbol{\varphi}_1\!\downarrow\!\boldsymbol{\omega}^k - \boldsymbol{\varphi}_0\!\downarrow\!\boldsymbol{\omega}^k = d\int_0^1 \boldsymbol{\varphi}_t\!\downarrow\!(\boldsymbol{\omega}^k \cdot \mathbf{v})\,dt + \int_0^1 \boldsymbol{\varphi}_t\!\downarrow\!(d\boldsymbol{\omega}^k \cdot \mathbf{v})\,dt\,.$$

Recalling that $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ is a contraction to $\mathbf{x}_0 \in \mathbb{M}$, we have that

$$\boldsymbol{\varphi}_1\!\uparrow\!\mathbf{w} = \mathbf{w}\,, \quad \boldsymbol{\varphi}_0\!\uparrow\!\mathbf{w} = 0\,, \quad \forall\,\mathbf{w} \in \mathbb{TM}\,,$$

and hence $\boldsymbol{\varphi}_1\!\downarrow\!\boldsymbol{\omega}^k = \boldsymbol{\omega}^k\,, \quad \boldsymbol{\varphi}_0\!\downarrow\!\boldsymbol{\omega}^k = 0\,.$

We have thus proved the formula

$$\boldsymbol{\omega}^k = d\boldsymbol{\alpha}^{(k-1)} + \boldsymbol{\beta}^k\,,$$

with

$$\boldsymbol{\alpha}^{(k-1)} = \int_0^1 \boldsymbol{\varphi}_t\!\downarrow\!(\boldsymbol{\omega}^k \cdot \mathbf{v})\,dt\,, \qquad \boldsymbol{\beta}^k = \int_0^1 \boldsymbol{\varphi}_t\!\downarrow\!(d\boldsymbol{\omega}^k \cdot \mathbf{v})\,dt\,,$$

If $d\boldsymbol{\omega}^k = 0$ the form $\boldsymbol{\omega}^k$ is exact and we get:

Lemma 1.6.1 (Poincare lemma) *In a star-shaped manifold any closed form is exact.*

1.6.13 Potentials in a linear space

If the manifold is a linear space $\mathbb{S}$ we may set $\boldsymbol{\varphi}_t(\mathbf{x}) = t\mathbf{x}$ so that $\mathbf{v}(\mathbf{x}) = \mathbf{x}$ and $T\boldsymbol{\varphi}_t(\mathbf{x}) = t\,\mathbf{I}$. Then we get the following expressions for $\boldsymbol{\alpha}^{(k-1)}$ and $\boldsymbol{\beta}^k$:

$$\boldsymbol{\alpha}^{(k-1)}(\mathbf{x}) = \int_0^1 t^{(k-1)}\boldsymbol{\omega}^k(t\mathbf{x}) \cdot \mathbf{x}\, dt\,,$$

$$\boldsymbol{\beta}^k(\mathbf{x}) = \int_0^1 t^k d\boldsymbol{\omega}^k(t\mathbf{x}) \cdot \mathbf{x}\, dt\,.$$

From the formula $\boldsymbol{\omega}^k = d\boldsymbol{\alpha}^{(k-1)} + \boldsymbol{\beta}^k$ we may directly infer some classical integrability conditions in a linear space and the explicit expressions of the relevant potentials.

To this end, let us recall the definitions of cross product, gradient, curl and divergence in an inner product linear space $\{\mathbb{S}\,,\mathbf{g}\}$:

cross product:	$\mathbf{u} \times \mathbf{v} = \boldsymbol{\mu}_{\mathbf{g}}\mathbf{u}\mathbf{v}\,,$	$\dim \mathbb{S} = 2$
cross product:	$\mathbf{g}(\mathbf{u} \times \mathbf{v}) = \boldsymbol{\mu}_{\mathbf{g}}\mathbf{u}\mathbf{v}\,,$	$\dim \mathbb{S} = 3$
gradient:	$d\,f = \mathbf{g}\nabla f\,,$	$\dim \mathbb{S} = n$
curl:	$d(\mathbf{g}\mathbf{v}) = (\mathrm{rot}\,\mathbf{v})\,\boldsymbol{\mu}_{\mathbf{g}}\,,$	$\dim \mathbb{S} = 2$
curl:	$d(\mathbf{g}\mathbf{v}) = \boldsymbol{\mu}_{\mathbf{g}}(\mathrm{rot}\,\mathbf{v})\,,$	$\dim \mathbb{S} = 3$
divergence:	$d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = (\mathrm{div}\,\mathbf{v})\boldsymbol{\mu}_{\mathbf{g}}\,.$	$\dim \mathbb{S} = n$

Remark 1.6.1 *An important special result is peculiar to* $\dim \mathbb{S} = 2$ *. Indeed we may set* $\boldsymbol{\mu}_{\mathbf{g}} = \mathbf{g}\mathbf{R}$ *to define uniquely the operator* $\mathbf{R} \in BL(\mathbb{S}\,;\mathbb{S})$ *. Then, by the skew-symmetry of* $\boldsymbol{\mu}_{\mathbf{g}}$ *we get* $\mathbf{R}^T = -\mathbf{R}$ *and* $\mathbf{g}(\mathbf{R}\mathbf{a},\mathbf{a}) = 0$ *for all* $\mathbf{a} \in \mathbb{S}$ *. Moreover, being* $\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{a},\mathbf{R}\mathbf{a})^2 = \mathbf{g}(\mathbf{a},\mathbf{a})\mathbf{g}(\mathbf{R}\mathbf{a},\mathbf{R}\mathbf{a})$ *, we infer that* $\mathbf{g}(\mathbf{R}\mathbf{a},\mathbf{R}\mathbf{a}) = \boldsymbol{\mu}_{\mathbf{g}}(\mathbf{a},\mathbf{R}\mathbf{a}) = \mathbf{g}(\mathbf{a},\mathbf{a})$ *so that, by polarization,* $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ *and* $\mathbf{R}^2 = -\mathbf{I}$ *. The operator* $\mathbf{R} \in BL(\mathbb{S}\,;\mathbb{S})$ *is an isometry which changes any vector in* $\mathbb{S}$ *into its orthogonal such that the oriented square* $\{\mathbf{a}\,,\mathbf{R}\mathbf{a}\}$ *has a positive area, and we have that:*

$$(\mathrm{div}\,\mathbf{v})\boldsymbol{\mu}_{\mathbf{g}} = d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = d(\mathbf{g}\mathbf{R}\mathbf{v}) = \mathrm{rot}\,(\mathbf{R}\mathbf{v})\boldsymbol{\mu}_{\mathbf{g}}\,.$$

that is $\mathrm{div}\,\mathbf{v} = \mathrm{rot}\,(\mathbf{R}\mathbf{v})$ *and* $\mathrm{div}\,(\mathbf{R}\mathbf{v}) = -\mathrm{rot}\,\mathbf{v}$ *.*

From POINCARÉ lemma we get the following results.

- Let $\dim \mathbb{S} = 3$ and $\boldsymbol{\omega}^1 = \mathbf{g}\mathbf{v}$ (analogous result for $\dim \mathbb{S} = 2$). Since $d(\mathbf{g}\mathbf{v}) = \boldsymbol{\mu}_{\mathbf{g}}(\mathrm{rot}\,\mathbf{v})$, the closedness of $\boldsymbol{\omega}^1$ is equivalent to the irrotationality condition, i.e. $d(\mathbf{g}\mathbf{v}) = 0 \iff \mathrm{rot}\,\mathbf{v} = 0$. Any irrotational vector field admits a scalar potential such that $\nabla f = \mathbf{v}$, given by

$$f(\mathbf{x}) = \int_0^1 \mathbf{g}(\mathbf{v}(t\mathbf{x}), \mathbf{x})\, dt$$

- Let $\dim \mathbb{S} = 2$ and $\boldsymbol{\omega}^1 = \boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}$. Since $d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = (\mathrm{div}\,\mathbf{v})\boldsymbol{\mu}_{\mathbf{g}}$, the closedness of $\boldsymbol{\omega}^1$ is equivalent to the solenoidality condition, i.e. $d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = 0 \iff \mathrm{div}\,\mathbf{v} = \mathrm{rot}\,(\mathbf{R}\mathbf{v}) = 0$. Then there exists a scalar potential such that $\nabla f = \mathbf{R}\mathbf{v}$, defined by

$$f(\mathbf{x}) = \int_0^1 \mathbf{g}(\mathbf{R}\mathbf{v}(t\mathbf{x}), \mathbf{x})\, dt = \int_0^1 \boldsymbol{\mu}_{\mathbf{g}}(\mathbf{v}(t\mathbf{x}), \mathbf{x})\, dt = \int_0^1 \mathbf{v}(t\mathbf{x}) \times \mathbf{x}\, dt\,.$$

- Let $\dim \mathbb{S} = 3$ and $\boldsymbol{\omega}^2 = \boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}$. Since $d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = (\mathrm{div}\,\mathbf{v})\boldsymbol{\mu}_{\mathbf{g}}$, the closedness of $\boldsymbol{\omega}^2$ is equivalent to the solenoidality condition, i.e. $d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = 0 \iff \mathrm{div}\,\mathbf{v} = 0$. Any solenoidal vector field admits then a vector potential, i.e. $\mathrm{div}\,\mathbf{v} = 0 \implies \mathbf{v} = \mathrm{rot}\,\mathbf{w}$, with

$$\mathbf{w}(\mathbf{x}) = \int_0^1 t\,\mathbf{v}(t\mathbf{x}) \times \mathbf{x}\, dt\,.$$

- Let $\dim \mathbb{S} = n$. Setting $\boldsymbol{\omega}^n = f\,\boldsymbol{\mu}_{\mathbf{g}}$ we have that $d\boldsymbol{\omega}^n = 0$ and hence any scalar field $f \in \mathrm{C}^1(\mathbb{S}\,;\Re)$ is the divergence of a vector field, that is: $f = \mathrm{div}\,\mathbf{w}$, with

$$\mathbf{w}(\mathbf{x}) = \int_0^1 t^{(n-1)}\, f(t\mathbf{x})\mathbf{x}\, dt\,.$$

1.6.14 de Rham cohomology and Betti's numbers

Let $\mathbb{M}$ be a finite dimensional, compact manifold with $n = \dim \mathbb{M}$. Then:

- Two k-chains are said to be *homological* if their difference is a boundary. The family of equivalence classes of k-chains so defined, endowed with the natural linear operations, is the *homology space* of dimension k and is denoted by $H_k(\mathbb{M})$.

- Two k-forms are said to be *cohomological* if their difference is a coboundary. The family of equivalence classes of k-forms so defined, endowed with the natural linear operations, is the *cohomology space* of dimension k and is denoted by $H^k(\mathbb{M})$.

- The dimension b_k of the linear space $H_k(\mathbb{M})$ is called the k-dimensional BETTI*'s number* of $\mathbb{M}$, [6].

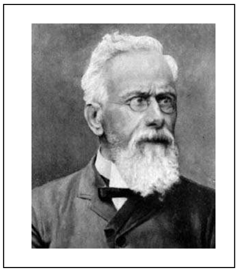

Figure 1.36: Enrico Betti (1823 - 1892)

- The EULER-POINCARÉ *characteristic* of $\mathbb{M}$ is the integer defined by:

$$\chi(\mathbb{M}) = \sum_{k=0}^{n} (-1)^k b_k \,,$$

The following result is due to S. CHERN [26], see also A. AVEZ [8].

Theorem 1.6.1 (Chern's theorem) *A finite dimensional, orientable and compact manifold $\mathbb{M}$ admits a regular vector field if and only if its* EULER-POINCARÉ *characteristic vanishes.*

We owe to GEORGES DE RHAM the following basic result [34].

Theorem 1.6.2 (de Rham's theorem) *A k-cocycle is a coboundary iff its integral over every k-cycle vanishes, and a k-cycle is a boundary iff the integral over it of every k-cocycle vanishes. The dimensions of the linear spaces $H^k(\mathbb{M})$ and $H_k(\mathbb{M})$ are the same and $b_k = b_{n-k}$ where $n = \dim \mathbb{M}$.*

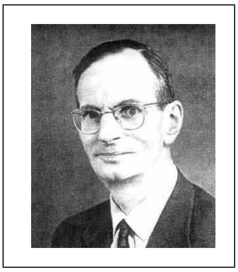

Figure 1.37: Georges de Rham (1903 - 1990)

A simple interpretation of DE RHAM's theorem may be given by rewriting STOKES formula as

$$\langle \boldsymbol{\omega}, \partial \mathbb{M} \rangle = \langle d\boldsymbol{\omega}, \mathbb{M} \rangle \,,$$

where $\dim \mathbb{M} = n$ and $\boldsymbol{\omega} \in \Lambda^{(n-1)}(\mathbb{M}\,;\Re)$. STOKES formula provides then a duality product between differential forms and manifolds with the operators ∂ (*boundary chain of*) and d (*exterior derivative of*) in duality.

The exterior derivative d is a linear operator in each linear space $\Lambda^k(\mathbb{M}\,;\Re)$ of differential k-forms on compact manifolds.

The *boundary chain* ∂ is a signed additive operator on each oriented chain of k-manifolds and positive homogeneity is granted by setting $\partial(\alpha\mathbb{M}) := \alpha(\partial\mathbb{M})$ for all $\alpha \in \Re$. Here $\alpha(\partial\mathbb{M})$ is the chain such that $\langle \boldsymbol{\omega}, \alpha(\partial\mathbb{M}) \rangle = \alpha \langle \boldsymbol{\omega}, \partial\mathbb{M} \rangle$.

The duality expressed by STOKES formula implies that kernels and ranges of the dual linear operators d and ∂ meet the properties:

$$\begin{cases} Ker\,\partial = (Im\,d)^0 \,, \\ Ker\,d = (Im\,\partial)^0 \,, \end{cases}$$

where the symbol $()^0$ denotes the annihilator according to the duality. Indeed from STOKES formula we infer that

$$\begin{cases} \boldsymbol{\omega} = d\boldsymbol{\alpha} \,, \\ \partial\Sigma = 0 \,, \end{cases} \implies \int_\Sigma \boldsymbol{\omega} = \int_\Sigma d\boldsymbol{\alpha} = \oint_{\partial\Sigma} \boldsymbol{\alpha} = 0 \,,$$

and that

$$\begin{cases} \Sigma = \partial\mathbb{M} \,, \\ d\boldsymbol{\omega} = 0 \,, \end{cases} \implies \int_\Sigma \boldsymbol{\omega} = \int_{\partial\mathbb{M}} \boldsymbol{\omega} = \oint_{\mathbb{M}} d\boldsymbol{\omega} = 0 \,,$$

which are the implications to be proved. ■

Then DE RHAM's theorem states that we have also the relations

$$\begin{cases} Im\,\partial = (Ker\,d)^0\,, \\ Im\,d = (Ker\,\partial)^0\,, \end{cases}$$

which provide two dual fundamental existence result.

1.6.15 Classical integral transformations

On an oriented finite dimensional manifold $\mathbb{M}$ endowed with a standard volume form $\boldsymbol{\mu} \in BL\,(\mathbb{TM}^{(\dim \mathbb{M})}\,; \Re)$ the *divergence* of a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TM})$ is defined as the constant of proportionality between the LIE derivative of the standard volume form along the flow of the vector field and the standard volume form itself:

$$\mathcal{L}_{\mathbf{v}}\boldsymbol{\mu} = (\operatorname{div} \mathbf{v})\,\boldsymbol{\mu}\,.$$

The divergence may be also defined in terms of the exterior derivative by the relation

$$d(\boldsymbol{\mu}\mathbf{v}) = (\operatorname{div} \mathbf{v})\,\boldsymbol{\mu}\,.$$

Indeed, being $d\boldsymbol{\mu} = 0$, the homotopy formula tells us that

$$\mathcal{L}_{\mathbf{v}}\,\boldsymbol{\mu} = (d\boldsymbol{\mu})\mathbf{v} + d(\boldsymbol{\mu}\mathbf{v}) = d(\boldsymbol{\mu}\mathbf{v}) = (\operatorname{div} \mathbf{v})\,\boldsymbol{\mu}\,.$$

From STOKES formula we may then derive the classical integral transformation theorems. Indeed by the definition of

gradient:	$d\,f = \mathbf{i}_{(\nabla f)}\mathbf{g} = \mathbf{g}\nabla f\,,$		$\dim \mathbb{M} = n$
curl:	$d(\mathbf{g}\mathbf{v}) = (\operatorname{rot} \mathbf{v})\,\boldsymbol{\mu}_{\mathbf{g}}\,,$		$\dim \mathbb{M} = 2$
curl:	$d(\mathbf{g}\mathbf{v}) = \mathbf{i}_{(\operatorname{rot} \mathbf{v})}\boldsymbol{\mu}_{\mathbf{g}} = \boldsymbol{\mu}_{\mathbf{g}}(\operatorname{rot} \mathbf{v})\,,$		$\dim \mathbb{M} = 3$
divergence:	$d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = (\operatorname{div} \mathbf{v})\boldsymbol{\mu}_{\mathbf{g}}\,,$		$\dim \mathbb{M} = n$

we get the following statements:

- $\dim \mathbb{M} = n\,, \quad \Gamma \subset \mathbb{M}\,, \quad \dim \Gamma = 1:$ the *gradient theorem*:

  $$\int_{\Gamma} df = \int_{\Gamma} \mathbf{g}\nabla f = \int_{\Gamma} \mathbf{g}(\nabla f, \mathbf{t})\ (\mathbf{g}\,\mathbf{t}) = \int_{\partial\Gamma} f = f(\mathbb{B}) - f(\mathbb{A})\,,$$

  where $\mathbb{A}$, $\mathbb{B}$ are the end points of the curve Γ and $(\mathbf{g}\,\mathbf{t}) = \mathbf{i}_{\mathbf{t}}\,\mathbf{g}$ is the volume form (the signed-length) induced along the curve Γ.

- $\dim \mathbb{M} = 3\,, \quad \Sigma \subset \mathbb{M}\,, \quad \dim \Sigma = 2\,:$ the *curl theorem*:

$$\int_{\Sigma} d(\mathbf{g}\mathbf{v}) = \int_{\Sigma} \boldsymbol{\mu}_{\mathbf{g}}(\mathrm{rot}\,\mathbf{v}) = \int_{\Sigma} \mathbf{g}(\mathrm{rot}\,\mathbf{v}, \mathbf{n})\ (\boldsymbol{\mu}_{\mathbf{g}}\mathbf{n}) = \int_{\partial\Sigma} \mathbf{g}\mathbf{v} = \int_{\partial\Sigma} \mathbf{g}(\mathbf{v}, \mathbf{t})\ (\mathbf{g}\,\mathbf{t})\,,$$

with $\mathbf{n}$ unit normal to the surface Σ and $\mathbf{t}$ unit tangent to the boundary of the surface. For $\dim \mathbb{M} = 2$ the *curl theorem* writes:

$$\int_{\mathbb{M}} d(\mathbf{g}\mathbf{v}) = \int_{\mathbb{M}} (\mathrm{rot}\,\mathbf{v})\boldsymbol{\mu}_{\mathbf{g}} = \int_{\mathbb{M}} d(\mathbf{g}\mathbf{v}) = \int_{\partial\mathbb{M}} \mathbf{g}\mathbf{v}\,.$$

- $\dim \mathbb{M} = n\,:$ the *divergence theorem*:

$$\int_{\mathbb{M}} d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = \int_{\mathbb{M}} (\mathrm{div}\,\mathbf{v})\,\boldsymbol{\mu}_{\mathbf{g}} = \int_{\partial\mathbb{M}} \boldsymbol{\mu}_{\mathbf{g}}\mathbf{v} = \int_{\partial\mathbb{M}} \mathbf{g}(\mathbf{v}, \mathbf{n})\ (\boldsymbol{\mu}_{\mathbf{g}}\mathbf{n})\,,$$

  with $\mathbf{n}$ unit normal to the boundary $\partial\mathbb{M}$.

Remark 1.6.2 *The definition of gradient, curl and divergence in $\Re^3$ given above are based on the following algebraic results.*

- *To any one-form df on $\Re^n$ there correspond a unique vector ∇f in $\Re^n$ such that $df = \mathbf{g}\nabla f$.*
- *To any two-form $\boldsymbol{\omega}^2$ on $\Re^3$ there correspond a unique vector $\mathbf{w}$ in $\Re^3$ such that $\boldsymbol{\omega}^2 = \boldsymbol{\mu}\mathbf{w}$, with $\boldsymbol{\mu}$ a given volume form.*
- *All volume forms $\boldsymbol{\mu}$ on $\Re^n$ are proportional one another.*

Let us prove the second statement. The linear space $\Lambda^2(\Re^3)$ is 3-dimensional since $C^3_2 = C^3_1 = 3$. The linear space $S \subseteq \Lambda^2(\Re^3)$, spanned by the 2-forms $\mathbf{i}_{\mathbf{w}}\boldsymbol{\mu}$ on $\Re^3$ when $\mathbf{w}$ ranges in $\Re^3$, is also 3-dimensional since the forms $\mathbf{i}_{\mathbf{e}_i}\boldsymbol{\mu}$, with $\{\mathbf{e}_i\,, i = 1, 2, 3\}$ a basis, are linearly independent. Indeed

$$\sum_{i=1,3} \lambda_i\,(\mathbf{i}_{\mathbf{e}_i}\boldsymbol{\mu}) = \sum_{i=1,3} (\mathbf{i}_{(\lambda_i\,\mathbf{e}_i)}\boldsymbol{\mu}) = 0 \implies \lambda_i = 0\,, i = 1, \ldots, 3\,,$$

since otherwise, taking a basis $\{\mathbf{a}, \mathbf{b}, \sum_{i=1,3} \lambda_i\,\mathbf{e}_i\}$ in $\Re^3$, the volume

$$\boldsymbol{\mu}(\sum_{i=1,3} \lambda_i\,\mathbf{e}_i, \mathbf{a}, \mathbf{b}) = (\mathbf{i}_{(\sum_{i=1,3} \lambda_i\,\mathbf{e}_i)}\boldsymbol{\mu})(\mathbf{a}, \mathbf{b})$$

would be zero, whilst volume forms are non vanishing when evaluated on a basis. Then $S = \Lambda^2(\Re^3)$ and the correspondence $\mathbf{i}_{(\cdot)}\boldsymbol{\mu} \in BL\,(\Re^3\,; \Lambda^2(\Re^3))$ is a linear isomorphism.

1.6.16 Curvilinear coordinates

The definition of gradient, divergence and curl in terms of exterior derivative leads to simple formulas in curvilinear coordinates. Indeed let $\{\partial_1, \ldots, \partial_n\}$ be a basis of a system of curvilinear coordinates and

$$(d\boldsymbol{\omega}^k)(\partial_0, \partial_1, \ldots, \partial_k) = \sum_{i=0,k} (-1)^i \, \partial_i \, (\omega^k(\partial_0, \partial_1, \ldots, \partial_k)_i) \, ,$$

the coordinate formula for the exterior derivative provided in section 1.6.6.

- The divergence of a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ is expressed in curvilinear coordinates by

$$\operatorname{div} \mathbf{v} = \frac{1}{\boldsymbol{\mu}_{\mathbf{g}}(\partial_1, \ldots, \partial_n)} \sum_{i=1,n} \partial_i \, (v^i \, \boldsymbol{\mu}_{\mathbf{g}}(\partial_1, \ldots, \partial_n)) \, .$$

Indeed $d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v}) = (\operatorname{div} \mathbf{v})\boldsymbol{\mu}_{\mathbf{g}}$ and

$$\begin{aligned} d(\boldsymbol{\mu}_{\mathbf{g}}\mathbf{v})(\partial_1, \ldots, \partial_n) &= \sum_{i=1,n} (-1)^{(i-1)} \, \partial_i \, (\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{v}, \partial_1, \ldots, \partial_n)_i) \\ &= \sum_{i=1,n} \partial_i \, (v^i \, \boldsymbol{\mu}_{\mathbf{g}}(\partial_1, \ldots, \partial_n)) \, , \end{aligned}$$

where we have made use of the formulas

$$\begin{aligned} \partial_i \, (\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{v}, \partial_1, \ldots, \partial_n)_i) &= \partial_i \, (\boldsymbol{\mu}_{\mathbf{g}}(\sum_{k=1,n} v^k \partial_k, \partial_1, \ldots, \partial_n)_i) \\ &= \partial_i \, (v^i (\boldsymbol{\mu}_{\mathbf{g}} \partial_i)(\partial_1, \ldots, \partial_n)_i) \, , \end{aligned}$$

and

$$(-1)^{(i-1)} \, \partial_i \, (v^i (\boldsymbol{\mu}_{\mathbf{g}} \partial_i)(\partial_1, \ldots, \partial_n)_i) = \partial_i \, (v^i \, \boldsymbol{\mu}_{\mathbf{g}}(\partial_1, \ldots, \partial_n)) \, .$$

In orthogonal curvilinear coordinates the metric volume form may be evaluated as:

$$\boldsymbol{\mu}_{\mathbf{g}}(\partial_1, \ldots, \partial_n) = \prod_{i=1,n} \sqrt{\mathbf{g}(\partial_i, \partial_i)} = \prod_{i=1,n} h_i \, .$$

In terms of the engineering components $\hat{v}^i = v^i \, h_i$, (not summed) with $h_i = ||\partial_i||$, the formula above takes the form

$$\operatorname{div} \mathbf{v} = \frac{1}{\Big(\prod\limits_{i=1,n} h_i\Big)} \sum_{i=1,n} \partial_i \Big(\hat{v}^i \prod_{\substack{j=1,n \\ j \neq i}} h_j \Big) .$$

We remark that the engineering components are evaluated with respect to the normalized basis

$$\{\hat{\partial}_1, \ldots, \hat{\partial}_n\} , \qquad \text{with} \qquad \hat{\partial}_i = \frac{\partial_i}{||\partial_i||} .$$

A similar analysis can be perfomed to derive the component expressions of the gradient of a scalar field and the curl of a vector field in curvilinear coordinates. The issue is briefly illustrated below.

- For the gradient of a scalar field $f \in \mathrm{C}^1(\mathbb{M} ; \Re)$ in curvilinear coordinates we have:

  $$df(\partial_i) = \partial_i \, f = \mathbf{g}(\nabla f, \partial_i) = \mathbf{g}(\partial_i, \partial_k)(\nabla f)_k ,$$

  so that

  $$\nabla f = (\mathbf{G}^{-1})_{ik} \, (\partial_k \, f) \, \partial_i ,$$

  and in orthogonal curvilinear coordinates:

  $$\nabla f = \frac{\partial_i \, f}{h_i^2} \, \partial_i = \frac{\partial_i \, f}{h_i} \, \hat{\partial}_i .$$

- For the curl of a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M} ; \mathbb{TM})$ in curvilinear coordinates we have:

$$\begin{aligned} d(\mathbf{g}\mathbf{v})(\partial_2, \partial_3) &= \partial_2 \, \mathbf{g}(\mathbf{v}, \partial_3) - \partial_3 \, \mathbf{g}(\mathbf{v}, \partial_2) \\ &= \mathbf{i}_{(\mathrm{rot}\, \mathbf{v})} \boldsymbol{\mu}_{\mathbf{g}}(\partial_2, \partial_3) = (\mathrm{rot}\, \mathbf{v})^1 \, \boldsymbol{\mu}(\partial_1, \partial_2, \partial_3) \end{aligned}$$

$$\begin{aligned} d(\mathbf{g}\mathbf{v})(\partial_1, \partial_3) &= \partial_1 \, \mathbf{g}(\mathbf{v}, \partial_3) - \partial_3 \, \mathbf{g}(\mathbf{v}, \partial_1) \\ &= \mathbf{i}_{(\mathrm{rot}\, \mathbf{v})} \boldsymbol{\mu}_{\mathbf{g}}(\partial_1, \partial_3) = -(\mathrm{rot}\, \mathbf{v})^2 \, \boldsymbol{\mu}(\partial_1, \partial_2, \partial_3) \end{aligned}$$

$$\begin{aligned} d(\mathbf{g}\mathbf{v})(\partial_1, \partial_2) &= \partial_1 \, \mathbf{g}(\mathbf{v}, \partial_2) - \partial_2 \, \mathbf{g}(\mathbf{v}, \partial_1) \\ &= \mathbf{i}_{(\mathrm{rot}\, \mathbf{v})} \boldsymbol{\mu}_{\mathbf{g}}(\partial_1, \partial_2) = (\mathrm{rot}\, \mathbf{v})^3 \, \boldsymbol{\mu}(\partial_1, \partial_2, \partial_3) , \end{aligned}$$

so that

$$(\mathrm{rot}\,\mathbf{v})^1 = \frac{1}{\boldsymbol{\mu}(\partial_1, \partial_2, \partial_3)}(\partial_2\,\mathbf{g}(\mathbf{v}, \partial_3) - \partial_3\,\mathbf{g}(\mathbf{v}, \partial_2))$$

$$(\mathrm{rot}\,\mathbf{v})^2 = \frac{1}{\boldsymbol{\mu}(\partial_1, \partial_2, \partial_3)}(\partial_3\,\mathbf{g}(\mathbf{v}, \partial_1) - \partial_1\,\mathbf{g}(\mathbf{v}, \partial_3))$$

$$(\mathrm{rot}\,\mathbf{v})^3 = \frac{1}{\boldsymbol{\mu}(\partial_1, \partial_2, \partial_3)}(\partial_1\,\mathbf{g}(\mathbf{v}, \partial_2) - \partial_2\,\mathbf{g}(\mathbf{v}, \partial_1))\,,$$

and in orthogonal curvilinear coordinates:

$$(\mathrm{rot}\,\mathbf{v})^1 = \frac{1}{h_1 h_2 h_3}(\partial_2 v^3 - \partial_3 v^2)$$

$$(\mathrm{rot}\,\mathbf{v})^2 = \frac{1}{h_1 h_2 h_3}(\partial_3 v^1 - \partial_1 v^3)$$

$$(\mathrm{rot}\,\mathbf{v})^3 = \frac{1}{h_1 h_2 h_3}(\partial_1 v^2 - \partial_2 v^1)\,.$$

1.6.17 Reynolds theorem

The classical form of REYNOLDS theorem may obtained from the transport Theorem 1.6.3 by setting $\boldsymbol{\omega}_t = f_t\,\boldsymbol{\mu}$, with $\boldsymbol{\mu}$ volume-form on the n-dimensional ambient manifold $\mathbb{M}$, and choosing Γ to be an n-dimensional submanifold embedded in $\mathbb{M}$. Then the transport formula writes

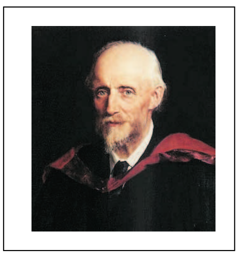

Figure 1.38: Osborne Reynolds (1842 - 1912)

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,t}(\Gamma)} f_t \, \boldsymbol{\mu} = \int_{\Gamma} \mathcal{L}_{t,\mathbf{v}} \, (f_t \, \boldsymbol{\mu})$$

$$= \int_{\Gamma} (\partial_{\tau=t} \; f_\tau) \, \boldsymbol{\mu} + \mathcal{L}_{\mathbf{v}} \, (f_t \, \boldsymbol{\mu}) \, .$$

Then from the formulas $\mathcal{L}_{\mathbf{v}} \, (f \, \boldsymbol{\mu}) = (\mathcal{L}_{\mathbf{v}} f) \, \boldsymbol{\mu} + f \, (\mathcal{L}_{\mathbf{v}} \boldsymbol{\mu})$ and $\mathcal{L}_{\mathbf{v}} \boldsymbol{\mu} = (\operatorname{div} \mathbf{v}) \, \boldsymbol{\mu}$ we infer that

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,t}(\Gamma)} f \, \boldsymbol{\mu} = \int_{\Gamma} (\partial_{\tau=t} \; f_\tau + \mathcal{L}_{\mathbf{v}} f) \, \boldsymbol{\mu} + \int_{\Gamma} f \, (\operatorname{div} \mathbf{v}) \, \boldsymbol{\mu}$$

$$= \int_{\Gamma} (\mathcal{L}_{t,\mathbf{v}} \, f + f \operatorname{div} \mathbf{v}) \, \boldsymbol{\mu} \, .$$

An alternative expression of the transport theorem may be obtained by formula v) of Proposition 1.3.11 and the definition of divergence of a vector field. Indeed, being $\boldsymbol{\mu}$ a volume-form, we have that

$$\mathcal{L}_{\mathbf{v}} \, (f \, \boldsymbol{\mu}) = \mathcal{L}_{(f \, \mathbf{v})} \, \boldsymbol{\mu} = \operatorname{div} \, (f \, \mathbf{v}) \, \boldsymbol{\mu} \, ,$$

we get

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,t}(\Gamma)} f \, \boldsymbol{\mu} = \int_{\Gamma} \mathcal{L}_{t,(f\mathbf{v})} \, \boldsymbol{\mu} = \int_{\Gamma} (\partial_t \, f + \mathcal{L}_{(f \, \mathbf{v})}) \, \boldsymbol{\mu} +$$

$$= \int_{\Gamma} (\partial_t \, f) \, \boldsymbol{\mu} + \int_{\Gamma} \operatorname{div} \, (f \, \mathbf{v}) \, \boldsymbol{\mu}$$

$$= \int_{\Gamma} (\partial_t \, f) \, \boldsymbol{\mu} + \int_{\partial\Gamma} f \, \mathbf{g}(\mathbf{v}, \mathbf{n}) \; (\boldsymbol{\mu}\mathbf{n}) \, ,$$

where the last formula follows from the divergence theorem (see section 1.6.15).

This last expression of the transport theorem tells us that

- the time-rate of increase of an extensive quantity evaluated on a flowing manifold is equal to the time-rate of increase evaluated by frozing the flow plus the time-rate of supply of its density thru the boundary.

It should be noted that the transport theorem for vector or tensor fields, other than volume-form fields, is not feasible on differentiable manifolds since the integral of such fields makes no sense.

The extension of these results from the euclidean space to manifolds can be performed by adopting a variational form which requires only the integration of volume-form fields.

1.7 Graded derivation algebra

- A *differential k-form* $\boldsymbol{\omega} \in \Lambda^k(\mathbb{M}\,;\Re)$ on a manifold $\mathbb{M}$ is a differentiable field of k-forms on $\mathbb{M}$.

We denote by $\Lambda(\mathbb{M}\,;\Re)$ the graded commutative algebra of differential forms on $\mathbb{M}$ with the associative and graded commutative exterior multiplication:

$$\boldsymbol{\omega}^k \wedge (\boldsymbol{\omega}^h \wedge \boldsymbol{\omega}^l) = (\boldsymbol{\omega}^k \wedge \boldsymbol{\omega}^h) \wedge \boldsymbol{\omega}^l\,, \qquad \boldsymbol{\omega}^k \wedge \boldsymbol{\omega}^h = (-1)^{kh}\, \boldsymbol{\omega}^h \wedge \boldsymbol{\omega}^k\,.$$

Definition 1.7.1 *The space* $\mathrm{DER}_s\,\Lambda(\mathbb{M}\,;\Re)$ *of graded derivations of degree* s *, is made of the linear maps* $D \in BL\,(\Lambda(\mathbb{M}\,;\Re)\,;\Lambda(\mathbb{M}\,;\Re))$ *with*

$$D(\Lambda^q(\mathbb{M}\,;\Re)) \subset \Lambda^{q+s}(\mathbb{M}\,;\Re)$$

fulfilling, for any $\boldsymbol{\alpha} \in \Lambda(\mathbb{M}\,;\Re)$ *, the graded* LEIBNIZ *rule:*

$$D(\boldsymbol{\omega} \wedge \boldsymbol{\alpha}) = D(\boldsymbol{\omega}) \wedge \boldsymbol{\alpha} + (-1)^{\deg D \deg \boldsymbol{\omega}} \boldsymbol{\omega} \wedge D(\boldsymbol{\alpha})\,.$$

By virtue of the graded LEIBNIZ rule, a graded derivation is completely defined by its action on 0-forms and 1-forms, since any k-form is pointwise uniquely expressed as a linear combination of k-th exterior products of 1-forms.

The space $\mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ of graded derivations of any degree, is a *graded* LIE *algebra* whose bracket is the *graded commutator*

$$[D_1\,, D_2] := D_1 \circ D_2 - (-1)^{\deg D_1 \deg D_2} D_2 \circ D_1\,,$$

fulfilling the *graded anticommutativity* relation:

$$[D_1\,, D_2] := -(-1)^{\deg D_1 \deg D_2} [D_2\,, D_1]\,,$$

and the *graded* JACOBI *identity*:

$$[D_1\,, [D_2\,, D_3]] = [[D_1\,, D_2]\,, D_3] + (-1)^{\deg D_1 \deg D_2} [D_2\,, [D_1\,, D_3]]\,.$$

To any derivation $D \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ there corresponds an *adjoint derivation*, defined by

$$\mathrm{ADJ}_D(\cdot) := [D\,, \cdot\,]\,,$$

Indeed, by the graded JACOBI identity we have that

$$\mathrm{ADJ}_D([D_1\,, D_2]) = [\mathrm{ADJ}_D(D_1)\,, D_2] + (-1)^{\deg D \deg D_1} [D_1\,, \mathrm{ADJ}_D(D_2)]\,.$$

and hence also $\mathrm{ADJ}_D \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ with $\deg \mathrm{ADJ}_D = \deg D$.

Let $\boldsymbol{\omega} \in \Lambda(\mathbb{M}\,;\Re)$. It is easy to see that:

- The *insertion* operator $\mathbf{i_v} : \Lambda^q(\mathbb{M}\,;\Re) \mapsto \Lambda^{q-1}(\mathbb{M}\,;\Re)$ is a derivation of degree -1. Indeed

$$\mathbf{i_v}(\boldsymbol{\omega} \wedge \boldsymbol{\alpha}) = (\mathbf{i_v}\boldsymbol{\omega}) \wedge \boldsymbol{\alpha} + (-1)^{\deg \boldsymbol{\omega}}\, \boldsymbol{\omega} \wedge \mathbf{i_v}\boldsymbol{\alpha}\,.$$

- The *Lie* derivation $\mathcal{L}_\mathbf{v} : \Lambda^q(\mathbb{M}\,;\Re) \mapsto \Lambda^q(\mathbb{M}\,;\Re)$ is of degree 0. Indeed

$$\mathcal{L}_\mathbf{v}(\boldsymbol{\omega} \wedge \boldsymbol{\alpha}) = (\mathcal{L}_\mathbf{v}\boldsymbol{\omega}) \wedge \boldsymbol{\alpha} + \boldsymbol{\omega} \wedge \mathcal{L}_\mathbf{v}\boldsymbol{\alpha}\,.$$

- The *exterior* derivation $d : \Lambda^q(\mathbb{M}\,;\Re) \mapsto \Lambda^{q+1}(\mathbb{M}\,;\Re)$ is of degree $+1$. Indeed

$$d(\boldsymbol{\omega} \wedge \boldsymbol{\alpha}) = (d\boldsymbol{\omega}) \wedge \boldsymbol{\alpha} + (-1)^{\deg \boldsymbol{\omega}}\, \boldsymbol{\omega} \wedge d\boldsymbol{\alpha}\,.$$

The *graded commutation* rule is in accordance with the formula

$$[\mathcal{L}_\mathbf{u}\,, \mathbf{i_v}] = \mathcal{L}_\mathbf{u} \circ \mathbf{i_v} - \mathbf{i_v} \circ \mathcal{L}_\mathbf{u}\,,$$

and the *differential homotopy formula* takes the simple expression

$$\mathcal{L}_\mathbf{v} = [\mathbf{i_v}\,, d] = \mathbf{i_v} \circ d + d \circ \mathbf{i_v}\,.$$

Definition 1.7.2 *A derivation* $D \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ *is algebraic if it vanishes on* 0*-forms:* $D(f) = 0\,, \quad \forall\, f \in \mathrm{C}^\infty(\mathbb{M}\,;\Re)$.

Being

$$D(f\boldsymbol{\omega}) = f\, D(\boldsymbol{\omega})\,, \quad \forall\, f \in \mathrm{C}^\infty(\mathbb{M}\,;\Re)\,,$$

we infer that a derivation is algebraic if and only if it is tensorial, i.e. lives at points. By virtue of the graded LEIBNIZ rule, an algebraic graded derivation is completely defined by its action on 1-forms and hence on differentials of functions which generate the space of one-forms. The action of an algebraic derivation $D \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ is equivalent to the action of an insertion operator and we may write

$$D(\boldsymbol{\omega}) = \mathbf{i_L}\boldsymbol{\omega}\,,$$

where $\deg \mathbf{L} = \deg D + 1$ and $\mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ is a tangent-valued exterior form. Indeed the vectorial value of the form $\mathbf{L}$, evaluated on its multi-argument whose cardinality is $\deg \mathbf{L}$, takes the first position in the list of the multi-argument of $\boldsymbol{\omega}$, whose cardinality is $\deg \boldsymbol{\omega}$, so that the final list of arguments has cardinality $\deg \boldsymbol{\omega} + \deg \mathbf{L} - 1$ and hence $\deg \mathbf{i_L} = \deg \mathbf{L} - 1$. To get an alternating form, the definition of $\mathbf{i_L}(\boldsymbol{\omega}) \in \Lambda^{\ell+k-1}(\mathbb{M}\,;\mathbb{TM})$, with $\deg \mathbf{L} = \ell$ and $\deg \boldsymbol{\omega} = k$, is given by

$$(\mathbf{i_L}\boldsymbol{\omega})(\mathbf{X}_1, \dots \mathbf{X}_{\ell+k-1}) :=$$
$$\sum_{\sigma \in \Sigma(\ell+k-1)} \frac{\operatorname{sign}\sigma}{(k-1)!(\ell)!}\, \boldsymbol{\omega}(\mathbf{L}(\mathbf{X}_{\sigma(1)}, \dots \mathbf{X}_{\sigma(\ell)}), \mathbf{X}_{\sigma(\ell+1)}, \dots, \mathbf{X}_{\sigma(\ell+k-1)}),$$

where $\mathbf{X}_i \in \mathbb{TM}$. Let us write explicitly the following special cases. Then if $\ell = 0$ and $k = 2$ i.e. $\mathbf{L} \in \Lambda^0(\mathbb{M}\,;\mathbb{TM})$ and $\boldsymbol{\omega} \in \Lambda^2(\mathbb{M}\,;\Re)$, then

$$(\mathbf{i_L}\boldsymbol{\omega})(\mathbf{X}) = \boldsymbol{\omega}(\mathbf{L}, \mathbf{X}).$$

If $\ell = 1$ and $k = 2$ i.e. $\mathbf{L} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ and $\boldsymbol{\omega} \in \Lambda^2(\mathbb{M}\,;\Re)$ then

$$(\mathbf{i_L}\boldsymbol{\omega})(\mathbf{X}, \mathbf{Y}) = \boldsymbol{\omega}(\mathbf{L} \cdot \mathbf{X}, \mathbf{Y}) + \boldsymbol{\omega}(\mathbf{X}, \mathbf{L} \cdot \mathbf{Y}).$$

A non-algebraic derivation $D \in \text{DER}\,\Lambda(\mathbb{M}\,;\Re)$ writes

$$D(\boldsymbol{\omega}) = \mathcal{L}_{\mathbf{K}}(\boldsymbol{\omega}),$$

with the LIE-NIJENHUIS derivative along the tangent-valued form $\mathbf{K} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ is defined by a formal extension of the homotopy formula to a *graded homotopy formula* [58]:

$$\mathcal{L}_{\mathbf{K}} := [\mathbf{i_K}\,, d] = \mathbf{i_K} \circ d - (-1)^{(\deg \mathbf{K}-1)} d \circ \mathbf{i_K}\,,$$

where d is the exterior derivative. Note that $\deg \mathbf{i_K} = \deg \mathbf{K} - 1$ and $\deg d = 1$ so that $\deg D = \deg \mathcal{L}_{\mathbf{K}} = \deg \mathbf{K}$.

If $\mathbf{K} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ is a 0-form, it is in fact a vector field and its degree is zero. The graded homotopy formula for the LIE derivative reduces then to the homotopy formula.

Denoting by $\mathbf{I} = \mathbf{id}_{\mathbb{TM}} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ the identity form, we have: $\mathbf{i_I}\boldsymbol{\omega} = (\deg \boldsymbol{\omega})\boldsymbol{\omega}$. Hence $\mathcal{L}_{\mathbf{I}} = d$ since

$$\mathcal{L}_{\mathbf{I}}\boldsymbol{\omega} = [\mathbf{i_I}\,, d]\boldsymbol{\omega} = \mathbf{i_I} \circ d\boldsymbol{\omega} - d \circ \mathbf{i_I}\boldsymbol{\omega} = (\deg \boldsymbol{\omega} + 1)d\boldsymbol{\omega} - (\deg \boldsymbol{\omega})d\boldsymbol{\omega} = d\boldsymbol{\omega}.$$

The next Lemma plays a basic role in the theory of graded derivations.

Lemma 1.7.1 *The linear map $\mathcal{L} \in BL(\Lambda(\mathbb{M}\,;\mathbb{TM})\,;\text{DER}\,\Lambda(\mathbb{M}\,;\Re))$ which associates $\mathcal{L}(\mathbf{L}) := \mathcal{L}_{\mathbf{L}} \in \text{DER}\,\Lambda(\mathbb{M}\,;\Re)$ with $\mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ is injective. Moreover the LIE derivatives of tangent-valued forms are in the null space of the adjoint of the exterior derivative, that is:*

$$[\mathcal{L}_{\mathbf{K}}\,, d] = 0\,, \quad \forall\, \mathbf{K} \in \Lambda(\mathbb{M}\,;\mathbb{TM}),$$

and the only algebraic derivation in the null space of the adjoint of the exterior derivative is the null algebraic derivation, that is:

$$[\mathbf{i}_{\mathbf{L}}\,, d] = 0 \iff \mathbf{L} = 0\,, \qquad \mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})\,.$$

Proof. The first assertion follows from the fact that $\mathcal{L}_{\mathbf{L}} f = 0$ for all $f \in \mathrm{C}^{\infty}(\mathbb{M}\,;\Re)$ implies that $\mathbf{L} = 0$. The second assertion is proved by a direct computation. Indeed, being $[d\,, d] = 2\,d \circ d = 0$, the graded JACOBI identity yields

$$0 = [\mathbf{i}_{\mathbf{K}}\,, [d\,, d]] = [[\mathbf{i}_{\mathbf{K}}\,, d]\,, d] + (-1)^{\deg \mathbf{K}}[d\,, [\mathbf{i}_{\mathbf{K}}\,, d]] = 2\,[[\mathbf{i}_{\mathbf{K}}\,, d]\,, d] = 2\,[\mathcal{L}_{\mathbf{K}}\,, d]\,.$$

The third assertion is clear since by definition $\mathcal{L}_{\mathbf{L}} := [\mathbf{i}_{\mathbf{L}}\,, d]$. ■

From Lemma 1.7.1 we infer that the adjoint of the exterior derivative acts on a derivation as a projection on the space of the algebraic derivations of the same degree.

Proposition 1.7.1 (Graded derivations) *A derivation* $D \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ *may be written uniquely as*

$$D = \mathcal{L}_{\mathbf{K}} + \mathbf{i}_{\mathbf{L}}\,,$$

with $\deg \mathbf{L} = \deg D + 1$ *and* $\deg \mathbf{K} = \deg D$. *Then* $\mathbf{L} = 0$ *if and only if* $[D\,, d] = 0$ *and the derivation is algebraic if and only if* $\mathbf{K} = 0$.

Proof. Let us evaluate a derivation $D \in \mathrm{DER}_k\,\Lambda(\mathbb{M}\,;\Re)$ on smooth scalar fields (0-forms) $f \in \Lambda^0(\mathbb{M}\,;\Re)$. Then $Df \in \Lambda^k(\mathbb{M}\,;\Re)$ is a k-form but also a point derivation on $\mathbb{M}$. Then Df can be valuated as the derivative of $f \in \Lambda^0(\mathbb{M}\,;\Re)$ along the point value of a uniquely defined tangent-valued exterior k-form $\mathbf{K} \in \Lambda^k(\mathbb{M}\,;\mathbb{TM})$:

$$Df = Tf \cdot \mathbf{K} = (\mathbf{i}_{\mathbf{K}} \circ d)f = \mathcal{L}_{\mathbf{K}} f\,.$$

Then $D - \mathcal{L}_{\mathbf{K}}$ is an algebraic derivation and we may write that $D - \mathcal{L}_{\mathbf{K}} = \mathbf{i}_{\mathbf{L}}$ for a unique tangent-valued form $\mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ with $\deg \mathbf{L} = \deg D + 1$. Lemma 1.7.1 and the formula

$$[D\,, d] = [\mathcal{L}_{\mathbf{K}}\,, d] + [\mathbf{i}_{\mathbf{L}}\,, d] = [\mathbf{i}_{\mathbf{L}}\,, d] = \mathcal{L}_{\mathbf{L}}\,,$$

provide the proofs of the last two assertions. ■

1.7.1 Nijenhuis-Richardson bracket

The graded commutator of two algebraic derivations is still an algebraic derivation and we may define the NIJENHUIS-RICHARDSON bracket by:

$$\mathbf{i}_{\mathrm{NR}(\mathbf{K},\mathbf{L})} := [\mathbf{K}\,,\mathbf{L}]\,,$$

with $\deg \mathrm{NR}(\mathbf{K},\mathbf{L}) = \deg \mathbf{K} + \deg \mathbf{L} - 1$. The explicit expression is given by

$$\mathrm{NR}(\mathbf{K},\mathbf{L}) = \mathbf{i}_{\mathbf{K}}\mathbf{L} - (-1)^{(\deg \mathbf{K}-1)(\deg \mathbf{L}-1)}\,\mathbf{i}_{\mathbf{L}}\mathbf{K}\,,$$

where $\mathbf{i}_{\mathbf{L}}\mathbf{K}$ is defined by setting $\mathbf{K} = \boldsymbol{\omega} \otimes \mathbf{X}$ with $\deg(\boldsymbol{\omega}) = \deg(\mathbf{K})$ and

$$\mathbf{i}_{\mathbf{L}}(\boldsymbol{\omega} \otimes \mathbf{X}) := (\mathbf{i}_{\mathbf{L}}\boldsymbol{\omega}) \otimes \mathbf{X}\,.$$

1.7.2 Frölicher-Nijenhuis bracket

The FRÖLICHER-NIJENHUIS bracket is a generalization of the LIE bracket to tangent-valued forms $\mathbf{K} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ and $\mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ on the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$, the value of the bracket being still a tangent-valued form. The reader is referenced to [80] for an exhaustive exposition of the topic.

To define the FRÖLICHER-NIJENHUIS bracket, briefly the FN-bracket, we recall the decomposition formula for a derivation $D \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$:

$$D = \mathcal{L}_{\mathbf{K}} + \mathbf{i}_{\mathbf{L}}\,.$$

The property $[\mathcal{L}_{\mathbf{K}}\,, d] = 0$ and the graded JACOBI identity tell us that, for any two tangent-valued forms $\mathbf{K}, \mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$, it is

$$[[\mathcal{L}_{\mathbf{K}}\,, \mathcal{L}_{\mathbf{L}}]\,, d] = 0\,.$$

Then, by Proposition 1.7.1, the derivation $[\mathcal{L}_{\mathbf{K}}\,, \mathcal{L}_{\mathbf{L}}] \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ may be written as $\mathcal{L}_{[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}}} \in \mathrm{DER}\,\Lambda(\mathbb{M}\,;\Re)$ with $[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ a tangent-valued form uniquely defined by the property

$$\mathcal{L}_{[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}}} := [\mathcal{L}_{\mathbf{K}}\,, \mathcal{L}_{\mathbf{L}}]\,,$$

By bilinearity, the map $(\mathbf{K}\,,\mathbf{L}) \rightarrow \mathrm{FN}(\mathbf{K},\mathbf{L}) \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ is a bracket, the FRÖLICHER-NIJENHUIS bracket, and

$$\deg[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}} = \deg \mathbf{K} + \deg \mathbf{L}\,.$$

The space $\Lambda(\mathbb{M}\,;\mathbb{TM})$ is a graded LIE algebra for the FN-bracket, fulfilling the *graded anticommutativity* relation:

$$[\mathbf{K}_1\,,\mathbf{K}_2]_{\rm FN} := -(-1)^{\deg \mathbf{K}_1 \deg \mathbf{K}_2}[\mathbf{K}_2\,,\mathbf{K}_1]_{\rm FN}\,,$$

and the *graded* JACOBI *identity*:

$$\begin{aligned}[\mathbf{K}_1\,,[\mathbf{K}_2\,,\mathbf{K}_3]_{\rm FN}]_{\rm FN} &= [[\mathbf{K}_1\,,\mathbf{K}_2]_{\rm FN}\,,\mathbf{K}_3]_{\rm FN} \\ &\quad +(-1)^{\deg \mathbf{K}_1 \deg \mathbf{K}_2}[\mathbf{K}_2\,,[\mathbf{K}_1\,,\mathbf{K}_3]_{\rm FN}]_{\rm FN}\,,\end{aligned}$$

equivalent to

$$\begin{aligned}&(-1)^{\deg \mathbf{K}_1 \deg \mathbf{K}_3}[\mathbf{K}_1\,,[\mathbf{K}_2\,,\mathbf{K}_3]_{\rm FN}]_{\rm FN} \\ &+(-1)^{\deg \mathbf{K}_2 \deg \mathbf{K}_1}[\mathbf{K}_2\,,[\mathbf{K}_3\,,\mathbf{K}_1]_{\rm FN}]_{\rm FN} \\ &+(-1)^{\deg \mathbf{K}_3 \deg \mathbf{K}_2}[\mathbf{K}_3\,,[\mathbf{K}_1\,,\mathbf{K}_2]_{\rm FN}]_{\rm FN} = 0\,.\end{aligned}$$

Being $\mathcal{L}_{\mathbf{I}} = d$ and $[\mathcal{L}_{\mathbf{K}}\,,d] = 0$ we get $[\mathbf{K}\,,\mathbf{I}]_{\rm FN} = 0$ for any $\mathbf{K} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$.

For vector fields, which are tangent-valued 0-forms, i.e. elements of $\Lambda^0(\mathbb{M}\,;\mathbb{TM})$ the FN-bracket coincides with the LIE bracket.

Lemma 1.7.2 *For* $\mathbf{K}, \mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ *we have that*

$$[\mathbf{i}_{\mathbf{L}}\,,\mathcal{L}_{\mathbf{K}}]_{\rm FN} = \mathcal{L}(\mathbf{i}_{\mathbf{L}}\mathbf{K}) + (-1)^{\deg \mathbf{K}}\mathbf{i}_{[\mathbf{L}\,,\mathbf{K}]_{\rm FN}}\,.$$

Proof. For $f \in {\rm C}^\infty(\mathbb{M}\,;\Re)$ we have that

$$\begin{aligned}[\mathbf{i}_{\mathbf{L}}\,,\mathcal{L}_{\mathbf{K}}]\,f &= (\mathbf{i}_{\mathbf{L}} \circ \mathcal{L}_{\mathbf{K}})\,f = (\mathbf{i}_{\mathbf{L}} \circ \mathbf{i}_{\mathbf{K}})\,d\,f \\ &= \mathbf{i}_{\mathbf{L}}(d\,f \circ \mathbf{K}) = d\,f \circ (\mathbf{i}_{\mathbf{L}}\mathbf{K}) = \mathcal{L}(\mathbf{i}_{\mathbf{L}}\mathbf{K})\,f\,,\end{aligned}$$

and hence $[\mathbf{i}_{\mathbf{L}}\,,\mathcal{L}_{\mathbf{K}}] - \mathcal{L}(\mathbf{i}_{\mathbf{L}}\mathbf{K})$ is an algebraic derivation. Moreover, by the graded JACOBI identity, we have that

$$\begin{aligned}[[\mathbf{i}_{\mathbf{L}}\,,\mathcal{L}_{\mathbf{K}}], d\,,] &= [\mathbf{i}_{\mathbf{L}}\,,[\mathcal{L}_{\mathbf{K}}\,,d]] - (-1)^{\deg \mathbf{K} \deg \mathbf{L}}[\mathcal{L}_{\mathbf{K}}\,,[\mathbf{i}_{\mathbf{L}}\,,d]] \\ &= -(-1)^{\deg \mathbf{K} \deg \mathbf{L}}[\mathcal{L}_{\mathbf{K}}\,,\mathcal{L}_{\mathbf{L}}] = -(-1)^{\deg \mathbf{K} \deg \mathbf{L}}\mathcal{L}_{[\mathbf{K}\,,\mathbf{L}]_{\rm FN}} \\ &= (-1)^{\deg \mathbf{K}}[\mathbf{i}_{[\mathbf{L}\,,\mathbf{K}]_{\rm FN}}\,,d]\,.\end{aligned}$$

The algebraic part of $[\mathbf{i}_{\mathbf{L}}\,,\mathcal{L}_{\mathbf{K}}] \in {\rm DER}\,\Lambda(\mathbb{M}\,;\Re)$ is equal to $(-1)^{\deg \mathbf{K}}\mathbf{i}_{[\mathbf{L}\,,\mathbf{K}]_{\rm FN}}$, by Lemma 1.7.1. ■

For $\mathbf{K} \in \Lambda^k(\mathbb{M}\,;\mathbb{TM})$ and $\boldsymbol{\omega} \in \Lambda^\ell(\mathbb{M}\,;\Re)$ the LIE-NIJENHUIS derivative $\mathcal{L}_\mathbf{K}\boldsymbol{\omega} \in \Lambda^{(\ell+k)}(\mathbb{M}\,;\Re)$ is expressed in terms of the LIE derivative by the formula

$$\begin{aligned}
&\mathcal{L}_\mathbf{K}\boldsymbol{\omega}(\mathbf{X}_1,\ldots\mathbf{X}_{k+\ell}) \\
&= \frac{1}{k!\,\ell!}\sum_\sigma \operatorname{sign}\sigma\,\mathcal{L}(\mathbf{K}(\mathbf{X}_{\sigma(1)},\ldots,\mathbf{X}_{\sigma(k)}))\,\boldsymbol{\omega}(\mathbf{X}_{\sigma(k+1)},\ldots\mathbf{X}_{\sigma(k+\ell)}) \\
&+\frac{-1}{k!\,(\ell-1)!}\sum_\sigma \operatorname{sign}\sigma\,\boldsymbol{\omega}([\mathbf{K}(\mathbf{X}_{\sigma(1)},\ldots,\mathbf{X}_{\sigma(k)})\,,\mathbf{X}_{\sigma(k+1)}],\mathbf{X}_{\sigma(k+2)},\ldots) \\
&+\frac{(-1)^{k-1}}{(k-1)!\,(\ell-1)!\,2!}\sum_\sigma \operatorname{sign}\sigma\,\boldsymbol{\omega}(\mathbf{K}([\mathbf{X}_{\sigma(1)}\,,\mathbf{X}_{\sigma(2)}],\mathbf{X}_{\sigma(3)},\ldots),\mathbf{X}_{\sigma(k+2)},\ldots)\,.
\end{aligned}$$

For $\mathbf{K} \in \Lambda^k(\mathbb{M}\,;\mathbb{TM})$ and $\mathbf{L} \in \Lambda^\ell(\mathbb{M}\,;\mathbb{TM})$ the FRÖLICHER-NIJENHUIS bracket $[\mathbf{K}\,,\mathbf{L}] \in \Lambda^{(k+\ell)}(\mathbb{M}\,;\mathbb{TM})$ is expressed in terms of the LIE bracket by the formula [103], [115]:

$$\begin{aligned}
&[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}}(\mathbf{X}_1,\ldots\mathbf{X}_{k+\ell}) \\
&= \frac{1}{k!\,\ell!}\sum_\sigma \operatorname{sign}\sigma\,[\mathbf{K}(\mathbf{X}_{\sigma(1)},\ldots\mathbf{X}_{\sigma(k)})\,,\mathbf{L}(\mathbf{X}_{\sigma(k+1)},\ldots\mathbf{X}_{\sigma(k+\ell)})] \\
&+\frac{-1}{k!\,(\ell-1)!}\sum_\sigma \operatorname{sign}\sigma\,\mathbf{L}([\mathbf{K}(\mathbf{X}_{\sigma(1)},\ldots,\mathbf{X}_{\sigma(k)})\,,\mathbf{X}_{\sigma(k+1)}],\mathbf{X}_{\sigma(k+2)},\ldots) \\
&+\frac{(-1)^{k\ell}}{(k-1)!\,\ell!}\sum_\sigma \operatorname{sign}\sigma\,\mathbf{K}([\mathbf{L}(\mathbf{X}_{\sigma(1)},\ldots,\mathbf{X}_{\sigma(\ell)})\,,\mathbf{X}_{\sigma(\ell+1)}],\mathbf{X}_{\sigma(\ell+2)},\ldots) \\
&+\frac{(-1)^{k-1}}{(k-1)!\,(\ell-1)!\,2!}\sum_\sigma \operatorname{sign}\sigma\,\mathbf{L}(\mathbf{K}([\mathbf{X}_{\sigma(1)}\,,\mathbf{X}_{\sigma(2)}],\mathbf{X}_{\sigma(3)},\ldots),\mathbf{X}_{\sigma(k+2)},\ldots) \\
&+\frac{(-1)^{k-1}}{(k-1)!\,(\ell-1)!\,2!}\sum_\sigma \operatorname{sign}\sigma\,\mathbf{K}(\mathbf{L}([\mathbf{X}_{\sigma(1)}\,,\mathbf{X}_{\sigma(2)}],\mathbf{X}_{\sigma(3)},\ldots),\mathbf{X}_{\sigma(\ell+2)},\ldots)\,.
\end{aligned}$$

1.7.3 Frölicher-Nijenhuis bracket between one forms

Let us now consider the special case of tangent-valued one-forms.

Tangent-valued 1-forms $\mathbf{K} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ are in one-to-one correspondence with the linear maps $\hat{\mathbf{K}} \in BL(\mathbb{TM}\,;\mathbb{TM})$ and $\hat{\mathbf{K}}^* \in BL(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ defined by [104]:

$$\hat{\mathbf{K}}(\mathbf{v}) := \mathbf{K}\mathbf{v}\,, \qquad \forall\, \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})\,,$$

$$\hat{\mathbf{K}}^*(\mathbf{v}^*) := \mathbf{i}_{\mathbf{K}}\mathbf{v}^*\,, \quad \forall\, \mathbf{v}^* \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})\,.$$

Given a tangent-valued one-form $\mathbf{K} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ and a tangent-valued zero-form (vector field) $\mathbf{u} \in \Lambda^0(\mathbb{M}\,;\mathbb{TM}) = \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, their FN-bracket is the tangent-valued one-form $[\mathbf{K}, \mathbf{u}\,, \in]\Lambda^1(\mathbb{M}\,;\mathbb{TM})$ which, for any vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, is defined by

$$[\mathbf{K}\,, \mathbf{u}]_{\mathrm{FN}}(\mathbf{v}) := [\mathbf{K}\mathbf{v}\,, \mathbf{u}] + \mathbf{K} \cdot [\mathbf{u}\,, \mathbf{v}]\,.$$

The FN-bracket between the tangent-valued one-forms $\mathbf{K}, \mathbf{L} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ is the tangent-valued two-form $[\mathbf{K}\,, \mathbf{L}]_{\mathrm{FN}} \in \Lambda^2(\mathbb{M}\,;\mathbb{TM})$ which, for any pair of vector fields $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$, is given by

$$\begin{aligned}[\mathbf{K}\,, \mathbf{L}]_{\mathrm{FN}}(\mathbf{u}, \mathbf{v}) := {} & [\mathbf{K}\mathbf{u}\,, \mathbf{L}\mathbf{v}] - [\mathbf{K}\mathbf{v}\,, \mathbf{L}\mathbf{u}] \\ & - \mathbf{L} \cdot ([\mathbf{K}\mathbf{u}\,, \mathbf{v}] - [\mathbf{K}\mathbf{v}\,, \mathbf{u}]) \\ & - \mathbf{K} \cdot ([\mathbf{L}\mathbf{u}\,, \mathbf{v}] - [\mathbf{L}\mathbf{v}\,, \mathbf{u}]) \\ & + (\mathbf{K} \circ \mathbf{L} + \mathbf{L} \circ \mathbf{K}) \cdot [\mathbf{u}\,, \mathbf{v}]\,.\end{aligned}$$

We remark that, although the point values of each term at the r.h.s. of the previous formulae depend on the choice of the vector fields $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, the l.h.s., is tensorial. Indeed the Leibniz rule for the Lie derivative yields

$$\begin{aligned}[\mathbf{K}\mathbf{u}\,, \mathbf{L}\mathbf{v}] &= [\mathbf{K}\mathbf{u}\,, \mathbf{L}] \cdot \mathbf{v} + \mathbf{L} \cdot [\mathbf{K}\mathbf{u}\,, \mathbf{v}]\,, \\ [\mathbf{K}\mathbf{u}\,, \mathbf{v}] &= [\mathbf{u}\,, \mathbf{K}]\mathbf{v} + \mathbf{K} \cdot [\mathbf{u}\,, \mathbf{v}]\,,\end{aligned}$$

and the previous formula may be written as

$$[\mathbf{K}\,, \mathbf{L}]_{\mathrm{FN}}(\mathbf{u}, \mathbf{v}) := [\mathbf{K}\mathbf{u}\,, \mathbf{L}] \cdot \mathbf{v} + [\mathbf{L}\mathbf{u}\,, \mathbf{K}] \cdot \mathbf{v} - \mathbf{L} \cdot [\mathbf{u}\,, \mathbf{K}] \cdot \mathbf{v} - \mathbf{K} \cdot [\mathbf{u}\,, \mathbf{L}] \cdot \mathbf{v}\,,$$

which shows the tensoriality with respect to $\mathbf{v}$. A symmetric argument yields tensoriality with respect to $\mathbf{u}$. From the previous formula, for $\mathbf{L} = \mathbf{K}$ we get:

$$\begin{aligned}[\mathbf{K}\,, \mathbf{K}]_{\mathrm{FN}}(\mathbf{u}, \mathbf{v}) := {} & [\mathbf{K}\mathbf{u}\,, \mathbf{K}\mathbf{v}] - [\mathbf{K}\mathbf{v}\,, \mathbf{K}\mathbf{u}] \\ & - \mathbf{K} \cdot ([\mathbf{K}\mathbf{u}\,, \mathbf{v}] - [\mathbf{K}\mathbf{v}\,, \mathbf{u}]) \\ & - \mathbf{K} \cdot ([\mathbf{K}\mathbf{u}\,, \mathbf{v}] - [\mathbf{K}\mathbf{v}\,, \mathbf{u}]) \\ & + (\mathbf{K} \circ \mathbf{K} + \mathbf{K} \circ \mathbf{K}) \cdot [\mathbf{u}\,, \mathbf{v}]\,,\end{aligned}$$

and, grouping:

$$\tfrac{1}{2}\,[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}}(\mathbf{u},\mathbf{v}) := [\mathbf{K}\mathbf{u}\,,\mathbf{K}\mathbf{v}] - \mathbf{K}\cdot([\mathbf{K}\mathbf{u}\,,\mathbf{v}] - [\mathbf{K}\mathbf{v}\,,\mathbf{u}]) + (\mathbf{K}\circ\mathbf{K})\cdot[\mathbf{u}\,,\mathbf{v}]\,.$$

Lemma 1.7.3 *If* $\mathbf{K}\in\Lambda^1(\mathbb{M}\,;\mathbb{TM})$ *is idempotent, that is* $\mathbf{K}\circ\mathbf{K}=\mathbf{K}$*, then*

$$\begin{aligned}[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}} &= [\mathbf{I}-\mathbf{K}\,,\mathbf{I}-\mathbf{K}]_{\mathrm{FN}}\\ &= (\mathbf{I}-\mathbf{K})\cdot[\mathbf{K}\mathbf{u}\,,\mathbf{K}\mathbf{v}] + \mathbf{K}\cdot[(\mathbf{I}-\mathbf{K})\mathbf{u}\,,(\mathbf{I}-\mathbf{K})\mathbf{v}]\,.\end{aligned}$$

Proof. From the defining formula, rearranging:

$$\begin{aligned}\tfrac{1}{2}\,[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}}(\mathbf{u},\mathbf{v}) &:= [\mathbf{K}\mathbf{u}\,,\mathbf{K}\mathbf{v}] - \mathbf{K}\cdot([\mathbf{K}\mathbf{u}\,,\mathbf{v}] - [\mathbf{K}\mathbf{v}\,,\mathbf{u}]) + \mathbf{K}\cdot[\mathbf{u}\,,\mathbf{v}]\\ &= [\mathbf{K}\mathbf{u}\,,\mathbf{K}\mathbf{v}] - \mathbf{K}\cdot[\mathbf{K}\mathbf{u}\,,\mathbf{K}\mathbf{v}]\\ &\quad + \mathbf{K}\cdot[\mathbf{K}\mathbf{u}\,,\mathbf{K}\mathbf{v}] - \mathbf{K}\cdot([\mathbf{K}\mathbf{u}\,,\mathbf{v}] - [\mathbf{K}\mathbf{v}\,,\mathbf{u}]) + \mathbf{K}\cdot[\mathbf{u}\,,\mathbf{v}]\\ &= (\mathbf{I}-\mathbf{K})\cdot[\mathbf{K}\mathbf{u}\,,\mathbf{K}\mathbf{v}] + \mathbf{K}\cdot[(\mathbf{I}-\mathbf{K})\mathbf{u}\,,(\mathbf{I}-\mathbf{K})\mathbf{v}]\,,\end{aligned}$$

and the result follows. ■

Lemma 1.7.4 *Given* $\mathbf{L}\in\Lambda^1(\mathbb{M}\,;\mathbb{TM})$ *and* $\mathbf{X}\in\Lambda^0(\mathbb{M}\,;\mathbb{TM})$ *we have that*

$$\begin{aligned} &i) && [\mathbf{i_L}\,,\mathbf{i_X}] = -\mathbf{i_{LX}}\,,\\ &ii) && [\mathbf{i_L}\,,\mathcal{L}_\mathbf{X}] = \mathbf{i}_{[\mathbf{L}\,,\mathbf{X}]_{\mathrm{FN}}}\,,\\ &iii) && [\mathbf{i_X}\,,\mathcal{L}_\mathbf{L}] = \mathcal{L}_{\mathbf{LX}} + \mathbf{i}_{[\mathbf{L}\,,\mathbf{X}]_{\mathrm{FN}}}\,,\\ &iv) && [\mathbf{i_L}\,,\mathcal{L}_\mathbf{L}] = \mathcal{L}_{\mathbf{L}\circ\mathbf{L}} - \mathbf{i}_{[\mathbf{L}\,,\mathbf{L}]_{\mathrm{FN}}}\,.\end{aligned}$$

Proof. To get $i)$ we remark that the derivation $[\mathbf{i_L}\,,\mathbf{i_X}]$ is algebraic so that it is enough to compute it on a one-form $\boldsymbol{\omega}\in\Lambda^1(\mathbb{M}\,;\Re)$:

$$[\mathbf{i_L}\,,\mathbf{i_X}]\boldsymbol{\omega} = (\mathbf{i_L}\circ\mathbf{i_X})\boldsymbol{\omega} - (\mathbf{i_X}\circ\mathbf{i_L})\boldsymbol{\omega} = -(\mathbf{i_X}\circ\mathbf{i_L})\boldsymbol{\omega} = -\mathbf{i_{LX}}\boldsymbol{\omega}\,.$$

Indeed $(\mathbf{i_L}\circ\mathbf{i_X})\boldsymbol{\omega} = \mathbf{i_L}(\boldsymbol{\omega}\cdot\mathbf{X}) = 0$ since $\boldsymbol{\omega}\cdot\mathbf{X}$ is a 0-form. To get $ii)$ we recall Lemma 1.7.2 to write: $[\mathbf{i_L}\,,\mathcal{L}_\mathbf{X}] = \mathcal{L}(\mathbf{i_L}\mathbf{X}) + \mathbf{i}_{[\mathbf{L}\,,\mathbf{X}]_{\mathrm{FN}}}$ and observe that $\mathbf{i_L}\mathbf{X} = \mathbf{i_L}(1\otimes\mathbf{X}) = (\mathbf{i_L}1)\otimes\mathbf{X} = 0$. Formulae $iii)$ and $iv)$ follow again by Lemma 1.7.2 being

$$\begin{aligned}[\mathbf{i_X}\,,\mathcal{L}_\mathbf{L}] &= \mathcal{L}(\mathbf{i_X}\mathbf{L}) - \mathbf{i}_{[\mathbf{X}\,,\mathbf{L}]_{\mathrm{FN}}} = \mathcal{L}_{\mathbf{LX}} + \mathbf{i}_{[\mathbf{L}\,,\mathbf{X}]_{\mathrm{FN}}}\,,\\ [\mathbf{i_L}\,,\mathcal{L}_\mathbf{L}] &= \mathcal{L}(\mathbf{i_L}\mathbf{L}) - \mathbf{i}_{[\mathbf{L}\,,\mathbf{L}]_{\mathrm{FN}}} = \mathcal{L}_{\mathbf{L}\circ\mathbf{L}} - \mathbf{i}_{[\mathbf{L}\,,\mathbf{L}]_{\mathrm{FN}}}\,,\end{aligned}$$

since $[\mathbf{X}\,,\mathbf{L}]_{\mathrm{FN}} = -[\mathbf{L}\,,\mathbf{X}]_{\mathrm{FN}}$ and $\mathbf{i_L}\mathbf{L} = \mathbf{L}\circ\mathbf{L}$. ■

Lemma 1.7.5 (Naturality of the FN-bracket) *The push-forward of a tangent-valued form* $\mathbf{K} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ *, according to a morphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *, is defined by*

$$\boldsymbol{\varphi}{\uparrow}\mathbf{K} \cdot \boldsymbol{\varphi}{\uparrow}\mathbf{X} := \boldsymbol{\varphi}{\uparrow}(\mathbf{K} \cdot \mathbf{X})\,,$$

where $\mathbf{X} \in \Lambda^0(\mathbb{M}\,;\mathbb{TM})$ *, and similarly for higher degree tangent-valued forms. Then the* FN*-brackets of two tangent-valued forms* $\mathbf{K}, \mathbf{L} \in \Lambda(\mathbb{M}\,;\mathbb{TM})$ *is natural with repect to the push, i.e.:*

$$\boldsymbol{\varphi}{\uparrow}[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}} = [\boldsymbol{\varphi}{\uparrow}\mathbf{K}\,,\boldsymbol{\varphi}{\uparrow}\mathbf{L}]_{\mathrm{FN}}\,.$$

Proof. We have that $\mathbf{i}_{\boldsymbol{\varphi}\uparrow\mathbf{K}} = \boldsymbol{\varphi}{\uparrow}\mathbf{i}_{\mathbf{K}}$ and $\boldsymbol{\varphi}{\uparrow} \circ d = d \circ \boldsymbol{\varphi}{\uparrow}$. Then

$$\begin{aligned}\boldsymbol{\varphi}{\uparrow}\mathcal{L}_{\mathbf{K}} &= \boldsymbol{\varphi}{\uparrow}[\mathbf{i}_{\mathbf{K}}\,,d] = \boldsymbol{\varphi}{\uparrow}(\mathbf{i}_{\mathbf{K}} \circ d - (-1)^{\deg \mathbf{K}} d \circ \mathbf{i}_{\mathbf{K}} \\ &= \boldsymbol{\varphi}{\uparrow}\mathbf{i}_{\mathbf{K}} \circ d - (-1)^{\deg \mathbf{K}} \boldsymbol{\varphi}{\uparrow} \circ d \circ \mathbf{i}_{\mathbf{K}} = \boldsymbol{\varphi}{\uparrow}\mathbf{i}_{\mathbf{K}} \circ d - (-1)^{\deg \mathbf{K}} d \circ \boldsymbol{\varphi}{\uparrow}\mathbf{i}_{\mathbf{K}} \\ &= [\boldsymbol{\varphi}{\uparrow}\mathbf{i}_{\mathbf{K}}\,,d] = \mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{K}}\,,\end{aligned}$$

and also $\boldsymbol{\varphi}{\uparrow}[\mathcal{L}_{\mathbf{K}}\,,\mathcal{L}_{\mathbf{L}}] = [\mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{K}}\,,\mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{L}}]$ so that

$$\mathcal{L}_{\boldsymbol{\varphi}\uparrow[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}}} = \boldsymbol{\varphi}{\uparrow}\mathcal{L}_{[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}}} = \boldsymbol{\varphi}{\uparrow}[\mathcal{L}_{\mathbf{K}}\,,\mathcal{L}_{\mathbf{L}}] = [\mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{K}}\,,\mathcal{L}_{\boldsymbol{\varphi}\uparrow\mathbf{L}}] = \mathcal{L}_{[\boldsymbol{\varphi}\uparrow\mathbf{K}\,,\boldsymbol{\varphi}\uparrow\mathbf{L}]_{\mathrm{FN}}}\,.$$

and the result follows. ■

Definition 1.7.3 *Given* $\mathbf{K}, \mathbf{L} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ *, the* Nijenhuis *differential is defined by*

$$d_{\mathbf{K}}\mathbf{L} := [\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}}\,.$$

Note that $d_{\mathbf{K}}\mathbf{I} = 0$. By Jacobi identity, the Nijenhuis differential is a graded derivation on the FN-algebra, i,e:

$$d_{\mathbf{K}}[\mathbf{L}\,,\mathbf{M}]_{\mathrm{FN}} = [d_{\mathbf{K}}\mathbf{L}\,,\mathbf{M}]_{\mathrm{FN}} + (-1)^{\deg \mathbf{K} \deg \mathbf{L}}[\mathbf{L}\,,d_{\mathbf{K}}\mathbf{M}]_{\mathrm{FN}}\,.$$

Rewriting Jacobi identity, we also infer that:

$$d_{[\mathbf{K}\,,\mathbf{L}]_{\mathrm{FN}}} = d_{\mathbf{K}} d_{\mathbf{L}} - (-1)^{\deg \mathbf{K} \deg \mathbf{L}} d_{\mathbf{L}} d_{\mathbf{K}}\,.$$

As a special case we get

$$d_{[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}}} = (1 - (-1)^{\deg \mathbf{K}})\, d_{\mathbf{K}} \circ d_{\mathbf{K}}\,.$$

If $\deg \mathbf{K}$ is even, by the graded anticommutativity of the FN-bracket we have: $[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}} = -[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}} = 0$ and, by the previous formula, also $d_{[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}}} = 0$. If $\deg \mathbf{K}$ is odd, we have the identities

$$\begin{aligned} d_{[\mathbf{K}\,,\mathbf{K}]_{\mathrm{FN}}} &= 2\, d_{\mathbf{K}} \circ d_{\mathbf{K}}\,, \\ d_{\mathbf{K}} d_{\mathbf{K}} \mathbf{K} &= 0\,, \qquad \text{second BIANCHI identity} \\ [d_{\mathbf{K}}\mathbf{K}\,, d_{\mathbf{K}}\mathbf{K}]_{\mathrm{FN}} &= 0\,. \end{aligned}$$

1.7.4 Curvature and soldering form

Definition 1.7.4 (Horizontal form) *In a fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$, a form $\mathbf{K} \in \Lambda^k(\mathbb{E}\,;\mathbb{VE})$ is said horizontal if it vanishes when any of its arguments is a vertical tangent vector to $\mathbb{TE}$. This concept is independent of the choice of a connection. Vertical-valued horizontal forms are also called semi-basic (in French semi-basique [79]).*

In terms of the FN-bracket, the curvature of a connection is the vertical-valued horizontal 2-form:

$$\mathbf{R} := -\tfrac{1}{2}[P_{\mathrm{V}}\,, P_{\mathrm{V}}]_{\mathrm{FN}} = -\tfrac{1}{2}[P_{\mathrm{H}}\,, P_{\mathrm{H}}]_{\mathrm{FN}}\,.$$

Indeed, from Lemma 1.7.3, being $P_{\mathrm{H}} \cdot [P_{\mathrm{V}}\mathbf{X}\,, P_{\mathrm{V}}\mathbf{Y}] = 0$ by integrability of the vertical distribution, we get:

$$-\mathbf{R}(\mathbf{X}, \mathbf{Y}) = P_{\mathrm{H}} \cdot [P_{\mathrm{V}}\mathbf{X}\,, P_{\mathrm{V}}\mathbf{Y}] + P_{\mathrm{V}} \cdot [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}] = P_{\mathrm{V}} \cdot [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}]\,,$$

which is the formula introduced in Section 1.4.6.

Definition 1.7.5 (Soldering form) *In a fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ a soldering form is a vertical-valued horizontal 1-form: $\boldsymbol{\sigma} \in \Lambda^1(\mathbb{E}\,;\mathbb{VE})$.*

From the definition it follows that soldering forms are nilpotent: $\boldsymbol{\sigma} \circ \boldsymbol{\sigma} = 0$. The following simple lemma is useful.

Lemma 1.7.6 *In a fibre bundle $\mathbf{p} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{M})$ the LIE-bracket of a pair of tangent-valued 0-forms $\mathbf{V}, \mathbf{X} \in \Lambda^0(\mathbb{E}\,;\mathbb{TE})$, one vertical-valued and the other one projectable, is a vertical-valued 0-form $[\mathbf{V}\,, \mathbf{X}] \in \Lambda^0(\mathbb{E}\,;\mathbb{VE})$. The LIE-derivative of a vertical-valued 1-form $\mathbf{K} \in \Lambda^1(\mathbb{E}\,;\mathbb{TE})$, along a projectable tangent-valued 0-form $\mathbf{X} \in \Lambda^0(\mathbb{E}\,;\mathbb{TE})$ is a vertical-valued 1-form $\mathcal{L}_{\mathbf{X}}\mathbf{K} \in \Lambda^1(\mathbb{E}\,;\mathbb{VE})$.*

Proof. Let $\mathbf{V}, \mathbf{Y} \in \Lambda^0(\mathbb{E}\,;\mathbb{TE})$ with $\mathbf{V}$ vertical and $\mathbf{Y}$ projectable. Then by naturality of the Lie-bracket of projectable vector fields with respect to relatedness we have:

$$T\mathbf{p} \cdot \mathbf{V} = 0 \implies T\mathbf{p} \cdot [\mathbf{V}\,, \mathbf{Y}] = [T\mathbf{p} \cdot \mathbf{V}\,, T\mathbf{p} \cdot \mathbf{Y}] = [0\,, T\mathbf{p} \cdot \mathbf{Y}] = 0\,.$$

Being $T\mathbf{p} \cdot \mathbf{K} = 0$, the second statement follows then from Leibniz formula:

$$\mathcal{L}_{\mathbf{X}}\mathbf{K} \cdot \mathbf{Y} = [\mathbf{X}\,, \mathbf{KY}] + \mathbf{K} \cdot [\mathbf{X}\,, \mathbf{Y}]\,,$$

since: $T\mathbf{p} \cdot \mathcal{L}_{\mathbf{X}}\mathbf{K} \cdot \mathbf{Y} = T\mathbf{p} \cdot [\mathbf{X}\,, \mathbf{KY}] + T\mathbf{p} \cdot \mathbf{K} \cdot [\mathbf{X}\,, \mathbf{Y}] = 0\,.$ ■

Definition 1.7.6 (Canonical soldering form) *In the tangent bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$*, the canonical soldering form* $\mathbf{J} \in \Lambda^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *is the vertical-valued horizontal one-form defined by*

$$\mathbf{J} := \mathbf{Vl}_{\mathbb{TM}} \circ (\boldsymbol{\tau}_{\mathbb{TM}}\,, T\boldsymbol{\tau}_{\mathbb{M}})\,,$$

where $(\boldsymbol{\tau}_{\mathbb{TM}}\,, T\boldsymbol{\tau}_{\mathbb{M}}) \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM})$ *and* $\mathbf{Vl}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{TM}\times_{\mathbb{M}}\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$*.*

Since the vertical lift $\mathbf{Vl}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{TM} \times_{\mathbb{M}} \mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ defines at any $\mathbf{v} \in \mathbb{TM}$ a linear isomorphism $\mathbf{Vl}_{\mathbb{TM}}(\mathbf{v}) \in BL\,(\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}\,;\mathbb{V}_{\mathbf{v}}\mathbb{TM})$ it follows that $\mathbf{J}(\mathbf{v}) \in BL\,(\mathbb{T}_{\mathbf{v}}\mathbb{TM}\,;\mathbb{T}_{\mathbf{v}}\mathbb{TM})$ with

$$\mathbf{ker}(\mathbf{J}(\mathbf{v})) = \mathbf{im}(\mathbf{J}(\mathbf{v})) = \mathbb{V}_{\mathbf{v}}\mathbb{TM} = \mathbb{T}_{\mathbf{v}}\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v})}\mathbb{M}\,.$$

Denoting by $\mathbf{J}^* \in \mathrm{C}^1(\mathbb{TM}\,;BL\,(\mathbb{T}^*\mathbb{TM}\,;\mathbb{T}^*\mathbb{TM}))$ the dual tensor field defined by

$$\langle\, \mathbf{Y}^*(\mathbf{v}), \mathbf{J}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v})\,\rangle = \langle\, \mathbf{J}^*(\mathbf{v}) \cdot \mathbf{Y}^*(\mathbf{v}), \mathbf{X}(\mathbf{v})\,\rangle\,, \quad \forall \begin{cases} \mathbf{X}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{TM} \\ \mathbf{Y}^*(\mathbf{v}) \in \mathbb{T}^*_{\mathbf{v}}\mathbb{TM} \end{cases}$$

we have that $\mathbf{im}(\mathbf{J}^*(\mathbf{v})) = \mathbf{ker}(\mathbf{J}(\mathbf{v}))^{\circ}$ and $\mathbf{ker}(\mathbf{J}^*(\mathbf{v})) = \mathbf{im}(\mathbf{J}(\mathbf{v}))^{\circ}$. Hence the tensor $\mathbf{J}^*(\mathbf{v}) \in BL\,(\mathbb{T}^*_{\mathbf{v}}\mathbb{TM}\,;\mathbb{T}^*_{\mathbf{v}}\mathbb{TM})$ vanishes exactly on forms vanishing on vertical vectors and takes values exactly on forms vanishing on vertical vectors. Then we have the simple consequence:

Proposition 1.7.2 *The canonical soldering* $\mathbf{J} \in \mathrm{C}^1(\mathbb{TM}\,;BL\,(\mathbb{T}^2\mathbb{M}\,;\mathbb{T}^2\mathbb{M}))$ *and its dual form* $\mathbf{J}^* \in \mathrm{C}^1(\mathbb{TM}\,;BL\,(\mathbb{T}^*\mathbb{TM}\,;\mathbb{T}^*\mathbb{TM}))$ *are nilpotent:*

$$\mathbf{J} \circ \mathbf{J} = 0\,, \quad \mathbf{J}^* \circ \mathbf{J}^* = 0\,.$$

Proposition 1.7.3 (Sprays) *A spray* $\mathbf{S} \in \Lambda^0(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *is characterized by the equivalent conditions:*

$$(\boldsymbol{\tau}_{\mathbb{TM}}\,, T\boldsymbol{\tau}_{\mathbb{M}}) \circ \mathbf{S} = \mathrm{DIAG} \iff \mathbf{J} \cdot \mathbf{S} = \mathbf{C}\,.$$

Proof. By the injectivity of the vertical lift $\mathbf{Vl}_{\mathbb{TM}} \in \mathrm{C}^1(\mathbb{TM} \times_{\mathbb{M}} \mathbb{TM} ; \mathbb{T}^2\mathbb{M})$ and by definition of the Liouville vector field $\mathbf{C} \in \mathrm{C}^1(\mathbb{TM} ; \mathbb{T}^2\mathbb{M})$, the equality $\mathbf{J} \cdot \mathbf{S} = \mathbf{Vl}_{\mathbb{TM}} \circ (\boldsymbol{\tau}_{\mathbb{TM}} , T\boldsymbol{\tau}_{\mathbb{M}}) \circ \mathbf{S} = \mathbf{Vl}_{\mathbb{TM}} \circ \mathrm{DIAG} = \mathbf{C}$ is equivalent to $\boldsymbol{\tau}_{\mathbb{TM}} \circ \mathbf{S} = T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{S}$. ■

A spray can be canonically associated with a connection on the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM} ; \mathbb{M})$ by taking the horizontal projection of any spray $\overline{\mathbf{S}}$. Indeed the spray $\mathbf{S} = P_{\mathrm{H}} \circ \overline{\mathbf{S}}$ is independent of the choice of the spray $\overline{\mathbf{S}}$ since the difference of any two sprays is vertical.

In [66], Proposition I.37 the following formula is stated without proof:

$$\nabla_{\mathbf{z}} \mathbf{w} = \mathbf{vd}_{\mathbb{TM}} \circ ([P_{\mathrm{H}} , \mathbf{JW}]_{\mathrm{FN}} \cdot \mathbf{Z}) ,$$

with $\mathbf{W}, \mathbf{Z} \in \Lambda^0(\mathbb{TM} ; \mathbb{T}^2\mathbb{M})$ projecting on $\mathbf{w}, \mathbf{z} \in \mathrm{C}^1(\mathbb{M} ; \mathbb{TM})$ respectively. By the formula in section 1.7.3 we have that:

$$[P_{\mathrm{H}} , \mathbf{JW}]_{\mathrm{FN}} \cdot \mathbf{Z} = [P_{\mathrm{H}}\mathbf{Z} , \mathbf{JW}] - P_{\mathrm{H}} \cdot [\mathbf{JW} , \mathbf{Z}] = [P_{\mathrm{H}}\mathbf{Z} , \mathbf{JW}] ,$$

since $[\mathbf{JW} , \mathbf{Z}]$ is vertical by Lemma 1.7.6. Moreover, $\mathbf{JW} = \mathbf{J}(P_{\mathrm{H}}\mathbf{W})$ and $P_{\mathrm{H}}\mathbf{Z} = \mathbf{H}_{\mathbf{z}}$, $P_{\mathrm{H}}\mathbf{W} = \mathbf{H}_{\mathbf{w}}$ by Lemma 1.4.5. The formula may then be rewritten as

$$\nabla_{\mathbf{z}} \mathbf{w} = \mathbf{vd}_{\mathbb{TM}} \circ [\mathbf{H}_{\mathbf{z}} , \mathbf{JH}_{\mathbf{w}}] ,$$

and coincides with the one proved in Lemma 1.7.12.

In [60], Propositions X.1.5 and X.1.6 on page 160 provided the following properties on the basis of computations in coordinates:

$$\begin{aligned} &[\mathbf{JX} , \mathbf{JY}] = \mathbf{J} \cdot [\mathbf{JX} , \mathbf{Y}] + \mathbf{J} \cdot [\mathbf{X} , \mathbf{JY}] , \\ &\mathbf{JX} = [\mathbf{JX} , \mathbf{C}] + \mathbf{J} \cdot [\mathbf{C} , \mathbf{X}] , \qquad \forall\, \mathbf{X}, \mathbf{Y} \in \Lambda^0(\mathbb{TM} ; \mathbb{T}^2\mathbb{M}) . \end{aligned}$$

so that

$$\begin{aligned} &[\mathbf{J} , \mathbf{J}]_{\mathrm{FN}} = 0 \iff [\mathcal{L}_{\mathbf{J}} , \mathcal{L}_{\mathbf{J}}] = 0 \iff \mathcal{L}_{\mathbf{J}} \circ \mathcal{L}_{\mathbf{J}} = 0 , \\ &[\mathbf{J} , \mathbf{C}]_{\mathrm{FN}} = \mathbf{J} . \end{aligned}$$

Related properties of the canonical soldering form have been investigated in [66] and referred to in [79], [120] and [68]. From the properties of the canonical soldering form we get the following formulas.

Lemma 1.7.7 *The canonical soldering form* $\mathbf{J} \in \Lambda^1(\mathbb{TM} ; \mathbb{T}^2\mathbb{M})$ *, the Liouville canonical field* $\mathbf{C} \in \Lambda^0(\mathbb{TM} ; \mathbb{T}^2\mathbb{M})$ *and a spray* $\mathbf{S} \in \Lambda^0(\mathbb{TM} ; \mathbb{T}^2\mathbb{M})$ *fulfil the* *properties:*

$$
\begin{aligned}
&i) \quad [\mathbf{i_J}\,,\mathbf{i_C}] \;= -\mathbf{i_{JC}} = 0\\
&ii) \quad [\mathbf{i_J}\,,\mathcal{L}_{\mathbf{C}}] = \mathbf{i}_{[\mathbf{J}\,,\mathbf{C}]_{\mathrm{FN}}}\\
&iii) \quad [\mathbf{i_C}\,,\mathcal{L}_{\mathbf{J}}] = \mathcal{L}_{\mathbf{JC}} + \mathbf{i}_{[\mathbf{J}\,,\mathbf{C}]_{\mathrm{FN}}} = \mathbf{i}_{[\mathbf{J}\,,\mathbf{C}]_{\mathrm{FN}}}\\
&iv) \quad [\mathbf{i_J}\,,\mathcal{L}_{\mathbf{J}}] = \mathcal{L}_{\mathbf{J}\circ\mathbf{J}} - \mathbf{i}_{[\mathbf{J}\,,\mathbf{J}]_{\mathrm{FN}}} = -\mathbf{i}_{[\mathbf{J}\,,\mathbf{J}]_{\mathrm{FN}}}\\
&v) \quad [\mathbf{i_S}\,,\mathcal{L}_{\mathbf{J}}] = \mathcal{L}_{\mathbf{JS}} + \mathbf{i}_{[\mathbf{J}\,,\mathbf{S}]_{\mathrm{FN}}} = \mathcal{L}_{\mathbf{C}} + \mathbf{i}_{[\mathbf{J}\,,\mathbf{S}]_{\mathrm{FN}}}\,.
\end{aligned}
$$

Proof. By the verticality of the LIOUVILLE canonical field, it is:

$$
\begin{aligned}
&\mathbf{J}\cdot\mathbf{C} = 0\,,\\
&[\mathbf{J}\,,\mathbf{C}]\cdot\mathbf{X} = [\mathbf{JX}\,,\mathbf{C}] + \mathbf{J}\cdot[\mathbf{C}\,,\mathbf{X}] = [\mathbf{JX}\,,\mathbf{C}]\,,
\end{aligned}
$$

and the results follow from Lemma 1.7.4. ∎

1.7.5 Generalized connections

A *generalized connection* on the tangent bundle $\boldsymbol{\pi} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ is an idempotent tangent-valued one-form $P_{\mathrm{V}} \in \Lambda^1(\mathbb{M}\,;\mathbb{TM})$ so that $P_{\mathrm{V}} \circ P_{\mathrm{V}} = P_{\mathrm{V}}$. Then a connection is pointwise a linear projection. Its image $\mathbb{VM} := \mathbf{im}(P_{\mathrm{V}})$ is called the *vertical* subspace and its kernel $\mathbb{HM} := \mathbf{ker}(P_{\mathrm{V}})$ is called the *horizontal* subspace.

A tangent-valued form $\boldsymbol{\omega} \in \Lambda^k(\mathbb{M}\,;\mathbb{TM})$ is said to be *horizontal* if it vanishes if any of its arguments is a vertical vector. If a connection is defined on the tangent bundle $\boldsymbol{\pi} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$, a tangent-valued form $\boldsymbol{\omega} \in \Lambda^k(\mathbb{M}\,;\mathbb{TM})$ is said to be *vertical* if it vanishes on horizontal vectors.

Setting $P_{\mathrm{H}} = \mathbf{id}_{\mathbb{TM}} - P_{\mathrm{V}}$, by the formula above for the FN tangent-valued two-form, we get:

$$
-\tfrac{1}{2}[P_{\mathrm{V}}\,,P_{\mathrm{V}}]_{\mathrm{FN}} = -\tfrac{1}{2}[P_{\mathrm{H}}\,,P_{\mathrm{H}}]_{\mathrm{FN}} = \mathbf{R} + \mathbf{R^c}\,,
$$

with the tangent-valued 2-forms $\mathbf{R}, \mathbf{R^c} \in \Lambda^2(\mathbb{M}\,;\mathbb{TM})$, respectively called the *curvature* and the *cocurvature* of the connection, given by

$$
\begin{aligned}
\mathbf{R}(\mathbf{u},\mathbf{v}) &:= -P_{\mathrm{V}}\cdot[P_{\mathrm{H}}\mathbf{u}\,,P_{\mathrm{H}}\mathbf{v}]\,,\\
\mathbf{R^c}(\mathbf{u},\mathbf{v}) &:= -P_{\mathrm{H}}\cdot[P_{\mathrm{V}}\mathbf{u}\,,P_{\mathrm{V}}\mathbf{v}]\,, \quad \forall\,\mathbf{u},\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})\,,
\end{aligned}
$$

so that

$$
\begin{aligned}
2\,\mathbf{R} &= -P_{\mathrm{V}}\circ[P_{\mathrm{V}}\,,P_{\mathrm{V}}]_{\mathrm{FN}} = -P_{\mathrm{V}}\circ[P_{\mathrm{H}}\,,P_{\mathrm{H}}]_{\mathrm{FN}}\,,\\
2\,\mathbf{R^c} &= -P_{\mathrm{H}}\circ[P_{\mathrm{H}}\,,P_{\mathrm{H}}]_{\mathrm{FN}} = -P_{\mathrm{H}}\circ[P_{\mathrm{V}}\,,P_{\mathrm{V}}]_{\mathrm{FN}}\,.
\end{aligned}
$$

Then the curvature $\mathbf{R} \in \Lambda^2(\mathbb{M}\,;\mathbb{VM})$ is a vertical-valued horizontal 1-form, while the cocurvature $\mathbf{R^c} \in \Lambda^2(\mathbb{M}\,;\mathbb{HM})$ is a horizontal-valued vertical 1-form.

Tensoriality of $\mathbf{R}, \mathbf{R^c} \in \Lambda^2(\mathbb{M}\,;\mathbb{TM})$ is easily inferred from Lemma 1.2.1 and the property of complementary projections: $P_{\rm V} \circ P_{\rm H} = 0$ and $P_{\rm H} \circ P_{\rm V} = 0$.

By Frobenius theorem, the *curvature* and the *cocurvature* are respectively obstructions against integrability of the horizontal and the vertical subbundle. The graded Jacobi identity implies that

$$[P_{\rm V}\,, [P_{\rm V}\,, P_{\rm V}]_{\rm FN}]_{\rm FN} = 0\,,$$

and this result is known as the (generalized second) Bianchi identity for the connection [80]. Moreover we have that

$$[\mathbf{R}\,, P_{\rm V}]_{\rm FN} = \mathbf{i}_{\mathbf{R}}\mathbf{R^c} + \mathbf{i}_{\mathbf{R^c}}\mathbf{R}\,.$$

Indeed

$$-2\,\mathbf{R} = P_{\rm V} \circ [P_{\rm V}\,, P_{\rm V}]_{\rm FN} = \mathbf{I}_{[P_{\rm V}\,,P_{\rm V}]_{\rm FN}} P_{\rm V}\,,$$

and from Lemma 1.7.2 (see [80] Theorem 8.11 (2)):

$$\mathbf{i}_{[P_{\rm V}\,,P_{\rm V}]_{\rm FN}}[P_{\rm V}\,, P_{\rm V}]_{\rm FN} = 2\,[\mathbf{i}_{[P_{\rm V}\,,P_{\rm V}]_{\rm FN}} P_{\rm V}\,, P_{\rm V}]_{\rm FN} = 4\,[\mathbf{R}\,, P_{\rm V}]_{\rm FN}\,.$$

Therefore

$$[\mathbf{R}\,, P_{\rm V}]_{\rm FN} = \tfrac{1}{4}\mathbf{i}_{[P_{\rm V}\,,P_{\rm V}]_{\rm FN}}[P_{\rm V}\,, P_{\rm V}]_{\rm FN} = \mathbf{i}_{\mathbf{R}+\mathbf{R^c}}(\mathbf{R} + \mathbf{R^c}) = \mathbf{i}_{\mathbf{R}}\mathbf{R^c} + \mathbf{i}_{\mathbf{R^c}}\mathbf{R}\,,$$

since $\mathbf{i}_{\mathbf{R}}\mathbf{R} = 0$ and $\mathbf{i}_{\mathbf{R^c}}\mathbf{R^c} = 0$.

Lemma 1.7.8 (Characterization of a connection) *A connection on the tangent bundle of a manifold* $\mathbb{M}$ *is characterized by a tangent-valued one-form* $\mathbf{\Gamma} \in \Lambda^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *fulfilling the properties*

$$\begin{cases} \mathbf{J} \cdot \mathbf{\Gamma} = \mathbf{J}\,, \\ \mathbf{\Gamma} \cdot \mathbf{J} = -\mathbf{J}\,. \end{cases}$$

By the theory of the Frölicher-Nijenhuis bracket developed in section 1.7 on page 166, it can be proved that to any spray $\mathbf{S}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{TM}$ there corresponds a connection $\mathbf{\Gamma} \in \Lambda^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ given by

$$\mathbf{\Gamma} = [\mathbf{J}\,, \mathbf{S}]\,,$$

and that the spray canonically associated with this connection is given by [66], [67], [68]:

$$\mathbf{S} + \tfrac{1}{2}([\mathbf{C}\,, \mathbf{S}] - \mathbf{S}) = \tfrac{1}{2}([\mathbf{C}\,, \mathbf{S}] + \mathbf{S})\,.$$

1.7.6 Generalized torsion of a connection

Let $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ be a fibre bundle, $\mathbb{V}\mathbb{E} = \mathbf{ker}(T\mathbf{p})$ the corresponding vertical bundle and $P_V \in \Lambda^1(\mathbb{E};\mathbb{T}\mathbb{E})$ the projector on the vertical bundle defining the connection. By FROBENIUS theorem, the integrability of the vertical bundle ensures that

$$\mathbf{R^c}(\mathbf{X},\mathbf{Y}) := P_H \cdot [P_V\mathbf{X}, P_V\mathbf{Y}] = 0\,, \quad \forall\, \mathbf{X},\mathbf{Y} \in C^1(\mathbb{E};\mathbb{T}\mathbb{E})\,,$$

Then, the curvature of the connection, which is the vertical-valued horizontal 2-form defined by:

$$\mathbf{R}(\mathbf{X},\mathbf{Y}) := P_V \cdot [P_H\mathbf{X}, P_H\mathbf{Y}]\,, \quad \forall\, \mathbf{X},\mathbf{Y} \in C^1(\mathbb{E};\mathbb{T}\mathbb{E})\,,$$

may be defined as: $\mathbf{R} = \frac{1}{2}[P_V, P_V]_{FN} = \frac{1}{2}[P_H, P_H]_{FN} = \frac{1}{2}d_{P_V}P_V = \frac{1}{2}d_{P_H}P_H$. Following [121], the torsion in a fibre bundle is defined as follows.

Definition 1.7.7 *In a fibre bundle* $\mathbf{p} \in C^1(\mathbb{E};\mathbb{M})$ *the torsion* $\boldsymbol{T} \in \Lambda^2(\mathbb{E};\mathbb{V}\mathbb{E})$ *of the connection* $P_H \in \Lambda^1(\mathbb{E};\mathbb{H}\mathbb{E})$ *with respect to the soldering form* $\boldsymbol{\sigma} \in \Lambda^1(\mathbb{E};\mathbb{V}\mathbb{E})$ *is defined by the* NIJENHUIS *differential:*

$$\boldsymbol{T} := d_{\boldsymbol{\sigma}} P_H\,.$$

The next result was provided in [121] without proof (with a spurious factor $\frac{1}{2}$ in the explicit formula).

Lemma 1.7.9 *The torsion* $\boldsymbol{T} = d_{\boldsymbol{\sigma}}P_H = d_{P_H}\boldsymbol{\sigma} \in C^1(\mathbb{E};\mathbb{V}\mathbb{E})$ *of the connection* $P_H \in C^1(\mathbb{E};\mathbb{H}\mathbb{E})$ *with respect to the soldering form* $\boldsymbol{\sigma} \in \Lambda^1(\mathbb{E};\mathbb{V}\mathbb{E})$ *is the vertical-valued horizontal* 2*-form on* $\mathbb{E}$ *given by:*

$$\begin{aligned}\boldsymbol{T} &:= d_{\boldsymbol{\sigma}}P_H = d_{P_H}\boldsymbol{\sigma} = [\boldsymbol{\sigma}, P_H]_{FN} = [P_H, \boldsymbol{\sigma}]_{FN}\\ &= -d_{\boldsymbol{\sigma}}P_V = -d_{P_V}\boldsymbol{\sigma} = \tfrac{1}{2}d_{\boldsymbol{\sigma}}\boldsymbol{\Gamma} = \tfrac{1}{2}d_{\boldsymbol{\Gamma}}\boldsymbol{\sigma}\,,\end{aligned}$$

and explicitly, in terms of LIE *brackets:*

$$\boldsymbol{T}(\mathbf{H_u},\mathbf{H_v}) = [\mathbf{H_u}, \boldsymbol{\sigma}\mathbf{H_v}] - [\mathbf{H_v}, \boldsymbol{\sigma}\mathbf{H_u}] - \boldsymbol{\sigma} \cdot [\mathbf{H_u}, \mathbf{H_v}]\,,$$

for all $\mathbf{u},\mathbf{v} \in C^1(\mathbb{M};\mathbb{T}\mathbb{M})$.

Proof. The equality $[\boldsymbol{\sigma}\,, P_{\mathrm{H}}]_{\mathrm{FN}} = [P_{\mathrm{H}}\,, \boldsymbol{\sigma}]_{\mathrm{FN}}$ follows directly from the graded anticommutativity of the FN-bracket. To prove the latter formula, we recall the defining formula:

$$\begin{aligned}[P_{\mathrm{H}}\,, \boldsymbol{\sigma}]_{\mathrm{FN}}(\mathbf{X}, \mathbf{Y}) := \; & [P_{\mathrm{H}}\mathbf{X}\,, \boldsymbol{\sigma}\mathbf{Y}] - [P_{\mathrm{H}}\mathbf{Y}\,, \boldsymbol{\sigma}\mathbf{X}] \\ & - \boldsymbol{\sigma} \cdot ([P_{\mathrm{H}}\mathbf{X}\,, \mathbf{Y}] - [P_{\mathrm{H}}\mathbf{Y}\,, \mathbf{X}]) \\ & - P_{\mathrm{H}} \cdot ([\boldsymbol{\sigma}\mathbf{X}\,, \mathbf{Y}] - [\boldsymbol{\sigma}\mathbf{Y}\,, \mathbf{X}]) \\ & + (P_{\mathrm{H}} \circ \boldsymbol{\sigma} + \boldsymbol{\sigma} \circ P_{\mathrm{H}}) \cdot [\mathbf{X}\,, \mathbf{Y}]\,. \end{aligned}$$

By tensoriality of the torsion, the vector fields $\mathbf{X}, \mathbf{Y} \in \mathrm{C}^1(\mathbb{E}\,; \mathbb{TE})$ may be assumed to be projectable. Then, being $T\mathbf{p} \circ \boldsymbol{\sigma} = 0$, both brackets in the third line are vertical by Lemma 1.7.6 and hence the line vanishes. Moreover, observing that:

$$\boldsymbol{\sigma} \circ P_{\mathrm{H}} = \boldsymbol{\sigma}\,, \qquad P_{\mathrm{H}} \circ \boldsymbol{\sigma} = 0\,, \qquad \boldsymbol{\sigma} \circ \mathbf{R} = 0\,,$$

the sum of the second and of the fourth lines may be written as:

$$\begin{aligned} & -\boldsymbol{\sigma} \circ P_{\mathrm{H}} \cdot ([P_{\mathrm{H}}\mathbf{X}\,, \mathbf{Y}] - [P_{\mathrm{H}}\mathbf{Y}\,, \mathbf{X}] - [\mathbf{X}\,, \mathbf{Y}]) = \boldsymbol{\sigma} \cdot (\tfrac{1}{2}[P_{\mathrm{H}}\,, P_{\mathrm{H}}]_{\mathrm{FN}}(\mathbf{X}, \mathbf{Y}) - [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}]) \\ & = \boldsymbol{\sigma} \cdot (\mathbf{R}(\mathbf{X}, \mathbf{Y}) - [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}]) = -\boldsymbol{\sigma} \cdot [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}]\,, \end{aligned}$$

so that

$$\boldsymbol{T}(\mathbf{X}, \mathbf{Y}) = [P_{\mathrm{H}}\mathbf{X}\,, \boldsymbol{\sigma}\mathbf{Y}] - [P_{\mathrm{H}}\mathbf{Y}\,, \boldsymbol{\sigma}\mathbf{X}] - \boldsymbol{\sigma} \cdot [P_{\mathrm{H}}\mathbf{X}\,, P_{\mathrm{H}}\mathbf{Y}]\,,$$

and the result follows by observing that the horizontal projection of a projectable vector field is equal to the horizontal lift of the projected vector field by Lemma 1.4.5 on page 88. Then setting $\mathbf{u} \circ \mathbf{p} = T\mathbf{p} \circ \mathbf{X}$ and $\mathbf{v} \circ \mathbf{p} = T\mathbf{p} \circ \mathbf{Y}$ so that $P_{\mathrm{H}}\mathbf{X} = \mathbf{H}_{\mathbf{u}}$ and $P_{\mathrm{H}}\mathbf{Y} = \mathbf{H}_{\mathbf{v}}$, we get the formula in the statement which also shows the vertical-valuedness of the torsion by Lemma 1.7.6. The formulas in terms of the connection $\mathbf{\Gamma} \in \Lambda^1(\mathbb{E}\,; \mathbb{TE})$ are direct consequences of the property:

$$\left.\begin{aligned} & [\boldsymbol{\sigma}\,, P_{\mathrm{H}} + P_{\mathrm{V}}]_{\mathrm{FN}} = [\boldsymbol{\sigma}\,, \mathbf{I}]_{\mathrm{FN}} = 0 \\ & [\boldsymbol{\sigma}\,, P_{\mathrm{H}} - P_{\mathrm{V}}]_{\mathrm{FN}} = [\boldsymbol{\sigma}\,, \mathbf{\Gamma}]_{\mathrm{FN}} \end{aligned}\right\} \implies [\boldsymbol{\sigma}\,, P_{\mathrm{H}}]_{\mathrm{FN}} = -[\boldsymbol{\sigma}\,, P_{\mathrm{V}}]_{\mathrm{FN}} = \tfrac{1}{2}[\boldsymbol{\sigma}\,, \mathbf{\Gamma}]_{\mathrm{FN}}\,,$$

which holds for any vector-form $\boldsymbol{\sigma} \in \Lambda(\mathbb{M}\,; \mathbb{TM})$. ■

Lemma 1.7.10 *The torsion* $\boldsymbol{T} = d_{P_{\mathrm{H}}}\boldsymbol{\sigma} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{V}\mathbb{E})$ *of the connection* $P_{\mathrm{H}} \in \mathrm{C}^1(\mathbb{E}\,;\mathbb{H}\mathbb{E})$ *with respect to the soldering form* $\boldsymbol{\sigma} \in \Lambda^1(\mathbb{E}\,;\mathbb{V}\mathbb{E})$ *fulfils the relations:*

$$d_{P_{\mathrm{H}}}\boldsymbol{T} = d^2_{P_{\mathrm{H}}}\boldsymbol{\sigma} = [\mathbf{R}\,,\boldsymbol{\sigma}]_{\mathrm{FN}} = -d_{\boldsymbol{\sigma}}\mathbf{R}\,, \qquad \textit{first } \textsc{Bianchi} \textit{ identity}\,,$$

$$\overline{P_{\mathrm{H}}} = P_{\mathrm{H}} + \boldsymbol{\sigma} \implies \begin{cases} \overline{\boldsymbol{T}} = \boldsymbol{T} + d_{\boldsymbol{\sigma}}\boldsymbol{\sigma}\,, \\ \overline{\mathbf{R}} = \mathbf{R} + \boldsymbol{T} + \tfrac{1}{2} d_{\boldsymbol{\sigma}}\boldsymbol{\sigma}\,. \end{cases}$$

Proof. By the graded JACOBI identity:

$$[P_{\mathrm{H}}\,,[P_{\mathrm{H}}\,,\boldsymbol{\sigma}]_{\mathrm{FN}}]_{\mathrm{FN}} = [[P_{\mathrm{H}}\,,P_{\mathrm{H}}]_{\mathrm{FN}}\,,\boldsymbol{\sigma}]_{\mathrm{FN}} - (-1)^{\deg P_{\mathrm{H}}}[P_{\mathrm{H}}\,,[P_{\mathrm{H}}\,,\boldsymbol{\sigma}]_{\mathrm{FN}}]_{\mathrm{FN}}\,,$$

we get the former formula:

$$[P_{\mathrm{H}}\,,\boldsymbol{T}]_{\mathrm{FN}} = [P_{\mathrm{H}}\,,[P_{\mathrm{H}}\,,\boldsymbol{\sigma}]_{\mathrm{FN}}]_{\mathrm{FN}} = \tfrac{1}{2}[[P_{\mathrm{H}}\,,P_{\mathrm{H}}]_{\mathrm{FN}}\,,\boldsymbol{\sigma}]_{\mathrm{FN}} = [\mathbf{R}\,,\boldsymbol{\sigma}]_{\mathrm{FN}} = -[\boldsymbol{\sigma}\,,\mathbf{R}]_{\mathrm{FN}}\,.$$

The latter result is got by a direct computation based on the bilinearity of the FN-bracket. In this respect we remark that, being $\mathbf{ker}(P_{\mathrm{H}} + \boldsymbol{\sigma}) = \mathbf{ker}(T\mathbf{p})$ and

$$(P_{\mathrm{H}} + \boldsymbol{\sigma}) \circ (P_{\mathrm{H}} + \boldsymbol{\sigma}) = P_{\mathrm{H}} \circ P_{\mathrm{H}} + P_{\mathrm{H}} \circ \boldsymbol{\sigma} + \boldsymbol{\sigma} \circ P_{\mathrm{H}} + \boldsymbol{\sigma} \circ \boldsymbol{\sigma} = P_{\mathrm{H}} + \boldsymbol{\sigma}\,,$$

the sum $P_{\mathrm{H}} + \boldsymbol{\sigma}$ is a connection. On the other hand, the difference of any two connections is a soldering form. Indeed $\mathbb{V}\mathbb{E} \subseteq \mathbf{ker}(P_{\mathrm{H}} - \overline{P_{\mathrm{H}}})$ since $\mathbf{X} \in \mathbb{V}\mathbb{E} \implies P_{\mathrm{H}}\mathbf{X} = \overline{P_{\mathrm{H}}}\mathbf{X} = 0 \implies (P_{\mathrm{H}} - \overline{P_{\mathrm{H}}})\mathbf{X} = 0$. Moreover $\mathbf{im}(P_{\mathrm{H}} - \overline{P_{\mathrm{H}}}) \subseteq \mathbb{V}\mathbb{E}$ since $T\mathbf{p} \cdot (P_{\mathrm{H}} - \overline{P_{\mathrm{H}}})\mathbf{X} = T\mathbf{p} \cdot P_{\mathrm{H}}\mathbf{X} - T\mathbf{p} \cdot \overline{P_{\mathrm{H}}}\mathbf{X} = 0$. ■

1.7.7 Canonical torsion in a tangent bundle

Let us state a preparatory result.

Lemma 1.7.11 *Let* $\boldsymbol{\Phi} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{T}\mathbb{M})$ *be an automorphism of the vector bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,;\mathbb{M})$ *, so that:*

$$\begin{cases} \boldsymbol{\Phi}(\alpha\,\mathbf{a}) = \alpha\,\boldsymbol{\Phi}(\mathbf{a})\,, \\ \boldsymbol{\Phi}(\mathbf{a} + \mathbf{b}) = \boldsymbol{\Phi}(\mathbf{a}) + \boldsymbol{\Phi}(\mathbf{b})\,, \end{cases}$$

for all $\mathbf{a}, \mathbf{b} \in \mathbb{T}\mathbb{M}$ *such that* $\boldsymbol{\tau}_{\mathbb{M}}(\mathbf{a}) = \boldsymbol{\tau}_{\mathbb{M}}(\mathbf{b})$ *and the same with* $\boldsymbol{\Phi}$ *replaced by* $\boldsymbol{\Phi}^{-1}$ *. Then:*

$$T\boldsymbol{\Phi}^{-1}(\mathbf{Vl}_{(\mathbb{T}\mathbb{M},\boldsymbol{\tau}_{\mathbb{T}\mathbb{M}},\mathbb{T}\mathbb{M})}(\boldsymbol{\Phi}(\mathbf{a})\,,\boldsymbol{\Phi}(\mathbf{b}))) = \mathbf{Vl}_{(\mathbb{T}\mathbb{M},\boldsymbol{\tau}_{\mathbb{T}\mathbb{M}},\mathbb{T}\mathbb{M})}(\mathbf{a}\,,\mathbf{b})\,.$$

Proof. By assumption we have that:

$$\boldsymbol{\Phi}(\mathbf{a}) + t\,\boldsymbol{\Phi}(\mathbf{b}) = \boldsymbol{\Phi}(\mathbf{a} + t\,\mathbf{b})\,.$$

A direct computation then gives:

$$\begin{aligned} T\boldsymbol{\Phi}^{-1}(\mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{TM}},\mathbb{TM})}(\boldsymbol{\Phi}(\mathbf{a}),\boldsymbol{\Phi}(\mathbf{b}))) &= T\boldsymbol{\Phi}^{-1}(\partial_{t=0}\,(\boldsymbol{\Phi}(\mathbf{a}) + t\,\boldsymbol{\Phi}(\mathbf{b}))) \\ &= \partial_{t=0}\,\boldsymbol{\Phi}^{-1}(\boldsymbol{\Phi}(\mathbf{a}) + t\,\boldsymbol{\Phi}(\mathbf{b})) \\ &= \partial_{t=0}\,\mathbf{a} + t\mathbf{b} = \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{TM}},\mathbb{TM})}(\mathbf{a},\mathbf{b})\,, \end{aligned}$$

where fibrewise linearity of $\boldsymbol{\Phi}^{-1} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ has been invoked. ■

The next result provides a formula which will be recalled hereafter in deriving the expression of the canonical torsion in a tangent bundle in terms of covariant derivatives.

Lemma 1.7.12 *Given a connection on a tangent bundle* $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{M})$ *and two tangent vector fields* $\mathbf{u},\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$*, the covariant derivative field* $\nabla_{\mathbf{v}}\mathbf{u} \in \mathrm{C}^0(\mathbb{M}\,;\mathbb{TM})$ *may be expressed as*

$$\mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}\,, \nabla_{\mathbf{v}}\mathbf{u} \circ \boldsymbol{\tau}_{\mathbb{M}}) = [\mathbf{H}_{\mathbf{v}}\,,\mathbf{JH}_{\mathbf{u}}]\,,$$

or equivalently as $\nabla_{\mathbf{v}}\mathbf{u} = \mathbf{vd}_{\mathbb{TM}} \circ [\mathbf{H}_{\mathbf{v}}\,,\mathbf{JH}_{\mathbf{u}}]$.

Proof. By Lemma 1.7.6 $P_{\mathrm{H}} \cdot [\mathbf{JH}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}] = 0$ and we get the equality

$$[P_{\mathrm{H}}\,,\mathbf{JH}_{\mathbf{u}}]_{\mathrm{FN}} \cdot \mathbf{H}_{\mathbf{v}} = [P_{\mathrm{H}}\mathbf{H}_{\mathbf{v}}\,,\mathbf{JH}_{\mathbf{u}}] + P_{\mathrm{H}} \cdot [\mathbf{JH}_{\mathbf{u}}\,,\mathbf{H}_{\mathbf{v}}] = [\mathbf{H}_{\mathbf{v}}\,,\mathbf{JH}_{\mathbf{u}}]\,,$$

which shows that the r.h.s. is tensorial in $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$. Moreover, we have that $\mathbf{JH}_{\mathbf{u}} = \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}\,, T\boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{H}_{\mathbf{u}}) = \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}\,, \mathbf{u} \circ \boldsymbol{\tau}_{\mathbb{M}})$. Observing that by Lemmas 1.2.12 and 1.4.8 the flow of an horizontal lift $\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{\lambda} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ is an automorphism, we may apply Lemma 1.7.11, the definition of Lie derivative, and the linearity of the vertical lift in its second argument, to show that:

$$\begin{aligned} [\mathbf{H}_{\mathbf{v}}\,,\mathbf{JH}_{\mathbf{u}}] = \mathcal{L}_{\mathbf{H}_{\mathbf{v}}}(\mathbf{JH}_{\mathbf{u}}) &:= \partial_{\lambda=0}\,T\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{-\lambda} \circ \mathbf{JH}_{\mathbf{u}} \circ \mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{\lambda} \\ &= \partial_{\lambda=0}\,T\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{-\lambda} \circ \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{\lambda}\,, \mathbf{u} \circ \boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{\lambda}) \\ &= \partial_{\lambda=0}\,T\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{-\lambda} \circ \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{\lambda}\,, \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \boldsymbol{\tau}_{\mathbb{M}}) \\ &= \partial_{\lambda=0}\,\mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}\,, \mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{-\lambda} \circ \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \boldsymbol{\tau}_{\mathbb{M}}) \\ &= \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}\,, \partial_{\lambda=0}\,(\mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_{-\lambda} \circ \mathbf{u} \circ \mathbf{Fl}^{\mathbf{v}}_{\lambda} \circ \boldsymbol{\tau}_{\mathbb{M}})\,. \end{aligned}$$

The result then follows from Lemma 1.4.10 on page 93. ■

In the tangent bundle $\boldsymbol{\tau}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM};\mathbb{M})$, the canonical torsion of a connection is the one induced by the canonical soldering form $\mathbf{J} := \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\boldsymbol{\tau}_{\mathbb{TM}}, T\boldsymbol{\tau}_{\mathbb{M}})$, so that $\boldsymbol{T} = d_{\mathbf{J}} P_{\mathrm{H}} = d_{P_{\mathrm{H}}} \mathbf{J} \in \Lambda^2(\mathbb{TM};\mathbb{T}^2\mathbb{M})$. Since the torsion is a vertical-valued horizontal 2-form on $\mathbb{TM}$, it may be evaluated on the horizontal lifts $\mathbf{H_u}, \mathbf{H_v} \in \mathrm{C}^1(\mathbb{TM};\mathbb{T}^2\mathbb{M})$ of tangent vector fields $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ and its value may be expressed as the vertical lift of the value of a 2-form on $\mathbb{M}$ with values in $\mathbb{TM}$ by

$$\begin{aligned}\mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}, \mathrm{TORS}(\mathbf{u},\mathbf{v}) \circ \boldsymbol{\tau}_{\mathbb{M}}) &:= \boldsymbol{T}(\mathbf{H_u}, \mathbf{H_v}) \\ &= [\mathbf{H_u}, \mathbf{J} \cdot \mathbf{H_v}] - [\mathbf{H_v}, \mathbf{J} \cdot \mathbf{H_u}] - \mathbf{J} \cdot [\mathbf{H_u}, \mathbf{H_v}] \,.\end{aligned}$$

Recalling that

$$\begin{aligned}\mathbf{J} \cdot \mathbf{H_v} &= \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}, T\boldsymbol{\tau}_{\mathbb{M}} \cdot \mathbf{H_v}) \\ &= \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}, \mathbf{v} \circ \boldsymbol{\tau}_{\mathbb{M}}) \,,\end{aligned}$$

and that $P_{\mathrm{H}} \cdot [\mathbf{H_u}, \mathbf{H_v}] = \mathbf{H}_{[\mathbf{u},\mathbf{v}]}$, we have:

$$\begin{aligned}\mathbf{J} \cdot [\mathbf{H_u}, \mathbf{H_v}] &= \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}, T\boldsymbol{\tau}_{\mathbb{M}} \cdot [\mathbf{H_u}, \mathbf{H_v}]) \\ &= \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}, [\mathbf{u}, \mathbf{v}] \circ \boldsymbol{\tau}_{\mathbb{M}}) \,.\end{aligned}$$

Moreover, by Lemma 1.7.12, we have that

$$[\mathbf{H_v}, \mathbf{J} \cdot \mathbf{H_u}] = \mathbf{Vl}_{(\mathbb{TM},\boldsymbol{\tau}_{\mathbb{M}},\mathbb{M})} \circ (\mathbf{id}_{\mathbb{TM}}, \nabla_{\mathbf{v}} \mathbf{u} \circ \boldsymbol{\tau}_{\mathbb{M}}) \,,$$

and hence

$$\mathrm{TORS}(\mathbf{u},\mathbf{v}) = \nabla_{\mathbf{u}} \mathbf{v} - \nabla_{\mathbf{v}} \mathbf{u} - [\mathbf{u}, \mathbf{v}] \,.$$

1.8 Symplectic manifolds

Definition 1.8.1 (Symplectic manifold) *A symplectic manifold is a manifold $\mathbb{M}$ modeled on a linear* Banach *space and endowed with a closed differential two-form $\boldsymbol{\omega} \in \Lambda^2(\mathbb{TM};\Re)$ which is weakly nondegenerate.*

This means that $\boldsymbol{\omega}^\flat_{\mathbf{x}} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M};\mathbb{T}^*_{\mathbf{x}}\mathbb{M})$ is injective for all $\mathbf{x} \in \mathbb{M}$, i.e. $\ker \boldsymbol{\omega}^\flat_{\mathbf{x}} = \{0\}$ or

$$\boldsymbol{\omega}_{\mathbf{x}}(\mathbf{u_x}, \mathbf{v_x}) = 0 \,, \quad \forall\, \mathbf{v_x} \in \mathbb{T}_{\mathbf{x}}\mathbb{M} \implies \mathbf{u_x} = 0 \,,$$

Definition 1.8.2 (Strong nondegeneracy) *The strong nondegeneracy (or nondegeneracy) of* $\boldsymbol{\omega}^\flat_{\mathbf{x}} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,; \mathbb{T}^*_{\mathbf{x}}\mathbb{M})$ *means that it is one-to-one, i.e.* $\ker \boldsymbol{\omega}^\flat_{\mathbf{x}} = \{0\}$ *and* $Im\, \boldsymbol{\omega}^\flat_{\mathbf{x}} = \mathbb{T}^*_{\mathbf{x}}\mathbb{M}$.

By the open mapping theorem, a one-to-one bounded linear map, between Banach spaces, has a bounded linear inverse. Then a weakly nondegenerate two-form $\boldsymbol{\omega}_{\mathbf{x}} \in \Lambda^2(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\Re)$ is nondegenerate if $\boldsymbol{\omega}^\flat_{\mathbf{x}} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,; \mathbb{T}^*_{\mathbf{x}}\mathbb{M})$ is onto.

Definition 1.8.3 (Exact symplectic manifold) *A symplectic manifold is exact if the two-form* $\boldsymbol{\omega}$ *is exact, that is* $\boldsymbol{\omega} = d\boldsymbol{\theta}$ *for a differential one-form* $\boldsymbol{\theta} \in \Lambda^1(\mathbb{M}\,;\Re)$.

1.8.1 Canonical forms

The standard example of a symplectic manifold is the cotangent vector bundle $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$ to a given manifold $\mathbb{M}$. The interest for this peculiar symplectic manifold is motivated by the hamiltonian description of mechanics.

Applying the tangent functor to the projector $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$ we get the fibre-wise surjective map $T\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{T}^*\mathbb{M}\,;\mathbb{T}\mathbb{M})$:

$$\left.\begin{array}{r}\forall\, \mathbf{v}^* \in \mathbb{T}^*\mathbb{M} \\ \forall\, \mathbf{v} \in \mathbb{T}_{\boldsymbol{\tau}^*_{\mathbb{M}}(\mathbf{v}^*)}\mathbb{M}\end{array}\right\} \implies \exists\, \mathbf{X}(\mathbf{v}^*) \in \mathbb{T}_{\mathbf{v}^*}\mathbb{T}^*\mathbb{M} \,:\, T\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{X}(\mathbf{v}^*) = \mathbf{v}\,.$$

The cotangent map of the projector $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$ is the map $T^*\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\boldsymbol{\tau}^*_{\mathbb{M}}{\downarrow}\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})$ introduced with Definition 1.2.14 on page 44. This is a homomorphism from the pull-back bundle $\boldsymbol{\tau}^*_{\mathbb{M}}{\downarrow}\mathbb{T}^*\mathbb{M} = \mathbb{T}^*\mathbb{M} \times_{\mathbb{M}} \mathbb{T}^*\mathbb{M}$ to the bundle $(\mathbb{T}^*\mathbb{T}^*\mathbb{M}\,, \boldsymbol{\tau}^*_{\mathbb{T}^*\mathbb{M}}\,, \mathbb{T}^*\mathbb{M})$ which, by Lemma 1.2.3 on page 45, is fibrewise injective and horizontal valued. Let us recall here Definition 1.2.15 of page 45.

Definition 1.8.4 (Liouville one-form) *The canonical one-form or* Liouville *one-form is defined by:*

$$\boldsymbol{\theta}\mathbb{M} := T^*\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathrm{DIAG} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})\,,$$

or, explicitly:

$$\langle\, \boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*), \mathbf{X}(\mathbf{v}^*)\,\rangle := \langle\, \mathbf{v}^*, T_{\mathbf{v}^*}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{X}(\mathbf{v}^*)\,\rangle\,,$$

for all $\mathbf{v}^* \in \mathbb{T}^*\mathbb{M}$ *and* $\mathbf{X}(\mathbf{v}^*) \in \mathbb{T}_{\mathbf{v}^*}\mathbb{T}^*\mathbb{M}$, *or also, briefly*

$$\langle\, \boldsymbol{\theta}\mathbb{M}, \mathbf{X}\,\rangle := \langle\, \boldsymbol{\tau}_{\mathbb{T}^*\mathbb{M}}(\mathbf{X}), T\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \mathbf{X}\,\rangle \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\Re)\,,$$

for all sections $\mathbf{X} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}\mathbb{T}^*\mathbb{M})$ *of the bundle* $\boldsymbol{\tau}_{\mathbb{T}^*\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$.

The definition of the canonical one-form in terms of the cotangent map appears to be new. On its grounds many related properties are clarified and their proof is greatly simplified.

Lemma 1.8.1 *In a local chart* $\varphi \in \mathrm{C}^1(U_{\mathbb{M}}\,; U_E)$ *the canonical one-form is given by:*

$$\boldsymbol{\theta}E(v^*)) \cdot X(v^*) = \langle \mathrm{pr}_2(v^*), \mathrm{pr}_1(X(v^*)) \rangle_{E^* \times E} \,.$$

Proof. The induced local charts in the tangent manifold, in the cotangent manifold, in the manifold tangent to the cotangent manifold and in the manifold cotangent to the cotangent manifold, are respectively given by:

$$\begin{aligned}
&T\varphi \in \mathrm{C}^1(\mathbb{T}\mathbb{M}\,; E \times E)\,,\\
&T^*\varphi^{-1} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,; E \times E^*)\,,\\
&TT^*\varphi^{-1} \in \mathrm{C}^1(\mathbb{T}\mathbb{T}^*\mathbb{M}\,; E \times E^* \times E \times E^*)\,,\\
&T^*T^*\varphi \in \mathrm{C}^1(\mathbb{T}^*\mathbb{T}^*\mathbb{M}\,; E \times E^* \times E^* \times E^{**})\,.
\end{aligned}$$

Let us set

$$\begin{aligned}
v^* &:= T^*\varphi^{-1}(\mathbf{v}^*) \in E \times E^*\,,\\
X &:= T^*\varphi^{-1}{\uparrow}\mathbf{X} \in E \times E^* \times E \times E^*\,,\\
\boldsymbol{\theta}E &:= (T^*\varphi^{-1}{\uparrow}\boldsymbol{\theta}\mathbb{M}) \circ T^*\varphi \in E \times E^* \times E^* \times E^{**}\,,\\
\pi^*(v^*) &:= \mathrm{pr}_1(v^*) = (\varphi \circ \boldsymbol{\tau}^*_{\mathbb{M}} \circ T^*\varphi)(v^*) \in E\,,
\end{aligned}$$

where the last equality follows from the commutative diagram:

$$\begin{array}{ccc}
\mathbb{T}^*\mathbb{M} & \xrightarrow{T^*\varphi^{-1}} & E \times E^* \\
{\scriptstyle \boldsymbol{\tau}^*_{\mathbb{M}}}\big\downarrow & & \big\downarrow{\scriptstyle \mathrm{pr}_1} \\
\mathbb{M} & \xrightarrow{\varphi} & E
\end{array}
\qquad \Longleftrightarrow \qquad \varphi \circ \boldsymbol{\tau}^*_{\mathbb{M}} = \mathrm{pr}_1 \circ T^*\varphi^{-1}\,.$$

The canonical one-form is the expressed in the model space by

$$\begin{aligned}
\boldsymbol{\theta}E(v^*)) \cdot X(v^*) &:= \langle v^*, T\pi^*(v^*) \cdot X(v^*) \rangle_{E^* \times E}\\
&= \langle \mathrm{pr}_2(v^*), \mathrm{pr}_1(X(v^*)) \rangle_{E^* \times E}\,,
\end{aligned}$$

and the result is proved. ■

The statement of next result is a modification of a from [3], Proposition 3.2.11 on page 179, by adding the assumption of horizontal-valuedness of the one-form $\boldsymbol{\theta}$.

Theorem 1.8.1 (Characterization of the canonical one-form) *The canonical or* LIOUVILLE *one-form* $\boldsymbol{\theta}\mathbb{M} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})$ *defined by* $\boldsymbol{\theta}\mathbb{M} := T^*\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathrm{DIAG}$ *is equivalently characterized as the unique horizontal-valued one-form* $\boldsymbol{\theta} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})$ *fulfilling the property:*

$$\boldsymbol{\alpha}{\downarrow}\boldsymbol{\theta}\mathbb{M} = \boldsymbol{\alpha}\,,$$

for any section $\boldsymbol{\alpha} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ *of the cotangent bundle* $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$.

Proof. By definition of the pull-back one-form $\boldsymbol{\alpha}{\downarrow}\boldsymbol{\theta}\mathbb{M} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ we have that:

$$\begin{aligned}
\langle \boldsymbol{\alpha}{\downarrow}\boldsymbol{\theta}\mathbb{M}(\mathbf{x}), \mathbf{v}_{\mathbf{x}} \rangle &= \langle \boldsymbol{\theta}\mathbb{M}(\boldsymbol{\alpha}(\mathbf{x})), T_{\mathbf{x}}\boldsymbol{\alpha} \cdot \mathbf{v}_{\mathbf{x}} \rangle \\
&= \langle T^*\boldsymbol{\tau}^*_{\mathbb{M}}(\boldsymbol{\alpha}(\mathbf{x})\,, \boldsymbol{\alpha}(\mathbf{x})), T_{\mathbf{x}}\boldsymbol{\alpha} \cdot \mathbf{v}_{\mathbf{x}} \rangle \\
&= \langle T^*_{\boldsymbol{\alpha}(\mathbf{x})}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot \boldsymbol{\alpha}(\mathbf{x}), T_{\mathbf{x}}\boldsymbol{\alpha} \cdot \mathbf{v}_{\mathbf{x}} \rangle \\
&= \langle \boldsymbol{\alpha}(\mathbf{x}), T_{\boldsymbol{\alpha}(\mathbf{x})}\boldsymbol{\tau}^*_{\mathbb{M}} \cdot T_{\mathbf{x}}\boldsymbol{\alpha} \cdot \mathbf{v}_{\mathbf{x}} \rangle \\
&= \langle \boldsymbol{\alpha}(\mathbf{x}), T_{\mathbf{x}}(\boldsymbol{\tau}^*_{\mathbb{M}} \circ \boldsymbol{\alpha}) \cdot \mathbf{v}_{\mathbf{x}} \rangle \\
&= \langle \boldsymbol{\alpha}(\mathbf{x}), \mathbf{v}_{\mathbf{x}} \rangle\,, \quad \forall\, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,,
\end{aligned}$$

and hence the property is fulfilled. Vice versa if $\boldsymbol{\theta} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})$ and $\boldsymbol{\alpha}{\downarrow}\boldsymbol{\theta} = \boldsymbol{\alpha}$ for any section $\boldsymbol{\alpha} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}^*\mathbb{M})$ of the cotangent bundle $\boldsymbol{\tau}^*_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\mathbb{M})$, from the previous equalities we infer that

$$\langle \boldsymbol{\theta}(\boldsymbol{\alpha}(\mathbf{x})), T_{\mathbf{x}}\boldsymbol{\alpha} \cdot \mathbf{v}_{\mathbf{x}} \rangle = \langle (T^*\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathrm{DIAG})(\boldsymbol{\alpha}(\mathbf{x})), T_{\mathbf{x}}\boldsymbol{\alpha} \cdot \mathbf{v}_{\mathbf{x}} \rangle\,, \quad \forall\, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,.$$

Hence, by the horizontal-valuedness of $\boldsymbol{\theta}$ and $\boldsymbol{\theta}\mathbb{M}$ we get:

$$\langle \boldsymbol{\theta}(\boldsymbol{\alpha}(\mathbf{x})), \mathbf{H}(\boldsymbol{\alpha}_{\mathbf{x}}\,, \mathbf{v}_{\mathbf{x}}) \rangle = \langle (T^*\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathrm{DIAG})(\boldsymbol{\alpha}(\mathbf{x})), \mathbf{H}(\boldsymbol{\alpha}_{\mathbf{x}}\,, \mathbf{v}_{\mathbf{x}}) \rangle\,, \quad \forall\, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,.$$

The surjectivity of the linear map $\mathbf{H}(\boldsymbol{\alpha}_{\mathbf{x}}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{H}_{\boldsymbol{\alpha}_{\mathbf{x}}}\mathbb{M})$ then yields the equality $\boldsymbol{\theta} = T^*\boldsymbol{\tau}^*_{\mathbb{M}} \circ \mathrm{DIAG}$. ■

Definition 1.8.5 *The* ***canonical two-form*** $\boldsymbol{\omega}\mathbb{M} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,;\Lambda(\mathbb{T}\mathbb{T}^*\mathbb{M}^2\,;\Re))$ *is the negative exterior derivative of the canonical one-form:*

$$\boldsymbol{\omega}\mathbb{M} := -d\boldsymbol{\theta}\mathbb{M}\,.$$

The map $\boldsymbol{\omega}\mathbb{M}^\flat \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,; BL\,(\mathbb{T}\mathbb{T}^*\mathbb{M}\,; \mathbb{T}^*\mathbb{T}^*\mathbb{M}))$ provides an isomorphism between the bundles $\boldsymbol{\tau}_{\mathbb{T}^*\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}\mathbb{T}^*\mathbb{M}\,; \mathbb{T}^*\mathbb{M})$ and $\boldsymbol{\tau}^*_{\mathbb{T}^*\mathbb{M}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{T}^*\mathbb{M}\,; \mathbb{T}^*\mathbb{M})$ if it is surjective. Injectivity follows by the next Theorem.

Theorem 1.8.2 (Weak nondegeneracy of the canonical two-form) *The canonical two-form* $\boldsymbol{\omega}\mathbb{M}$ *is weakly nondegenerate*

$$\boldsymbol{\omega}\mathbb{M}(\mathbf{v}^*) \cdot \mathbf{X}_{\mathbf{v}^*} \cdot \mathbf{Y}_{\mathbf{v}^*} = 0 \quad \forall\, \mathbf{Y}_{\mathbf{v}^*} \in \mathbb{T}_{\mathbf{v}^*}\mathbb{T}\mathbb{M} \implies \mathbf{X}_{\mathbf{v}^*} = 0\,,$$

and hence is an exact symplectic two-form.

Proof. In a local chart $\varphi \in \mathrm{C}^1(U_\mathbb{M}\,; U_E)$, by PALAIS' formula for the exterior derivative, we have that:

$$\begin{aligned} d\boldsymbol{\theta}E(v^*) \cdot X(v^*) \cdot Y(v^*) = {} & d_{X(v^*)}\boldsymbol{\theta}E(v^*) \cdot Y(v^*) \\ & - d_{Y(v^*)}\boldsymbol{\theta}E(v^*) \cdot X(v^*) \\ & - \boldsymbol{\theta}E(v^*) \cdot [X(v^*), Y(v^*)\,,\,]. \end{aligned}$$

By tensoriality of the exterior derivative, may assume that the vector fields $\mathbf{X}, \mathbf{Y} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,; \mathbb{T}\mathbb{T}^*\mathbb{M})$ are such that their images through the local chart $\varphi \in \mathrm{C}^1(U_\mathbb{M}\,; U_E)$ are constant vector fields $X, Y \in \mathrm{C}^1(E \times E^*\,; E \times E^*)$.

Then the flows $\mathbf{Fl}^X_\lambda$ and $\mathbf{Fl}^Y_\lambda$ commute and $[X\,, Y] = 0$. PALAIS' formula then gives

$$\begin{aligned} d\boldsymbol{\theta}E(v^*) \cdot X(v^*) \cdot Y(v^*) = {} & d_{X(v^*)}\boldsymbol{\theta}E(v^*) \cdot Y(v^*) \\ & - d_{Y(v^*)}\boldsymbol{\theta}E(v^*) \cdot X(v^*) \\ = {} & d_{X(v^*)}\langle v^*, T\pi^*(v^*) \cdot Y(v^*)\rangle \\ & - d_{Y(v^*)}\langle v^*, T\pi^*(v^*) \cdot X(v^*)\rangle\,. \end{aligned}$$

The duality pairing $\langle v^*, T\pi^*(v^*)\cdot X(v^*)\rangle_{E^*\times E}$ is performed between the components $\mathrm{pr}_2(v^*) \in E^*$ and $\mathrm{pr}_2(T\pi^*(v^*)\cdot X(v^*)) \in E$ since the vectors $v^* \in E\times E^*$ and $T\pi^*(v^*) \cdot X(v^*) \in E \times E$ are based at the same point in E.

Next we observe that $\pi^* = \mathrm{pr}_1$ is a constant linear map from $E \times E^*$ onto E and that the vectors $X(v^*), Y(v^*) \in E \times E^*$ are independent of the point $v^* \in E\times E^*$. The vectors $T\pi^*(v^*)\cdot X(v^*) \in E\times E$ and $T\pi^*(v^*)\cdot Y(v^*) \in E\times E$ are then also independent of the point $v^* \in E \times E^*$. The derivatives in PALAIS' formula may thus be easily evaluated to give:

$$\begin{aligned} d\boldsymbol{\theta}E(v^*) \cdot X(v^*) \cdot Y(v^*) = {} & \langle X(v^*), T\pi^*(v^*) \cdot Y(v^*)\rangle_{E^*\times E} \\ & - \langle Y(v^*), T\pi^*(v^*) \cdot X(v^*)\rangle_{E^*\times E}\,, \end{aligned}$$

where the duality pairings are performed between the second cartesian components of the involved pairs. Observing that $\mathrm{pr}_1(X(v^*)) = \mathrm{pr}_2(T\pi^*(v^*) \cdot X(v^*))$, the condition $d(\boldsymbol{\theta}E(v^*)) \cdot X(v^*) \cdot Y(v^*) = 0$ for all $Y(v^*) \in E \times E^*$, may be conveniently rewritten as

$$\langle \mathrm{pr}_2(X(v^*)), \mathrm{pr}_1(Y(v^*)) \rangle_{E^* \times E} - \langle \mathrm{pr}_2(Y(v^*)), \mathrm{pr}_1(X(v^*)) \rangle_{E^* \times E} = 0 \,.$$

Assuming that $\mathrm{pr}_1(Y(v^*)) = 0$, we get the implication

$$\langle \mathrm{pr}_2(Y(v^*)), \mathrm{pr}_1(X(v^*)) \rangle_{E^* \times E} = 0 \,, \quad \forall\, \mathrm{pr}_2(Y(v^*)) \in E^* \implies \mathrm{pr}_1(X(v^*)) = 0 \,.$$

Then we may conclude that

$$\langle \mathrm{pr}_2(X(v^*)), \mathrm{pr}_1(Y(v^*)) \rangle_{E^* \times E} = 0 \,, \quad \forall\, \mathrm{pr}_1(Y(v^*)) \in E \iff \mathrm{pr}_2(X(v^*)) = 0 \,,$$

and the proposition is proved. ■

The proof of Theorem 1.8.2 provides the following simple representation of the canonical two form in a local chart:

$$-d\boldsymbol{\theta}E(v^*)^\flat = J^\flat \,,$$

where the linear operator $J^\flat \in BL(E \times E^* ; E^* \times E)$ is defined by

$$J^\flat(X(v^*)) := \{ -\mathrm{pr}_2(X(v^*)) \,, \mathrm{pr}_1(X(v^*)) \} \in E^* \times E \,, \quad \forall\, X(v^*) \in E \times E^* \,,$$

with the block-matrix representation

$$J^\flat = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} .$$

Accordingly, the result of Theorem 1.8.2 may be stated as

$$\ker J^\flat = \{0, 0\} \in E \times E^* \,.$$

Moreover, defining the linear operator $J^\sharp \in BL(E^* \times E ; E \times E^*)$ by

$$J^\sharp(X^*(v^*)) := \{ -\mathrm{pr}_2(X^*(v^*)) \,, \mathrm{pr}_1(X^*(v^*)) \} \in E \times E^* \,, \quad \forall\, X^*(v^*) \in E \times E^* \,,$$

we have that $J^\sharp \circ J^\flat = -\mathrm{id}_{E \times E^*}$ and $J^\flat \circ J^\sharp = -\mathrm{id}_{E^* \times E}$. Moreover:

$$\begin{aligned} -d\boldsymbol{\theta}E(v^*) \cdot X(v^*) \cdot Y(v^*) &= \langle J^\flat \cdot X(v^*), Y(v^*) \rangle_{(E^* \times E) \times (E \times E^*)} \\ &= -\langle J^\flat \cdot Y(v^*), X(v^*) \rangle_{(E^* \times E) \times (E \times E^*)} \,, \end{aligned}$$

that is $J^A = -J^\flat$ where $J^A \in BL(E \times E^* ; E^* \times E)$ is the adjoint operator of $J^\flat \in BL(E \times E^* ; E^* \times E)$. Hence $J^A J^\flat = I$.

1.8.2 Darboux theorem

We have seen that the exterior derivative of the canonical one-form on $\mathbb{T}^*\mathbb{M}$ is an exact two-form with a trivial kernel, i.e. $Ker\,\boldsymbol{\omega}\mathbb{M} = \{0\}$, and that its push forward along any local chart is a constant two-form in the linear model space $\mathbb{E}$. The corresponding linear operator performs a conterclockwise block rotation of $\pi/2$ in the product space $\mathbb{E} \times \mathbb{E}^*$.

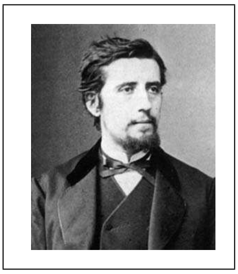

Figure 1.39: Jean Gaston Darboux (1842 - 1917)

A classical result due to Darboux [31] applies to closed nondegenerate two-forms on a simplectic manifold. We reproduce here the modern, elegant proof due to Jürgen Moser [2].

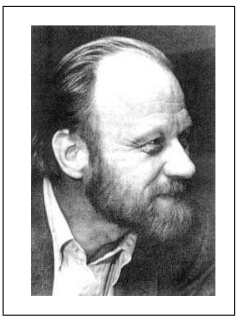

Figure 1.40: Jürgen Kurt Moser (1928 - 1999)

Theorem 1.8.3 (Darboux theorem) *A closed, nondegenerate two-form* $\boldsymbol{\omega} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{M}\,; \mathbb{T}^*\mathbb{T}^*\mathbb{M})$ *can be locally mapped to a constant two-form on the model linear space.*

Proof. Let $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;E)$ be a local chart to the model BANACH space E. Being $\boldsymbol{\varphi}\!\downarrow\! d\boldsymbol{\omega} = d(\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega})$, a diffeomorphic chart preserve closedness.

We may then assume that a chart has been applied, so that we have to look for a change of chart $\mathrm{C}^1(E\,;E)$ in a ball around the origin of E which makes the two-form a constant two-form.

To this end, we set $\boldsymbol{\omega}_0 = \boldsymbol{\omega}(0)$ and $\boldsymbol{\omega}_t = \boldsymbol{\omega}_0 + t(\boldsymbol{\omega}_x - \boldsymbol{\omega}_0)$, so that $\boldsymbol{\omega}_1 = \boldsymbol{\omega}_x$.

By POINCARÉ Lemma we may assume that $\boldsymbol{\omega}_x - \boldsymbol{\omega}_0 = d\boldsymbol{\alpha}$.

Moreover the assumed property that $\boldsymbol{\omega}_x$ and $\boldsymbol{\omega}_0$ are nondegenerate ensure that $\boldsymbol{\omega}_t$ is nondegenerate for all $t \in [0,1]$. Then the equation $\boldsymbol{\omega}_t \cdot \mathbf{X}_t = -\boldsymbol{\alpha}$ admits a unique solution for any given $\boldsymbol{\alpha}$.

The associated time-dependent flow $\mathbf{Fl}^{\mathbf{X}}_{\tau,t}$ drags the two-form $\boldsymbol{\omega}_t$ according to the rule

$$\partial_{\tau=t}\,\mathbf{Fl}^{\mathbf{X}}_{\tau,s}\!\downarrow\!\boldsymbol{\omega}_\tau = \mathbf{Fl}^{\mathbf{X}}_{\tau,t}\!\downarrow\!\mathcal{L}_{t,\mathbf{X}}\boldsymbol{\omega}_t\,.$$

Being $d\boldsymbol{\omega}_t = 0$ we have that

$$\begin{aligned}
\mathcal{L}_{t,\mathbf{X}}\boldsymbol{\omega}_t &= \partial_{\tau=t}\,\boldsymbol{\omega}_\tau + (\mathcal{L}_{\mathbf{X}}\boldsymbol{\omega})_t \\
&= \boldsymbol{\omega}_x - \boldsymbol{\omega}_0 + (\mathcal{L}_{\mathbf{X}}\boldsymbol{\omega})_t \\
&= \boldsymbol{\omega}_x - \boldsymbol{\omega}_0 + d(\boldsymbol{\omega}_t \cdot \mathbf{X}_t) + d\boldsymbol{\omega}_t \cdot \mathbf{X}_t \\
&= d(\boldsymbol{\alpha} + \boldsymbol{\omega}_t \cdot \mathbf{X}_t) = 0\,,
\end{aligned}$$

so that $\partial_{\tau=t}\,\mathbf{Fl}^{\mathbf{X}}_{\tau,s}\!\downarrow\!\boldsymbol{\omega}_\tau = 0$ and hence $\mathbf{Fl}^{\mathbf{X}}_{1,0}\!\downarrow\!\boldsymbol{\omega}_x = \boldsymbol{\omega}_0$. ■

1.8.3 Finite dimensional symplectic manifolds

Let us consider a n-D differentiable manifold $\mathbb{M}$ with model linear space E and dual space E^*.

Let $\{\mathbf{e}_j \in E, j = 1, \dots, n\}$ and $\{\mathbf{e}^i \in E^*, i = 1, \dots, n\}$ be dual bases in E and E^* so that $\langle \mathbf{e}^i, \mathbf{e}_j \rangle = I^{i}_{.j}$ with $I \in BL(\Re^n\,;\Re^n)$ the identity matrix.

Given a diffeomorphic local chart $\varphi \in \mathrm{C}^1(U_{\mathbb{M}}\,;U_E)$, the bases in the tangent bundle, the cotangent bundle, the tangent bundle to the cotangent bundle and the cotangent bundle to the cotangent bundle, are then generated by the coordinate maps:

$$\begin{aligned}
T\varphi^{-1} &\in \mathrm{C}^1(E \times E\,;\mathbb{TM})\,, \\
T^*\varphi &\in \mathrm{C}^1(E \times E^*\,;\mathbb{T}^*\mathbb{M})\,, \\
TT^*\varphi &\in \mathrm{C}^1(E \times E^* \times E \times E^*\,;\mathbb{TT}^*\mathbb{M})\,, \\
T^*T^*\varphi^{-1} &\in \mathrm{C}^1(E \times E^* \times E^* \times E^{**}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{M})\,.
\end{aligned}$$

The tangent space $\mathbb{T}_{\mathbf{x}}\mathbb{M}$ is generated by the velocities of the coordinate lines and the chart-induced basis is denoted by $\{\partial\mathbf{x}_i := T\varphi^{-1}(\mathbf{e}_i), i = 1,\dots,n\}$. The dual basis in the cotangent space $\mathbb{T}^*_{\mathbf{x}}\mathbb{M}$ is given by $\{d\mathbf{x}^i := T^*\varphi(\mathbf{e}^i), i = 1,\dots,n\}$ so that

$$\langle d\mathbf{x}^i, \partial\mathbf{x}_j \rangle_{\mathbb{T}^*_{\mathbf{x}}\mathbb{M}\times\mathbb{T}_{\mathbf{x}}\mathbb{M}} = \langle T^*\varphi(\mathbf{e}^i), T\varphi^{-1}(\mathbf{e}_j)\rangle_{\mathbb{T}^*_{\mathbf{x}}\mathbb{M}\times\mathbb{T}_{\mathbf{x}}\mathbb{M}} = \langle \mathbf{e}^i, \mathbf{e}_j \rangle_{E^*\times E} = I^{i.}_{.j}\,,$$

Setting $\mathbf{v}^* = p_i\, d\mathbf{x}^i$ and $\mathbf{x} = \boldsymbol{\tau}^*(\mathbf{v}^*)$, we have that

$$\begin{aligned}
\boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*) &= \{p_i\, d\mathbf{x}^i\,, 0\,\partial\mathbf{x}^k\} = \{p_i\, d\mathbf{x}^i\,, 0\}\,,\\
\mathbf{X}(\mathbf{v}^*) &= \{\alpha^k\,\partial\mathbf{x}_k\,, \beta_j\, d\mathbf{x}^j\}\,,\\
\mathbf{Y}(\mathbf{v}^*) &= \{\gamma^k\,\partial\mathbf{x}_k\,, \delta_j\, d\mathbf{x}^j\}\,,\\
\boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*)\cdot\mathbf{X}(\mathbf{v}^*) &= p_i(\mathbf{x})\,\alpha^i(\mathbf{x})\,,\\
\boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*)\cdot\mathbf{Y}(\mathbf{v}^*) &= p_i(\mathbf{x})\,\gamma^i(\mathbf{x})\,.
\end{aligned}$$

Hence

$$\begin{aligned}
d\boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*)\cdot\mathbf{X}(\mathbf{v}^*)\cdot\mathbf{Y}(\mathbf{v}^*) &= d_{\mathbf{X}(\mathbf{v}^*)}(\boldsymbol{\theta}\mathbb{M}\cdot\mathbf{Y})(\mathbf{v}^*) - d_{\mathbf{Y}(\mathbf{v}^*)}(\boldsymbol{\theta}\mathbb{M}\cdot\mathbf{X})(\mathbf{v}^*)\\
&\quad - \boldsymbol{\theta}\mathbb{M}(\mathbf{v}^*)\cdot[\mathbf{X}(\mathbf{v}^*)\,, \mathbf{Y}(\mathbf{v}^*)]\\
&= \delta_i(\mathbf{x})\,\alpha^i(\mathbf{x}) - \beta_i(\mathbf{x})\,\gamma^i(\mathbf{x})\,.
\end{aligned}$$

1.8.4 Symplectic maps

- A map $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ between two symplectic manifolds $\{\mathbb{M}\,,\boldsymbol{\omega}_{\mathbb{M}}\}$ and $\{\mathbb{N}\,,\boldsymbol{\omega}_{\mathbb{N}}\}$ is *symplectic* if it preserves the simplectic forms: $\boldsymbol{\omega}_{\mathbb{M}} = \boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega}_{\mathbb{N}}$.

In section 1.2.4 we introduced the *cotangent map* $T^*\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{T}^*\mathbb{M})$ of a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$, pointwise defined as the dual $T^*\boldsymbol{\varphi}(\mathbf{y}) \in BL\,(\mathbb{T}^*_{\mathbf{y}}\mathbb{N}\,;\mathbb{T}^*_{\boldsymbol{\varphi}^{-1}(\mathbf{y})}\mathbb{M})$ of the *tangent map* $T\boldsymbol{\varphi}(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{N})$, according to the relation:

$$\langle T^*_{\boldsymbol{\varphi}(\mathbf{x})}\boldsymbol{\varphi}\cdot\mathbf{v}^*_{\mathbb{N}}(\boldsymbol{\varphi}(\mathbf{x})), \mathbf{v}(\mathbf{x})\rangle = \langle \mathbf{v}^*_{\mathbb{N}}(\boldsymbol{\varphi}(\mathbf{x})), T_{\mathbf{x}}\boldsymbol{\varphi}\cdot\mathbf{v}(\mathbf{x})\rangle\,,$$

with the commutative diagram:

$$\begin{array}{ccc}
\mathbb{T}^*\mathbb{M} & \xleftarrow{\;T^*\boldsymbol{\varphi}\;} & \mathbb{T}^*\mathbb{N}\\
{\scriptstyle\boldsymbol{\tau}^*_{\mathbb{M}}}\big\downarrow & & {\scriptstyle\boldsymbol{\tau}^*_{\mathbb{N}}}\big\downarrow\\
\mathbb{M} & \xleftarrow{\;\boldsymbol{\varphi}^{-1}\;} & \mathbb{N}
\end{array}
\quad\iff\quad \boldsymbol{\varphi}^{-1}\circ\boldsymbol{\tau}^*_{\mathbb{N}} = \boldsymbol{\tau}^*_{\mathbb{M}}\circ T^*\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{M})\,.$$

Theorem 1.8.4 (Symplecticity of cotangent maps) *The cotangent map* $T^*\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{T}^*\mathbb{M})$ *of a diffeomorphism* $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ *between the symplectic spaces* $\{\mathbb{T}^*\mathbb{M}\,,\boldsymbol{\omega}\mathbb{M}\}$ *and* $\{\mathbb{T}^*\mathbb{N}\,,\boldsymbol{\omega}\mathbb{N}\}$ *meets the invariance property:*

$$(T^*\boldsymbol{\varphi})\!\downarrow\!\boldsymbol{\theta}\mathbb{M} = \boldsymbol{\theta}\mathbb{N}\,.$$

and hence is symplectic.

Proof. For any $\mathbf{b}^* \in \mathbb{T}^*\mathbb{N}$ and $\mathbf{Y}(\mathbf{b}^*) \in \mathbb{T}_{\mathbf{b}^*}\mathbb{T}^*\mathbb{N}$ we have that

$$\begin{aligned}
(T^*\boldsymbol{\varphi})\!\downarrow\!\boldsymbol{\theta}\mathbb{M}(\mathbf{b}^*)\cdot\mathbf{Y}(\mathbf{b}^*) &= \boldsymbol{\theta}\mathbb{M}(T^*\boldsymbol{\varphi}(\mathbf{b}^*))\cdot T^*\boldsymbol{\varphi}\!\uparrow\!\mathbf{Y}(\mathbf{b}^*)\\
&= \langle T^*\boldsymbol{\varphi}(\mathbf{b}^*), T\boldsymbol{\tau}^*_{\mathbb{M}}(T^*\boldsymbol{\varphi}(\mathbf{b}^*))\cdot TT^*\boldsymbol{\varphi}(\mathbf{b}^*)\cdot\mathbf{Y}(\mathbf{b}^*)\rangle\\
&= \langle T^*\boldsymbol{\varphi}(\mathbf{b}^*), T(\boldsymbol{\tau}^*_{\mathbb{M}}\circ T^*\boldsymbol{\varphi})(\mathbf{b}^*)\cdot\mathbf{Y}(\mathbf{b}^*)\rangle\\
&= \langle T^*\boldsymbol{\varphi}(\mathbf{b}^*), T(\boldsymbol{\varphi}^{-1}\circ\boldsymbol{\tau}^*_{\mathbb{N}})(\mathbf{b}^*)\cdot\mathbf{Y}(\mathbf{b}^*)\rangle\\
&= \langle \mathbf{b}^*, ((T\boldsymbol{\varphi}\circ T\boldsymbol{\varphi}^{-1})\circ T\boldsymbol{\tau}^*_{\mathbb{N}})(\mathbf{b}^*)\cdot\mathbf{Y}(\mathbf{b}^*)\rangle\\
&= \langle \mathbf{b}^*, T\boldsymbol{\tau}^*_{\mathbb{N}}(\mathbf{b}^*)\cdot\mathbf{Y}(\mathbf{b}^*)\rangle\\
&= \boldsymbol{\theta}\mathbb{M}(\mathbf{b}^*)\cdot\mathbf{Y}(\mathbf{b}^*)\,,
\end{aligned}$$

which gives

$$(T^*\boldsymbol{\varphi})\!\downarrow\!\boldsymbol{\theta}\mathbb{M} = \boldsymbol{\theta}\mathbb{M}\,.$$

By naturality of the exterior derivative with respect to the push:

$$T^*\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}\mathbb{M} = d(T^*\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\theta}\mathbb{M}) = d\boldsymbol{\theta}\mathbb{N} = \boldsymbol{\omega}\mathbb{N}\,.$$

we may conclude that the cotangent lift is a symplectic map. ■

A simpler proof of the statement in Theorem 1.8.4 is got by recalling that $\boldsymbol{\theta}\mathbb{M} = T^*\boldsymbol{\tau}^*_{\mathbb{M}}$ and that

$$(T^*\boldsymbol{\varphi})\!\downarrow\!\boldsymbol{\theta}\mathbb{M} = \boldsymbol{\theta}\mathbb{N} \iff T^*T^*\boldsymbol{\varphi}\circ\boldsymbol{\theta}\mathbb{M}\circ T^*\boldsymbol{\varphi} = \boldsymbol{\theta}\mathbb{N}\,.$$

Indeed a direct computation gives

$$T^*T^*\boldsymbol{\varphi}\circ\boldsymbol{\theta}\mathbb{M}\circ T^*\boldsymbol{\varphi} = T^*(\boldsymbol{\varphi}\circ\boldsymbol{\tau}^*_{\mathbb{M}}\circ T^*\boldsymbol{\varphi}) = T^*\boldsymbol{\tau}^*_{\mathbb{N}}\,.$$

A symplectic map which is the *cotangent map* $T^*\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{T}^*\mathbb{N}\,;\mathbb{T}^*\mathbb{M})$ of a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ is called a *point transformation* [3]. A map which preserves the canonical one-form is called a *homogeneous canonical transformation* or a Mathieu *transformation* [191].

1.8.5 Poincare-Cartan one-form

In section 1.4.13 we have shown how LEGENDRE transform induced by a Lagrangian $L \in \mathrm{C}^1(\mathbb{TM};\Re)$ provides a homeomorphism between the tangent and the cotangent bundles. The fibre-preserving property of this homeomorphism is expressed by the relation $\boldsymbol{\tau}^*_{\mathbb{M}} \circ d_{\mathrm{F}} L = \boldsymbol{\tau}_{\mathbb{M}}$.

Definition 1.8.6 *The* **Poincaré-Cartan** *one-form* $\boldsymbol{\theta}_L \in \mathrm{C}^1(\mathbb{TM};\mathbb{T}^*\mathbb{TM})$ *is the pull-back by the* LEGENDRE *transform* $d_{\mathrm{F}} L \in \mathrm{C}^1(\mathbb{TM};\mathbb{T}^*\mathbb{M})$ *of the canonical one-form on the tangent bundle:*

$$\boldsymbol{\theta}_L := d_{\mathrm{F}} L \downarrow \boldsymbol{\theta}\mathbb{M}\,.$$

Lemma 1.8.2 *The* **Poincaré-Cartan** *one-form may be written as*

$$\boldsymbol{\theta}_L = T^* \boldsymbol{\tau}_{\mathbb{M}} \cdot d_{\mathrm{F}} L = \mathbf{J}^* \cdot dL\,.$$

and explicitly $\boldsymbol{\theta}_L(\mathbf{v}) = T^* \boldsymbol{\tau}_{\mathbb{M}}(\mathbf{v}) \cdot d_{\mathrm{F}} L(\mathbf{v}) = \mathbf{J}^*(\mathbf{v}) \cdot dL(\mathbf{v})$ *.*

Proof. From section 1.4.12 on page 111 it is $d_{\mathrm{F}} L := \mathbf{Vl}^*_{(\mathbb{TM}\,,\boldsymbol{\tau}_{\mathbb{M}}\,,\mathbb{M})} \circ d\mathrm{L}$, and hence

$$\begin{aligned}
\boldsymbol{\theta}_L = d_{\mathrm{F}} L \downarrow \boldsymbol{\theta}\mathbb{M} &= T^* d_{\mathrm{F}} L \circ \boldsymbol{\theta}\mathbb{M} \circ d_{\mathrm{F}} L = T^* d_{\mathrm{F}} L \circ T^* \boldsymbol{\tau}^*_{\mathbb{M}} \circ d_{\mathrm{F}} L \\
&= T^*(\boldsymbol{\tau}^*_{\mathbb{M}} \circ d_{\mathrm{F}} L) \circ d_{\mathrm{F}} L = T^* \boldsymbol{\tau}_{\mathbb{M}} \circ d_{\mathrm{F}} L \\
&= T^* \boldsymbol{\tau}_{\mathbb{M}} \circ \mathbf{Vl}^*_{(\mathbb{TM}\,,\boldsymbol{\tau}_{\mathbb{M}}\,,\mathbb{M})} \circ dL = (\mathbf{Vl}_{(\mathbb{TM}\,,\boldsymbol{\tau}_{\mathbb{M}}\,,\mathbb{M})} \circ T\boldsymbol{\tau}_{\mathbb{M}})^* \circ dL \\
&= \mathbf{J}^* \circ dL\,,
\end{aligned}$$

which is the result. ■

Defining the derivative $d_{\mathbf{J}}$ of a Lagrangian $L \in \mathrm{C}^1(\mathbb{TM};\Re)$ by the relation

$$\langle d_{\mathbf{J}} L(\mathbf{v}), \mathbf{X}(\mathbf{v}) \rangle = \langle dL(\mathbf{v}), \mathbf{J}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) \rangle\,, \quad \forall\, \mathbf{x}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{TM}\,,$$

we have that $d_{\mathbf{J}} L = \mathbf{J}^* \circ dL = \boldsymbol{\theta}_L$.

1.9 Riemannian manifolds

A RIEMANN's manifold is a differentiable manifold $\mathbb{M}$ endowed with a twice covariant metric tensor field $\mathbf{g} \in \mathrm{C}^1(\mathbb{M}; BL(\mathbb{TM}^2;\Re))$ which is symmetric and positive definite:

$$\begin{aligned}
&\mathbf{g}(\mathbf{u},\mathbf{v}) = \mathbf{g}(\mathbf{v},\mathbf{u})\,, \quad \forall\, \mathbf{u},\mathbf{v} \in \mathbb{TM}\,, \\
&\mathbf{g}(\mathbf{u},\mathbf{u}) \ge 0\,, \quad \forall\, \mathbf{u} \in \mathbb{TM}\,.
\end{aligned}$$

Figure 1.41: Georg Friedrich Bernhard Riemann (1826 - 1866)

1.9.1 Levi Civita connection

The Levi-Civita *connection* on a riemannian manifold is an affine connection which is torsion-free and metric, that is such that

$$\begin{aligned} &i)\ \text{TORS}(\mathbf{v},\mathbf{u}) = \nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v} - [\mathbf{v}\,,\mathbf{u}] = 0\,, \\ &ii)\, \nabla\,\mathbf{g} = 0\,. \end{aligned}$$

Property *ii)* of the Levi-Civita connection states that the *parallel transport* does not affect the metric, i.e. that the covariant derivative of the metric tensor vanishes.

As a consequence the value of the metric tensor evaluated on a pair of vector fields $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1 di\mathbb{M}\,; \mathbb{TM}$ generated by parallel transport of vectors along a curve $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{M})$ is constant:

$$\nabla_{\dot{\mathbf{c}}}(\mathbf{g}\,(\mathbf{u},\mathbf{v})) = \mathbf{g}\,(\nabla_{\dot{\mathbf{c}}}\mathbf{u},\mathbf{v}) + \mathbf{g}\,(\mathbf{u},\nabla_{\dot{\mathbf{c}}}\mathbf{v}) = 0\,.$$

A fortiori the norm of a vector parallel transported along a curve $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{M})$ is constant too.

Proposition 1.9.1 (Basic theorem of riemannian geometry) *Given a riemannian manifold* $\{\mathbb{M}\,,\mathbf{g}\}$ *, there exists a unique* Levi-Civita *connection. The corresponding covariant derivative is expressed in terms of quantities independent of the connection, by the following formula due to* Koszul*:*

$$\begin{aligned} 2\,\mathbf{g}\,(\nabla_{\mathbf{v}}\mathbf{u},\mathbf{w}) = d_{\mathbf{v}}\,(\mathbf{g}\,(\mathbf{u},\mathbf{w})) + d_{\mathbf{u}}\,(\mathbf{g}\,(\mathbf{w},\mathbf{v})) - d_{\mathbf{w}}\,(\mathbf{g}\,(\mathbf{v},\mathbf{u})) \\ + \mathbf{g}\,([\mathbf{v}\,,\mathbf{u}],\mathbf{w}) - \mathbf{g}\,([\mathbf{u}\,,\mathbf{w}],\mathbf{v}) + \mathbf{g}\,([\mathbf{w}\,,\mathbf{v}],\mathbf{u}) \end{aligned}$$

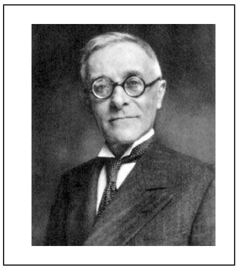

Figure 1.42: Tullio Levi Civita (1873 - 1941)

Proof. Since the connection is metric, we have that:

$$\begin{aligned} d_{\mathbf{w}}(\mathbf{g}(\mathbf{v},\mathbf{u})) &= \mathbf{g}(\nabla_{\mathbf{w}}\mathbf{v},\mathbf{u}) + \mathbf{g}(\mathbf{v},\nabla_{\mathbf{w}}\mathbf{u}), \\ d_{\mathbf{v}}(\mathbf{g}(\mathbf{u},\mathbf{w})) &= \mathbf{g}(\nabla_{\mathbf{v}}\mathbf{u},\mathbf{w}) + \mathbf{g}(\mathbf{u},\nabla_{\mathbf{v}}\mathbf{w}), \\ d_{\mathbf{u}}(\mathbf{g}(\mathbf{w},\mathbf{v})) &= \mathbf{g}(\nabla_{\mathbf{u}}\mathbf{w},\mathbf{v}) + \mathbf{g}(\mathbf{w},\nabla_{\mathbf{u}}\mathbf{v}). \end{aligned}$$

Then, adding the last two equalities and subtracting the first one, by taking into account that the torsion vanishes, so that $\nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v} - [\mathbf{v}\,,\mathbf{u}] = 0$ and the other two generated by ciclic permutation, we get result. ■

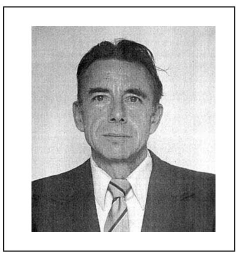

Figure 1.43: Jean-Louis Koszul (1921 -)

Since the connection defined by the Koszul formula is torsion-free and metric preserving existence and uniqueness of the Levi-Civita connection follows.

By Koszul formula, the Christoffel symbols can be expressed in terms of the metric tensor. Indeed we have that

$$2\,\mathbf{g}(\nabla_{\mathbf{e}_i}\mathbf{e}_j,\mathbf{e}_k) = d_{\mathbf{e}_i}(\mathbf{g}(\mathbf{e}_j,\mathbf{e}_k)) + d_{\mathbf{e}_j}(\mathbf{g}(\mathbf{e}_k,\mathbf{e}_i)) - d_{\mathbf{e}_k}(\mathbf{g}(\mathbf{e}_i,\mathbf{e}_j)),$$

that is

$$\boxed{2\,\Gamma_{ij}^{D}\,\mathbf{G}_{CD} = \mathbf{G}_{BC/A} + \mathbf{G}_{CA/B} - \mathbf{G}_{AB/C}\,.}$$

Remark 1.9.1 *The linear isomorphism $\mathbf{g} \in BL\,(\mathbb{TM}\,;\mathbb{T}^*\mathbb{M})$ induced by a metric tensor field doesn't commute in general with the covariant derivative, since by* LEIBNIZ *rule: $\nabla_{\mathbf{v}}(\mathbf{g}\mathbf{w}) = (\nabla_{\mathbf{v}}\mathbf{g})\mathbf{w} + \mathbf{g}(\nabla_{\mathbf{v}}\mathbf{w})$. Then the commutation property holds if and only if the connection is metric:*

$$\nabla_{\mathbf{v}}\mathbf{g} = 0 \iff \nabla_{\mathbf{v}}\mathbf{g} = \mathbf{g}\,\nabla_{\mathbf{v}}\,.$$

1.9.2 Weingarten map

Let $\mathbf{i} \in \mathrm{C}^1(\mathbb{Q}\,;\mathbb{M})$ be the inclusion map of a manifold $\mathbb{Q}$ in a riemannian manifold $\{\mathbb{M}\,,\mathbf{g}_{\mathbb{M}}\}$. A riemannian metric is induced in $\mathbb{Q}$ by setting

$$\mathbf{g}_{\mathbb{Q}} := \mathbf{i}{\downarrow}\mathbf{g}_{\mathbb{M}}\,,$$

that is $\mathbf{g}_{\mathbb{Q}}\,(\mathbf{X},\mathbf{Y}) := \mathbf{g}_{\mathbb{M}}\,(\mathbf{i}{\uparrow}\mathbf{X},\mathbf{i}{\uparrow}\mathbf{Y})\circ\mathbf{i}$ where $\mathbf{i}{\uparrow}\mathbf{X}\circ\mathbf{i} = T\mathbf{i}\circ\mathbf{X}$ and $\mathbf{X},\mathbf{Y} \in \mathbb{TQ}$.

Lemma 1.9.1 (Riemannian embedding) *Let $\{\mathbb{M}\,,\mathbf{g}_{\mathbb{M}}\}$ and $\{\mathbb{Q}\,,\mathbf{g}_{\mathbb{Q}}\}$ be riemannian manifolds with the inclusion map $\mathbf{i} \in \mathrm{C}^1(\mathbb{Q}\,;\mathbb{M})$ and the* LEVI-CIVITA *connections $\nabla_{\mathbb{M}}$ and $\nabla_{\mathbb{Q}}$. Then*

$$\mathbf{g}_{\mathbb{Q}} = \mathbf{i}{\downarrow}\mathbf{g}_{\mathbb{M}} \implies \nabla_{\mathbb{Q}}\mathbf{X}\cdot\mathbf{Y} = \mathbf{P}_{\mathbb{QM}}(\nabla_{\mathbb{M}}\mathbf{i}{\uparrow}\mathbf{X}\cdot\mathbf{i}{\uparrow}\mathbf{Y})\,,\quad \forall\,\mathbf{X},\mathbf{Y} \in \mathbb{TQ}\,,$$

where the linear map $\mathbf{P}_{\mathbb{QM}}(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{i}(\mathbf{x})}\mathbb{M}\,;\mathbb{T}_{\mathbf{i}(\mathbf{x})}(\mathbf{i}(\mathbf{Q})))$ is an orthogonal projector at $\mathbf{i}(\mathbf{x}) \in \mathbb{M}$.

Proof. The expression of the LEVI-CIVITA connection $\nabla_{\mathbb{Q}}$ follows from a direct verification by KOSZUL formula. ■

Definition 1.9.1 (Weingarten map) *Given a vectorial distribution $\Delta \subseteq \mathbb{TM}$, the* WEINGARTEN *map $\mathbf{W} \in \mathrm{C}^1(\Delta^2\,;\mathbb{TM})$ is the tangent-valued bilinear form defined by*

$$\mathbf{W}\,(\mathbf{X},\mathbf{Y}) := \mathbf{\Pi}(\nabla_{\mathbb{M}}\mathbf{X}\cdot\mathbf{Y})\,,\quad \forall\,\mathbf{X},\mathbf{Y} \in \Delta\,,$$

with $\mathbf{\Pi}_{\mathbf{x}} \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{M}\,;\Delta_{\mathbf{x}}^{\perp})$ the $\mathbf{g}_{\mathbb{M}}$-orthogonal projector.

We have the following result.

Lemma 1.9.2 (Symmetry of Weingarten map) *The* WEINGARTEN *map* $\mathbf{W} \in \mathrm{C}^1(\Delta^2\,;\mathbb{TM})$ *is a tensorial bilinear form which is symmetric if and only if the distribution* Δ *is involutive, i.e.*

$$\mathbf{W}(\mathbf{X},\mathbf{Y}) = \mathbf{W}(\mathbf{Y},\mathbf{X}) \iff [\mathbf{X}\,,\mathbf{Y}] \in \Delta\,, \quad \forall\, \mathbf{X},\mathbf{Y} \in \Delta\,.$$

Proof. Tensoriality of the WEINGARTEN map follows from the relation

$$\begin{aligned}\mathbf{\Pi}(\nabla_{\mathbb{M}}(f\,\mathbf{X})\cdot\mathbf{Y}) &= (\nabla_{\mathbb{M}}f\cdot\mathbf{Y})\,\mathbf{\Pi}(\mathbf{X}) + f\,\mathbf{\Pi}(\nabla_{\mathbb{M}}\mathbf{X}\cdot\mathbf{Y})\\ &= f\,\mathbf{\Pi}(\nabla_{\mathbb{M}}\mathbf{X}\cdot\mathbf{Y})\,,\end{aligned}$$

since $\mathbf{\Pi}(\mathbf{X}) = 0$. By expressing the orthogonal projector $\mathbf{\Pi}$ in terms of an orthonormal basis $\{\mathbf{e}_k | k = 1,\ldots,n\}$ of $\Delta^{\perp}_{\mathbf{x}}$, the WEINGARTEN map can be rewritten as:

$$\mathbf{W}(\mathbf{X},\mathbf{Y}) := \sum_{k=1}^{n} \mathbf{g}(\nabla_{\mathbb{M}}\mathbf{X}\cdot\mathbf{Y},\mathbf{e}_k)\,\mathbf{e}_k\,, \quad \forall\, \mathbf{X},\mathbf{Y} \in \Delta\,,$$

Recalling that the LEVI-CIVITA connection is torsion-free, we have that

$$\nabla_{\mathbb{M}}\mathbf{X}\cdot\mathbf{Y} - \nabla_{\mathbb{M}}\mathbf{Y}\cdot\mathbf{X} = [\mathbf{X}\,,\mathbf{Y}]\,,$$

and symmetry holds if and only if $\mathbf{g}([\mathbf{X}\,,\mathbf{Y}],\mathbf{e}_k) = 0$ which is equivalent to require that $[\mathbf{X}\,,\mathbf{Y}] \in \Delta$. ■

Figure 1.44: Julius Weingarten (1836 - 1910)

By the metric property of the LEVI-CIVITA connection it is:

$$\nabla_{\mathbf{Y}}\mathbf{g}(\mathbf{X},\mathbf{e}_k) = \mathbf{g}(\nabla_{\mathbf{Y}}\mathbf{X},\mathbf{e}_k) + \mathbf{g}(\mathbf{X},\nabla_{\mathbf{Y}}\mathbf{e}_k) = 0\,,$$

and the WEINGARTEN map can be rewritten as

$$\mathbf{W}(\mathbf{X},\mathbf{Y}) := -\sum_{k=1}^{n} \mathbf{g}(\mathbf{X}, \nabla_{\mathbb{M}}\mathbf{e}_k \cdot \mathbf{Y})\,\mathbf{e}_k\,, \quad \forall\, \mathbf{X},\mathbf{Y} \in \Delta\,.$$

If the manifold $\mathbb{Q}$ is an hypersurface in $\mathbb{M}$ with inclusion map $\mathbf{i} \in \mathrm{C}^1(\mathbb{Q}\,;\mathbb{M})$, setting $\Delta_{\mathbf{x}} = \mathbb{T}_{\mathbf{i}(\mathbf{x})}(\mathbf{i}(\mathbf{Q}))$, the distribution $\Delta_{\mathbf{x}}^{\perp}$ is one dimensional with orthonormal basis the unit normal $\mathbf{n}(\mathbf{x}) \in \mathbb{T}_{\mathbf{i}(\mathbf{x})}\mathbb{M}$. We may then define the WEINGARTEN tensor field $\mathbf{W} \in \mathrm{C}^1(\mathbb{T}\mathbb{Q}^2\,;\Re)$ as the unique component of the WEINGARTEN map, by

$$\mathbf{W}(\mathbf{X},\mathbf{Y}) := \mathbf{g}(\nabla_{\mathbb{M}}\mathbf{i}{\uparrow}\mathbf{X} \cdot \mathbf{i}{\uparrow}\mathbf{Y}, \mathbf{n})\,, = -\mathbf{g}(\mathbf{i}{\uparrow}\mathbf{X}, \nabla_{\mathbb{M}}\mathbf{n} \cdot \mathbf{i}{\uparrow}\mathbf{Y}) \quad \forall\, \mathbf{X},\mathbf{Y} \in \mathbb{T}\mathbb{Q}\,.$$

1.9.3 Gradient, hessian, divergence and laplacian

- The *gradient* $\nabla f \in \mathrm{C}(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ of a scalar field $f \in \mathrm{C}^1(\mathbb{M}\,;\Re)$ is the vector field associated with the directional derivative according to the pointwise relation

  $$\mathbf{g}(\nabla f, \mathbf{u}) := d_{\mathbf{u}} f\,, \quad \forall\, \mathbf{u} \in \mathbb{T}\mathbb{M} \iff df = \mathbf{g} \circ \nabla f\,.$$

  On the l.h.s the metric tensor is considered to be the symmetric bilinear map $\mathbf{g}(\mathbf{x}) \in BL(\mathbb{T}\mathbb{M}(\mathbf{x}) \times \mathbb{T}\mathbb{M}(\mathbf{x})\,;\Re)$ while on the r.h.s. we have made use of the equivalent characterization $\mathbf{g}(\mathbf{x}) \in BL(\mathbb{T}\mathbb{M}(\mathbf{x})\,;\mathbb{T}^*\mathbb{M}(\mathbf{x}))$.

- The *hessian* $\nabla^2 f \in \mathrm{C}(\mathbb{M}\,;BL(\mathbb{T}\mathbb{M}\,;\mathbb{T}\mathbb{M}))$ of a scalar field $f \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ is the $(1,1)$ tensor field associated with the covariant derivative of the gradient ∇f, according to the relation

  $$\mathbf{g}((\nabla^2 f)\,\mathbf{v}, \mathbf{u}) := \mathbf{g}(\nabla_{\mathbf{v}}(\nabla f), \mathbf{u})\,, \quad \forall\, \mathbf{v},\mathbf{u} : \mathbb{M} \mapsto \mathbb{T}\mathbb{M}\,.$$

- The *laplacian* $\Delta f \in \mathrm{C}(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ of a scalar field $f \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ is the scalar field defined by

  $$\Delta f := \mathrm{div}\,(\nabla f)\,.$$

In a riemannian manifold with the LEVI-CIVITA connection we have that

$$\Delta f = \mathrm{tr}(\nabla(\nabla f))\,.$$

Remark 1.9.2 *The second covariant derivative:*

$$(\nabla d)_{\mathbf{vu}}\, f := (\nabla_{\mathbf{v}}\, df) \cdot \mathbf{u} = d_{\mathbf{v}}\, d_{\mathbf{u}}\, f - d_{(\nabla_{\mathbf{v}}\mathbf{u})}\, f\,.$$

is related to the hessian by the formula

$$(\nabla d)\, f = \mathbf{g}\nabla^2 f\,.$$

Indeed, being $\nabla \mathbf{g} = 0$ *, we have that*

$$\begin{aligned}\mathbf{g}\,(\nabla_{\mathbf{v}}(\nabla f), \mathbf{u}) &= d_{\mathbf{v}}(\mathbf{g}\,(\nabla f, \mathbf{u})) - \mathbf{g}\,(\nabla f, \nabla_{\mathbf{v}}\mathbf{u}) \\ &= d_{\mathbf{v}}\, d_{\mathbf{u}}\, f - d_{(\nabla_{\mathbf{v}}\mathbf{u})}\, f\,.\end{aligned}$$

A connection ∇ in $\mathbb{M}$ is said to be $\boldsymbol{\mu}$-volumetric if the covariant derivative of the volume form $\boldsymbol{\mu} \in BL\,(\mathbb{TM}^3\,;\Re)$ vanishes identically: $\nabla\,\boldsymbol{\mu} = 0$. By the tensoriality of the torsion tensor, we may introduce the linear operator $\mathrm{T{\scriptsize ORS}}(\mathbf{v}) \in BL\,(\mathbb{TM}\,;\mathbb{TM})$ defined pointwise by

$$\mathrm{T{\scriptsize ORS}}(\mathbf{v}) \cdot \mathbf{u} = \mathrm{{\scriptsize TORS}}(\mathbf{v}, \mathbf{u})\,, \quad \forall\, \mathbf{v}, \mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})\,.$$

Then we may state the following result.

Lemma 1.9.3 *Let* $\mathbb{M}$ *be a manifold,* $\boldsymbol{\mu} \in BL\,(\mathbb{TM}^3\,;\Re)$ *a volume form, and* ∇ *a volumetric connection in* $\mathbb{M}$ *. The divergence of a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is then given by*

$$\operatorname{div} \mathbf{v} = \operatorname{tr}(\nabla \mathbf{v} + \mathrm{T{\scriptsize ORS}}(\mathbf{v}))\,.$$

Proof. Recalling that the torsion of the connection ∇ is defined by the formula $\mathrm{{\scriptsize TORS}}(\mathbf{v}, \mathbf{a}) = \nabla_{\mathbf{v}}\mathbf{a} - \nabla_{\mathbf{a}}\mathbf{v} - \mathcal{L}_{\mathbf{v}}\mathbf{a}$, we have, for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$, that

$$\begin{aligned}
&(\mathcal{L}_{\mathbf{v}}\boldsymbol{\mu})(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \\
&\mathcal{L}_{\mathbf{v}}\,(\boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \mathbf{c})) - \boldsymbol{\mu}(\mathcal{L}_{\mathbf{v}}\mathbf{a}, \mathbf{b}, \mathbf{c}) - \boldsymbol{\mu}(\mathbf{a}, \mathcal{L}_{\mathbf{v}}\mathbf{b}, \mathbf{c}) - \boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \mathcal{L}_{\mathbf{v}}\mathbf{c}) = \\
&\nabla_{\mathbf{v}}\,(\boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \mathbf{c})) - \boldsymbol{\mu}(\nabla_{\mathbf{v}}\mathbf{a}, \mathbf{b}, \mathbf{c}) - \boldsymbol{\mu}(\mathbf{a}, \nabla_{\mathbf{v}}\mathbf{b}, \mathbf{c}) - \boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \nabla_{\mathbf{v}}\mathbf{c}) \\
&\qquad + \boldsymbol{\mu}(\nabla_{\mathbf{a}}\mathbf{v}, \mathbf{b}, \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \nabla_{\mathbf{b}}\mathbf{v}, \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \nabla_{\mathbf{c}}\mathbf{v}) \\
&\qquad + \boldsymbol{\mu}(\mathrm{{\scriptsize TORS}}(\mathbf{v}, \mathbf{a}), \mathbf{b}, \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \mathrm{{\scriptsize TORS}}(\mathbf{v}, \mathbf{b}), \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \mathrm{{\scriptsize TORS}}(\mathbf{v}, \mathbf{c})) = \\
&(\nabla_{\mathbf{v}}\,\boldsymbol{\mu})(\mathbf{a}, \mathbf{b}, \mathbf{c}) + \boldsymbol{\mu}(\nabla_{\mathbf{a}}\mathbf{v}, \mathbf{b}, \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \nabla_{\mathbf{b}}\mathbf{v}, \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \nabla_{\mathbf{c}}\mathbf{v}) \\
&+\boldsymbol{\mu}(\mathrm{T{\scriptsize ORS}}(\mathbf{v}) \cdot \mathbf{a}, \mathbf{b}, \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \mathrm{T{\scriptsize ORS}}(\mathbf{v}) \cdot \mathbf{b}, \mathbf{c}) + \boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \mathrm{T{\scriptsize ORS}}(\mathbf{v}) \cdot \mathbf{c}) = \\
&\qquad = (\nabla_{\mathbf{v}}\,\boldsymbol{\mu})(\mathbf{a}, \mathbf{b}, \mathbf{c}) + (\operatorname{tr}(\nabla \mathbf{v} + \mathrm{T{\scriptsize ORS}}(\mathbf{v}))\,\boldsymbol{\mu}(\mathbf{a}, \mathbf{b}, \mathbf{c})\,,
\end{aligned}$$

or briefly

$$\mathcal{L}_{\mathbf{v}}\boldsymbol{\mu} = \nabla_{\mathbf{v}}\,\boldsymbol{\mu} + \mathrm{tr}(\nabla\mathbf{v} + \mathrm{TORS}(\mathbf{v}))\,\boldsymbol{\mu}\,.$$

The result follows since $\mathcal{L}_{\mathbf{v}}\boldsymbol{\mu} = (\mathrm{div}\,\mathbf{v})\,\boldsymbol{\mu}$ and $\nabla_{\mathbf{v}}\,\boldsymbol{\mu} = 0$ by assumption. ■

If the connection is torsion-free, the formula becomes

$$\mathrm{div}\,\mathbf{v} = \mathrm{tr}(\nabla\mathbf{v})\,.$$

This result then holds in any riemannian manifold endowed with the LEVI-CIVITA connection which is torsion-free and, being metric, is also volumetric with respect to the volume form associated with the metric.

1.9.4 Killing's vector fields

A connection ∇ in a riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$ is said to be $\mathbf{g}$-metric if the covariant derivative of the metric tensor $\mathbf{g} \in BL\,(\mathbb{TM}^2\,;\Re)$ vanishes identically:

$$\nabla\,\mathbf{g} = 0\,.$$

The *stretching* $\mathrm{DEF}(\mathbf{v}) \in BL\,(\mathbb{TM}\,;\mathbb{TM})$ of a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ is the $\mathbf{g}$-symmetric operator defined by

$$\mathbf{g} \circ \mathrm{DEF}(\mathbf{v}) := \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\,\mathbf{g}\,.$$

The *metric gradient* $\mathrm{MET}(\mathbf{v}) \in BL\,(\mathbb{TM}\,;\mathbb{TM})$, is the $\mathbf{g}$-symmetric operator which, by the tensoriality of the nabla operator, is defined pointwise by

$$\mathbf{g} \circ \mathrm{MET}(\mathbf{v}) := \tfrac{1}{2}\nabla_{\mathbf{v}}\,\mathbf{g}\,.$$

EULER*'s rigidity operator* $\mathrm{EUL}(\mathbf{v}) \in BL\,(\mathbb{TM}\,;\mathbb{TM})$ is the $\mathbf{g}$-symmetric operator defined by

$$\mathrm{EUL}(\mathbf{v}) := \mathrm{sym}\,\nabla\mathbf{v}\,.$$

We have the following result.

Proposition 1.9.2 *Let* $\{\mathbb{M}\,,\mathbf{g}\}$ *be a riemannian manifold and* ∇ *a connection in* $\mathbb{M}$. *The stretching of a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is given by*

$$\mathrm{DEF}(\mathbf{v}) = \mathrm{MET}(\mathbf{v}) + \mathrm{EUL}(\mathbf{v}) + \mathrm{sym}\,\mathrm{TORS}(\mathbf{v})\,.$$

Proof. Applying the LEIBNIZ rule to the LIE derivative and to the covariant derivative, we have that, for any $\mathbf{v}, \mathbf{u}, \mathbf{w} \in C^1(\mathbb{M}; \mathbb{TM})$:

$$(\mathcal{L}_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) = \mathcal{L}_{\mathbf{v}}\, (\mathbf{g}\,(\mathbf{u}, \mathbf{w})) - \mathbf{g}\,(\mathcal{L}_{\mathbf{v}}\mathbf{u}, \mathbf{w}) - \mathbf{g}\,(\mathbf{u}, \mathcal{L}_{\mathbf{v}}\mathbf{w})\,,$$
$$(\nabla_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) = \nabla_{\mathbf{v}}\, (\mathbf{g}\,(\mathbf{u}, \mathbf{w})) - \mathbf{g}\,(\nabla_{\mathbf{v}}\mathbf{u}, \mathbf{w}) - \mathbf{g}\,(\mathbf{u}, \nabla_{\mathbf{v}}\mathbf{w})\,.$$

Since the LIE derivative and the covariant derivative of a scalar field coincide, we also have that $\mathcal{L}_{\mathbf{v}}\, (\mathbf{g}\,(\mathbf{u}, \mathbf{w})) = \nabla_{\mathbf{v}}\, (\mathbf{g}\,(\mathbf{u}, \mathbf{w}))$ and hence:

$$\begin{aligned}(\mathcal{L}_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) = (\nabla_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) &+ \mathbf{g}\,(\nabla_{\mathbf{v}}\mathbf{u}, \mathbf{w}) + \mathbf{g}\,(\mathbf{u}, \nabla_{\mathbf{v}}\mathbf{w}) \\ &- \mathbf{g}\,(\mathcal{L}_{\mathbf{v}}\mathbf{u}, \mathbf{w}) - \mathbf{g}\,(\mathbf{u}, \mathcal{L}_{\mathbf{v}}\mathbf{w})\,.\end{aligned}$$

Moreover, since $\mathrm{TORS}(\mathbf{v}, \mathbf{u}) := (\nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v}) - [\mathbf{v}\,, \mathbf{u}]$ we may write

$$\begin{aligned}(\mathcal{L}_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) = (\nabla_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) &+ \mathbf{g}\,(\mathrm{TORS}(\mathbf{v}, \mathbf{u}), \mathbf{w}) + \mathbf{g}\,(\nabla_{\mathbf{u}}\mathbf{v}, \mathbf{w}) \\ &+ \mathbf{g}\,(\mathrm{TORS}(\mathbf{v}, \mathbf{w}), \mathbf{u}) + \mathbf{g}\,(\nabla_{\mathbf{w}}\mathbf{v}, \mathbf{u})\,,\end{aligned}$$

which, in terms of the operator $\mathrm{TORS}(\mathbf{v}) \in BL\,(\mathbb{TM}; \mathbb{TM})$ is given by

$$\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}}\, \mathbf{g}) = \tfrac{1}{2}(\nabla_{\mathbf{v}}\, \mathbf{g}) + \mathbf{g} \circ (\mathrm{sym}\, \nabla\mathbf{v}) + \mathbf{g} \circ (\mathrm{sym}\, \mathrm{TORS}(\mathbf{v}))\,.$$

Recalling the definitions $\mathbf{g} \circ \mathrm{DEF}(\mathbf{v}) = \frac{1}{2}(\mathcal{L}_{\mathbf{v}}\, \mathbf{g})$ and $\mathbf{g} \circ \mathrm{MET}(\mathbf{v}) = \frac{1}{2}\nabla_{\mathbf{v}}\, \mathbf{g}$ we get the result. ■

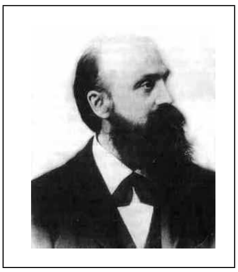

Figure 1.45: Wilhelm Karl Joseph Killing (1847 - 1923)

A vector field $\mathbf{v} \in C^1(\mathbb{M}; \mathbb{TM})$ on a riemannian manifold $\{\mathbb{M}\,, \mathbf{g}\}$ is a KILLING *field* if the metric tensor is dragged along its flow:

$$\mathcal{L}_{\mathbf{v}}\, \mathbf{g} = 0\,.$$

A KILLING vector field is also called an *infinitesimal isometry.*

Proposition 1.9.3 (Euler-Killing formula) *Let* $\{\mathbb{M}, \mathbf{g}\}$ *be a riemannian manifold endowed with the* LEVI-CIVITA *connection. Then the* LIE *derivative of the metric tensor is expressed in terms of the covariant derivative by the formula*

$$\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}}\,\mathbf{g}) = \mathbf{g} \circ (\operatorname{sym}\nabla\mathbf{v}), \quad \forall\,\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM}),$$

and the EULER-KILLING *condition may be written as* $\mathrm{DEF}(\mathbf{v}) = \operatorname{sym}\nabla\mathbf{v} = 0$ *.*

Proof. The statement follows from proposition 1.9.2 being the LEVI-CIVITA connection metric-preserving and torsion-free: $\mathrm{MET} = 0$ and $\mathrm{TORS} = 0$. ■

KILLING condition extends EULER's condition in the euclidean space to the more general case of riemannian manifolds. In the euclidean space endowed with the canonical connection, the LEVI-CIVITA parallel transport is simply the translation and the covariant derivative is the ordinary derivative. Hence KILLING's condition reduces to the classical one due to EULER.

1.9.5 Geodesics

- A *pathwork of manifolds* $\mathbb{M}$ is a finite family $\{\mathbb{M}_\alpha \mid \alpha \in \mathcal{A}\}$ whose elements are regular manifolds $\mathbb{M}_\alpha$, possibly with boundary, all modeled on the same BANACH space and with the element manifolds intersecting pairwise only at their boundaries. A pathwork $\{\mathbb{M}, \mathbf{g}\}$ of riemannian manifolds is a pathwork of manifolds $\mathbb{M}$ endowed with a metric tensor field which is regular in each element and may undergo finite jumps at the interfaces between the elements.

The simplest picture of a pathwork of manifolds is a parallelepiped in the euclidean 3-space (a candy box). In geometrical optics a pathwork of riemannian manifolds is naturally provided by optical media with different refraction properties.

Let us consider a pathwork of riemannian manifolds $\{\mathbb{M}, \mathbf{g}\}$ and a path $\gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$ which is piecewise regular according to a finite partition $\mathcal{T}(I)$. We denote by ∂I the boundary chain of the interval I and by $\partial\mathcal{T}(I)$, $\mathcal{I}(I)$ the union of the boundary chains of the elements and the family of interfaces between the elements of the partition $\mathcal{T}(I)$.

The *speed* or *velocity* of the path at a regular point $t \in I$ is given by $\mathbf{v}(\gamma(t)) = \mathbf{v}_t := \partial_{\tau=t}\,\gamma(\tau)$, and the *scalar speed* is its $\mathbf{g}$-norm: $\|\mathbf{v}_t\|_{\mathbf{g}} := \sqrt{\mathbf{g}\,(\mathbf{v}_t, \mathbf{v}_t)}$. We will assume that $\|\mathbf{v}_t\|_{\mathbf{g}} \in \mathrm{C}^0(I\,;\Re)$, i.e. that the scalar speed of the path is continuous over the whole interval of definition.

- The *length* of the path $\gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$ is the integral of its scalar speed:

$$\ell(\gamma) := \int_{\mathcal{T}(I)} \sqrt{\mathbf{g}\,(\mathbf{v}_t, \mathbf{v}_t)}\, dt\,,$$

  and is independent of the parametrization.

- The *energy* of the path $\gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$ is the integral of half its squared scalar speed:

$$\mathcal{E}(\gamma) := \int_{\mathcal{T}(I)} \tfrac{1}{2}\,\mathbf{g}\,(\mathbf{v}_t, \mathbf{v}_t)\, dt\,,$$

  which is dependent on the parametrization.

Let $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ be a flow with velocity field $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$.

- A path $\gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$ is said to have a *stationary length* if for any flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$:

$$\partial_{\lambda=0}\, \ell(\boldsymbol{\varphi}_\lambda \circ \gamma) := \partial_{\lambda=0} \int_{\mathcal{T}(I)} \sqrt{(\boldsymbol{\varphi}_\lambda{\downarrow}\mathbf{g})(\mathbf{v}_t, \mathbf{v}_t)}\, dt = \int_{\partial I} \frac{\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\,(\mathbf{v}_t, \mathbf{v}_{\boldsymbol{\varphi}_t})}{\sqrt{\mathbf{g}\,(\mathbf{v}_t, \mathbf{v}_t)}}\, dt\,,$$

where $\mathbf{v}_{\boldsymbol{\varphi}_t} := \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}(\mathbf{v}_t))$ and

$$(\boldsymbol{\varphi}_\lambda{\downarrow}\mathbf{g})(\mathbf{v}_t, \mathbf{v}_t) := \mathbf{g}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))}(\boldsymbol{\varphi}_\lambda{\uparrow}\mathbf{v}_t, \boldsymbol{\varphi}_\lambda{\uparrow}\mathbf{v}_t)\,.$$

The stationarity condition states that, if the path is dragged by the flow into the path $\boldsymbol{\varphi}_\lambda \circ \gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$, the initial rate of variation of the length, due to the variation of the scalar speed field (the l.h.s.), is equal to the gap of equiprojectivity of the flow velocity at the end points (the r.h.s.). A path $\gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$ with a stationary length is called a *geodesic*.

The strip fastened around a candy box is a geodesic. By FERMAT principle, the light ray thru optical media is a geodesic.

- A path $\gamma \in \mathrm{C}^1(I\,;\mathbb{M})$ is said to have a *stationary energy* if for any flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$:

$$\partial_{\lambda=0}\, \mathcal{E}(\boldsymbol{\varphi}_\lambda \circ \gamma) := \partial_{\lambda=0} \int_{\mathcal{T}(I)} \tfrac{1}{2}\,(\boldsymbol{\varphi}_\lambda{\downarrow}\mathbf{g})(\mathbf{v}_t, \mathbf{v}_t)\, dt = \int_{\partial I} \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\,(\mathbf{v}_t, \mathbf{v}_{\boldsymbol{\varphi}_t})\, dt\,.$$

Proposition 1.9.4 (Energy characterization of geodesics) *In a pathwork of riemannian manifolds* $\{\mathbb{M}\,,\mathbf{g}\}$ *, a geodesic* $\gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$ *, parametrized with a constant scalar speed, has a stationary energy. This is equivalent to the differential condition:*

$$\tfrac{1}{2}\,(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t) = \partial_{\tau=t}\;\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}\,(\mathbf{v}_\tau,\mathbf{v}_{\boldsymbol{\varphi}_\tau})\,,$$

in the elements of the pathwork, and to the jump conditions

$$\langle\, [[\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\mathbf{v}_t]], \mathbf{v}_{\boldsymbol{\varphi}_t}\,\rangle = 0\,,$$

at singular points.

Proof. Let the path $\gamma \in \mathrm{C}^1(\mathcal{T}(I)\,;\mathbb{M})$ have constant scalar speed $\alpha > 0$. Then

$$\partial_{\lambda=0}\sqrt{(\boldsymbol{\varphi}_\lambda{\downarrow}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t)} = \frac{\partial_{\lambda=0}\,(\boldsymbol{\varphi}_\lambda{\downarrow}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t)}{2\,\sqrt{\mathbf{g}(\mathbf{v}_t,\mathbf{v}_t)}} = \tfrac{1}{2\alpha}\,(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t)\,.$$

Stationarity of the lenght may then be written as

$$\tfrac{1}{2\alpha}\int_{\mathcal{T}(I)}(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t)\,dt = \tfrac{1}{\alpha}\int_{\partial I}\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\,(\mathbf{v}_t,\mathbf{v}_{\boldsymbol{\varphi}_t})\,dt\,,$$

which is equivalent to

$$\begin{aligned}\int_{\mathcal{T}(I)}\tfrac{1}{2}\,(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t)\,dt = &\int_{\mathcal{T}(I)}\partial_{\tau=t}\;\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}\,(\mathbf{v}_\tau,\mathbf{v}_{\boldsymbol{\varphi}_\tau})\,dt\\ &-\int_{\mathcal{I}(I)}\langle\,[[\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}\mathbf{v}_\tau]],\mathbf{v}_{\boldsymbol{\varphi}_\tau}\,\rangle\,dt\,,\end{aligned}$$

and, by the arbitrarity of the flow, to the differential and jump conditions in the statement. ■

Let us now consider, in a riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$ with a connection ∇, a curve $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{M})$ thru $\mathbf{x} = \mathbf{c}(0)$ with speed $\mathbf{w} = \partial_{\lambda=0}\,\mathbf{c}(\lambda)$, and a vector $\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$.

We recall that the *horizontal lift* of a functional $f \in \mathrm{C}^1(\mathbb{TM}\,;\Re)$ at $\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ is defined by:

$$\langle\, d_{\mathrm{B}}f(\mathbf{v}),\mathbf{w}\,\rangle := \partial_{\lambda=0}\;f(\mathbf{c}(\lambda){\Uparrow}\,\mathbf{v})\,,\quad \forall\,\mathbf{w} \in \mathbb{T}_{\boldsymbol{\tau}(\mathbf{v})}\mathbb{M}\,.$$

The definition is well-posed since the r.h.s. depends linearly on $\mathbf{w} \in \mathbb{T}_{\boldsymbol{\tau}(\mathbf{v})}\mathbb{M}$ for any fixed $\mathbf{v} \in \mathbb{TM}$. Accordingly, the *horizontal lift* of the quadratic metric form $q_{\mathbf{g}} \in \mathrm{C}^1(\mathbb{TM};\Re)$ defined by $q_{\mathbf{g}}(\mathbf{v}) := \mathbf{g}(\mathbf{v},\mathbf{v})$ at $\mathbf{v} \in \mathbb{TM}$ is given by:

$$\langle d_{\mathrm{B}} q_{\mathbf{g}}(\mathbf{v}), \mathbf{w} \rangle := \partial_{\lambda=0}\ \mathbf{g}_{\mathbf{c}(\lambda)}(\mathbf{c}(\lambda)\Uparrow \mathbf{v}, \mathbf{c}(\lambda)\Uparrow \mathbf{v}), \quad \forall\, \mathbf{w} \in \mathbb{T}_{\boldsymbol{\tau}(\mathbf{v})}\mathbb{M}.$$

Proposition 1.9.5 (Differential equation of a geodesic) *In a pathwork of riemannian manifolds* $\{\mathbb{M}, \mathbf{g}\}$ *with a connection* ∇*, a constant-speed curve* $\gamma \in \mathrm{C}^1(I;\mathbb{M})$ *is a geodesic if and only if it fulfills the differential equation:*

$$\partial_{\tau=t}\ \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \boldsymbol{\chi}_{\tau,t}\Uparrow \mathbf{v}_{\boldsymbol{\varphi}_t}) = \tfrac{1}{2}\langle d_{\mathrm{B}} q_{\mathbf{g}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle + \langle (\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\mathbf{v}_t)\mathrm{T{\small ORS}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle,$$

which may be written as

$$\nabla_{\mathbf{v}_t}(\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v})}\mathbf{v}) = \tfrac{1}{2}\, d_{\mathrm{B}} q_{\mathbf{g}}(\mathbf{v}_t) + (\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\mathbf{v}_t)\mathrm{T{\small ORS}}(\mathbf{v}_t),$$

at regular points, and the jump conditions

$$\langle [[\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\mathbf{v}_t]], \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle = 0,$$

at singular points.

Proof. We have that

$$\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{g})(\mathbf{v}_t, \mathbf{v}_t) = \tfrac{1}{2}\,\partial_{\lambda=0}\ \mathbf{g}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))}(\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t, \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t).$$

We extend the velocity along the path to a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ by pushing it along the flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M};\mathbb{M})$ according to the relation:

$$\mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))) := \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t.$$

Then, writing $\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t = \boldsymbol{\varphi}_\lambda\!\Uparrow \boldsymbol{\varphi}_\lambda\!\Downarrow \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t$ and applying L{\small EIBNIZ} rule, we get

$$\begin{aligned}
\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{g})(\mathbf{v}_t, \mathbf{v}_t) &= \tfrac{1}{2}\,\partial_{\lambda=0}\ \mathbf{g}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))}(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}_t, \boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}_t) \\
&\quad + \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\partial_{\lambda=0}\ \boldsymbol{\varphi}_\lambda\!\Downarrow \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t, \mathbf{v}_t) \\
&= \tfrac{1}{2}\langle d_{\mathrm{B}} q_{\mathbf{g}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle + \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_{\boldsymbol{\varphi}_t}}\mathbf{v}, \mathbf{v}_t).
\end{aligned}$$

Similarly, defining the trajectory-flow $\boldsymbol{\chi}_{\tau,t} \in \mathrm{C}^1(\mathbb{M};\mathbb{M})$ by $\boldsymbol{\chi}_{\tau,t} \circ \gamma_t = \gamma_\tau$, we have that

$$\begin{aligned}
\partial_{\tau=t}\ \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \mathbf{v}_{\boldsymbol{\varphi}_\tau}) &= \partial_{\tau=t}\ \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \boldsymbol{\chi}_{\tau,t}\Uparrow \boldsymbol{\chi}_{\tau,t}\Downarrow \mathbf{v}_{\boldsymbol{\varphi}_\tau}) \\
&= \partial_{\tau=t}\ \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \boldsymbol{\chi}_{\tau,t}\Uparrow \mathbf{v}_{\boldsymbol{\varphi}_t}) + \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_t) \\
&= \langle \nabla_{\mathbf{v}_t}(\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v})}\mathbf{v}), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle + \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_t).
\end{aligned}$$

By definition of the vector field $\mathbf{v} \in C^1(\mathbb{M}\,;\mathbb{TM})$ we have that $[\mathbf{v}_{\boldsymbol{\varphi}}\,, \mathbf{v}] = 0$ and hence

$$\text{TORS}(\mathbf{v}) \cdot \mathbf{v}_{\boldsymbol{\varphi}} = \text{TORS}(\mathbf{v}, \mathbf{v}_{\boldsymbol{\varphi}}) = \nabla_{\mathbf{v}} \mathbf{v}_{\boldsymbol{\varphi}} - \nabla_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{v}\,.$$

The differential condition of Proposition 1.9.4 may then be written as

$$\tfrac{1}{2}\,\langle\, d_{\text{B}} q_{\mathbf{g}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\,\rangle = \langle\, \nabla_{\mathbf{v}_t}(\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v})}\mathbf{v}), \mathbf{v}_{\boldsymbol{\varphi}_t}\,\rangle + \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\mathbf{v}_t, \text{TORS}(\mathbf{v}_t) \cdot \mathbf{v}_{\boldsymbol{\varphi}_t})\,.$$

and the statement is proven. ■

Proposition 1.9.6 (Geodesics in riemannian manifolds) *In a pathwork of riemannian manifolds* $\{\mathbb{M}\,, \mathbf{g}\}$ *with a torsion-free connection* ∇ *, a geodesic* $\gamma \in C^1(I\,;\mathbb{M})$ *fulfills the differential equation:*

$$\nabla_{\mathbf{v}_t}(\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v})}\mathbf{v}) = \tfrac{1}{2}\, d_{\text{B}} q_{\mathbf{g}}(\mathbf{v}_t)\,,$$

in the elements of the pathwork, and the jump conditions

$$\langle\, [[\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\mathbf{v}_t]], \mathbf{v}_{\boldsymbol{\varphi}_t}\,\rangle = 0\,,$$

at singular points. If the connection is LEVI-CIVITA*, the differential equation becomes*

$$\nabla_{\mathbf{v}_t}\mathbf{v} = 0\,.$$

Proof. The first statement follows directly from proposition 1.9.5. In a riemannian manifold $\{\mathbb{M}\,, \mathbf{g}\}$ the LEVI-CIVITA connection is torsion-free and metric-preserving. In a metric connection, being $\nabla \mathbf{g} = 0$, the norm is preserved by the parallel transport, so that

$$d_{\text{B}} q_{\mathbf{g}}(\mathbf{v}_t) = \partial_{\lambda=0}\, \mathbf{g}(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}_t, \boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}_t) = 0\,.$$

Moreover, for any $\mathbf{w} \in C^1(\mathbb{M}\,;\mathbb{TM})$, it is:

$$\langle\, \nabla_{\mathbf{v}_t}(\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v})}\mathbf{v}), \mathbf{w}\,\rangle = d_{\mathbf{v}_t}\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v})}(\mathbf{v}, \mathbf{w}) - \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\mathbf{v}_t, \nabla_{\mathbf{v}_t}\mathbf{w}) = \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}, \mathbf{w})\,.$$

Hence the latter statement follows from the former one. ■

The next result provides another proof of the second statement in proposition 1.9.6 extending a formula in [119], [91].

Proposition 1.9.7 (First variation of the energy) *In a pathwork of riemannian manifolds* $\{\mathbb{M}\,,\mathbf{g}\}$ *the first variation of the energy of a path is given by*

$$\partial_{\lambda=0}\;\mathcal{E}(\boldsymbol{\varphi}_\lambda\circ\gamma) := \partial_{\lambda=0}\int_I \tfrac{1}{2}\,(\boldsymbol{\varphi}_\lambda{\downarrow}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t)\,dt$$

$$= \int_I (\tfrac{1}{2}\,\langle\, d_{\mathrm{B}} q_{\mathbf{g}}(\mathbf{v}_t),\mathbf{v}_{\boldsymbol{\varphi}_t}\,\rangle + \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_{\boldsymbol{\varphi}_t}}\mathbf{v},\mathbf{v}_t))\,dt\,,$$

which, in the LEVI-CIVITA *connection, becomes:*

$$\partial_{\lambda=0}\;\mathcal{E}(\boldsymbol{\varphi}_\lambda\circ\gamma) = -\int_{\mathcal{T}(I)} \mathbf{g}\,(\nabla_{\mathbf{v}_t}\mathbf{v},\mathbf{v}_{\boldsymbol{\varphi}_t})\,dt - \int_{\mathcal{I}(I)} \langle\,[[\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\mathbf{v}_t]],\mathbf{v}_{\boldsymbol{\varphi}_t}\,\rangle\,dt\,.$$

Proof. The first formula follows from the proof of proposition 1.9.5. The latter is then deduced by observing that in the LEVI-CIVITA connection: $d_{\mathrm{B}} q_{\mathbf{g}}(\mathbf{v}_t) = 0$ and $\mathrm{TORS}(\mathbf{v},\mathbf{v}_{\boldsymbol{\varphi}}) = \nabla_{\mathbf{v}}\mathbf{v}_{\boldsymbol{\varphi}} - \nabla_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{v} = 0$, so that

$$\partial_{\lambda=0}\int_I (\boldsymbol{\varphi}_\lambda{\downarrow}\mathbf{g})(\mathbf{v}_t,\mathbf{v}_t)\,dt = \int_I \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_{\boldsymbol{\varphi}_t}}\mathbf{v},\mathbf{v}_t)\,dt = \int_I \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}_{\boldsymbol{\varphi}},\mathbf{v}_t)\,dt\,.$$

Moreover, at regular points in I, we have:

$$\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}_{\boldsymbol{\varphi}},\mathbf{v}_t) = d_{\mathbf{v}_t}(\mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\mathbf{v}_t,\mathbf{v}_{\boldsymbol{\varphi}_t}) - \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v},\mathbf{v}_{\boldsymbol{\varphi}_t})\,.$$

so that an integration by parts in each element of the partition $\mathcal{T}(I)$ yields the result. ■

Remark 1.9.3 *If a vector field* $\mathbf{w}\in\mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is parallel transported along the geodesic curve, that is* $\nabla_{\mathbf{v}_t}\mathbf{w} = 0$ *for all* $t\in I$, *its inner product with the tangent field* $\mathbf{v}\in\mathrm{C}^1(\gamma\,;\mathbb{T}\gamma)$ *is constant:*

$$d_{\mathbf{v}_t}(\mathbf{g}\,(\mathbf{w},\mathbf{v})) = \mathbf{g}\,(\nabla_{\mathbf{v}_t}\mathbf{w},\mathbf{v}) + \mathbf{g}\,(\mathbf{w},\nabla_{\mathbf{v}_t}\mathbf{v}) = 0\,.$$

Since the norms of the parallel transported fields $\mathbf{v}$ *and* $\mathbf{w}$ *are constant along the geodesic, the cosinus of the angle between the parallel transported vector and the tangent to the geodesic curve is constant too. Short geodesics, joining two sufficiently near points, are curves of minimal length [119], [91], [143].*

Remark 1.9.4 *The differential condition of stationary length may be also formulated as follows. Let us define the one-form* $\boldsymbol{\theta}_{\mathbf{g}}$ *on* $\mathbb{TM}$ *as*

$$\langle \boldsymbol{\theta}_{\mathbf{g}}(\mathbf{v}), \delta\mathbf{v} \rangle := \langle \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v})} \cdot \mathbf{v}, d\boldsymbol{\tau}(\mathbf{v}) \cdot \delta\mathbf{v} \rangle \,, \quad \forall\, \mathbf{v} \in \mathbb{TM} \,, \quad \forall\, \delta\mathbf{v} \in \mathbb{T}_{\mathbf{v}}\mathbb{TM} \,.$$

We have that $\boldsymbol{\tau}(\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t) = \boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))$ *Moreover, for any flow* $\boldsymbol{\psi}_\lambda \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{TM})$ *which projects to a flow* $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$*, it is* $T\boldsymbol{\tau}(\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t) \cdot \boldsymbol{\psi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t = \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t$ *. Then we have that*

$$\begin{aligned}
(\boldsymbol{\psi}_\lambda\!\downarrow\!\boldsymbol{\theta}_{\mathbf{g}})(\mathbf{v}_t) \cdot \dot{\mathbf{v}}_t &= \langle \boldsymbol{\theta}_{\mathbf{g}}(\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t), \boldsymbol{\psi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t \rangle \\
&= \langle \mathbf{g}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))} \cdot \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t, T\boldsymbol{\tau}(\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t) \cdot \boldsymbol{\psi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t \rangle \\
&= \langle \mathbf{g}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))} \cdot \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t, \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t \rangle = (\boldsymbol{\varphi}_\lambda\!\downarrow\!\mathbf{g})(\mathbf{v}_t, \mathbf{v}_t) \,, \\
\langle \boldsymbol{\theta}_{\mathbf{g}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\psi}}(\mathbf{v}_t) \rangle &= \langle \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)} \cdot \mathbf{v}_t, T\boldsymbol{\tau}(\mathbf{v}_t) \cdot \mathbf{v}_{\boldsymbol{\psi}}(\mathbf{v}_t) \rangle = \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_t)}\,(\mathbf{v}_t, \mathbf{v}_{\boldsymbol{\varphi}_t}) \,,
\end{aligned}$$

and the stationarity condition may be written

$$\partial_{\lambda=0} \int_{\boldsymbol{\psi}_\lambda(\Gamma)} \tfrac{1}{2}\, \boldsymbol{\theta}_{\mathbf{g}} = \int_{\partial\Gamma} \boldsymbol{\theta}_{\mathbf{g}} \cdot \mathbf{v}_{\boldsymbol{\psi}} \iff \int_{\Gamma} \tfrac{1}{2}\, \mathcal{L}_{\mathbf{v}_{\boldsymbol{\psi}}} \boldsymbol{\theta}_{\mathbf{g}} = \int_{\Gamma} d(\boldsymbol{\theta}_{\mathbf{g}} \cdot \mathbf{v}_{\boldsymbol{\psi}}) \,.$$

where $\Gamma \in \mathrm{C}^1(I\,;\mathbb{TM})$ *is the lifted curve of* $\gamma \in \mathrm{C}^1(I\,;\mathbb{M})$*, defined by* $\Gamma(t) = \mathbf{v}_t$ *.*

By the homotopy formula

$$\mathcal{L}_{\mathbf{v}_{\boldsymbol{\psi}}} \boldsymbol{\theta}_{\mathbf{g}} = d(\boldsymbol{\theta}_{\mathbf{g}} \cdot \mathbf{v}_{\boldsymbol{\psi}}) + (d\boldsymbol{\theta}_{\mathbf{g}}) \cdot \mathbf{v}_{\boldsymbol{\psi}} \,,$$

and the arbitrarity of the virtual velocity field $\mathbf{v}_{\boldsymbol{\psi}} = \partial_{\lambda=0}\ \boldsymbol{\psi}_\lambda \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$*, the differential condition of stationarity is given by*

$$d\boldsymbol{\theta}_{\mathbf{g}}(\mathbf{v}_t) \cdot \mathbf{v}_{\boldsymbol{\psi}}(\mathbf{v}_t) \cdot \dot{\mathbf{v}}_t = d_{\dot{\mathbf{v}}_t}(\boldsymbol{\theta}_{\mathbf{g}} \cdot \mathbf{v}_{\boldsymbol{\psi}})(\mathbf{v}_t) \,.$$

The tensoriality of the exterior derivative ensures that PALAIS *formula can be applied by extending the vector* $\dot{\mathbf{v}}_t \in \mathbb{T}_{\mathbf{v}_t}\Gamma$ *to a vector field* $\dot{\mathcal{F}} \in \mathrm{C}^1(\mathbb{TM}\,;\mathbb{T}^2\mathbb{M})$ *such that* $\dot{\mathcal{F}}(\mathbf{v}_t) = \dot{\mathbf{v}}_t$ *:*

$$d\boldsymbol{\theta}_{\mathbf{g}}(\mathbf{v}_t) \cdot \mathbf{v}_{\boldsymbol{\psi}}(\mathbf{v}_t) \cdot \dot{\mathbf{v}}_t = d_{\mathbf{v}_{\boldsymbol{\psi}}(\mathbf{v}_t)}(\boldsymbol{\theta}_{\mathbf{g}} \cdot \dot{\mathcal{F}})(\mathbf{v}_t) - d_{\dot{\mathbf{v}}_t}(\boldsymbol{\theta}_{\mathbf{g}} \cdot \mathbf{v}_{\boldsymbol{\psi}})(\mathbf{v}_t) - (\boldsymbol{\theta}_{\mathbf{g}} \cdot \mathcal{L}_{\mathbf{v}_{\boldsymbol{\psi}}} \dot{\mathcal{F}})(\mathbf{v}_t) \,.$$

It is expedient to assume that $\mathcal{L}_{\mathbf{v}_{\boldsymbol{\psi}}} \dot{\mathcal{F}} = 0$ *, so that*

$$d_{\dot{\mathbf{v}}_t}(\boldsymbol{\theta}_{\mathbf{g}} \cdot \mathbf{v}_{\boldsymbol{\psi}})(\mathbf{v}_t) = \partial_{\tau=t}\ \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \mathbf{v}_{\boldsymbol{\varphi}_\tau}) \,,$$

$$d_{\mathbf{v}_{\boldsymbol{\psi}}(\mathbf{v}_t)}(\boldsymbol{\theta}_{\mathbf{g}} \cdot \dot{\mathcal{F}})(\mathbf{v}_t) = \partial_{\lambda=0}\ \mathbf{g}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}(\mathbf{v}_t))}(\boldsymbol{\psi}_\lambda(\mathbf{v}_t), \boldsymbol{\psi}_\lambda(\mathbf{v}_t)) = (\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{g})(\mathbf{v}_t, \mathbf{v}_t) \,,$$

and the stationarity condition writes

$$\tfrac{1}{2}\, (\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{g})(\mathbf{v}_t, \mathbf{v}_t) = \partial_{\tau=t}\ \mathbf{g}_{\boldsymbol{\tau}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \mathbf{v}_{\boldsymbol{\varphi}_\tau}) \,.$$

1.9.6 Riemann-Christoffel curvature tensor

In a riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$ endowed with the LEVI-CIVITA connection, the curvature can be represented as a $(0,4)$ tensor field

$$\mathbf{R} : \mathbb{M} \mapsto BL\,(\mathbb{TM}^4\,;\Re)\,,$$

by setting

$$\boxed{\begin{aligned}\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a}) &:= \mathbf{g}\,(\nabla^2_{\mathbf{vu}}\,\mathbf{w}\,,\mathbf{a}) - \mathbf{g}\,(\nabla^2_{\mathbf{uv}}\,\mathbf{w}\,,\mathbf{a})\\ &= \mathbf{g}\,(\nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\,\mathbf{w} - \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\,\mathbf{w} - \nabla_{[\mathbf{v}\,,\mathbf{u}]}\,\mathbf{w}\,,\mathbf{a})\,.\end{aligned}}$$

In a local system of coordinates the components of the curvature tensor field are given by the formula

$$\boxed{\mathbf{R}(\mathbf{e}_i,\mathbf{e}_j,\mathbf{e}_k,\mathbf{e}_D) = \mathbf{g}\,(\nabla_{\mathbf{e}_i}\nabla_{\mathbf{e}_j}\,\mathbf{e}_k - \nabla_{\mathbf{e}_j}\nabla_{\mathbf{e}_i}\,\mathbf{e}_k - \nabla_{[\mathbf{e}_i\,,\mathbf{e}_j]}\,\mathbf{e}_k\,,\mathbf{e}_D)\,.}$$

Hence, recalling that $[\mathbf{e}_i\,,\mathbf{e}_j] = 0$, we get

$$\begin{aligned}\mathbf{R}_{ABCD} &= \mathbf{g}\,(\nabla_{\mathbf{e}_i}(\Gamma^E_{jk}\,\mathbf{e}_E) - \nabla_{\mathbf{e}_j}(\Gamma^E_{ik}\,\mathbf{e}_E),\mathbf{e}_D) =\\ &= \mathbf{g}\,(\Gamma^E_{jC/A}\,\mathbf{e}_E + \Gamma^E_{jk}\,\Gamma^F_{iE}\,\mathbf{e}_F - \Gamma^E_{iC/B}\,\mathbf{e}_E - \Gamma^E_{ik}\,\Gamma^F_{jE}\,\mathbf{e}_F,\mathbf{e}_D) =\\ &= \mathbf{g}\,(\Gamma^E_{jC/A}\,\mathbf{e}_E + \Gamma^F_{jk}\,\Gamma^E_{iF}\,\mathbf{e}_E - \Gamma^E_{iC/B}\,\mathbf{e}_E - \Gamma^F_{ik}\,\Gamma^E_{jF}\,\mathbf{e}_E,\mathbf{e}_D) =\\ &= \mathbf{G}_{ED}\left[\Gamma^E_{jC/A} - \Gamma^E_{iC/B} + \Gamma^F_{jk}\,\Gamma^E_{iF} - \Gamma^F_{ik}\,\Gamma^E_{jF}\right].\end{aligned}$$

Substituting the relations

$$\Gamma^D_{ij}\mathbf{G}_{CD} = \mathbf{G}_{BC/A} + \mathbf{G}_{CA/B} - \mathbf{G}_{AB/C}\,,$$

we obtain the expressions of the components of the curvature tensor in terms of the components of the metric tensor and of its derivatives.

In a riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$ the curvature tensor field $\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a})$ meets the following properties

- $\mathbf{R}$ is *antisymmetric* in the first and in the second pair of arguments

$$\boxed{\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a}) = -\mathbf{R}(\mathbf{u},\mathbf{v},\mathbf{w},\mathbf{a}) = \mathbf{R}(\mathbf{u},\mathbf{v},\mathbf{a},\mathbf{w})\,.}$$

- $\mathbf{R}$ is *symmetric* with respect to an exchange between the first and the second pair of arguments

$$\boxed{\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a}) = \mathbf{R}(\mathbf{w},\mathbf{a},\mathbf{v},\mathbf{u})\,.}$$

- $\mathbf{R}$ fulfills the *first* Bianchi *identity*

$$\boxed{\mathbf{R}(\mathbf{v},\mathbf{u})\,\mathbf{w} + \mathbf{R}(\mathbf{w},\mathbf{v})\,\mathbf{u} + \mathbf{R}(\mathbf{u},\mathbf{w})\,\mathbf{v} = 0\,, \quad \forall\, \mathbf{v},\mathbf{u},\mathbf{w} : \mathbb{M} \mapsto \mathbb{TM}\,.}$$

- $\mathbf{R}$ fulfills the *second* Bianchi *identity*

$$\boxed{(\nabla_{\mathbf{a}}\mathbf{R})(\mathbf{v},\mathbf{u})\,\mathbf{w} + (\nabla_{\mathbf{v}}\mathbf{R})(\mathbf{u},\mathbf{a})\,\mathbf{w} + (\nabla_{\mathbf{u}}\mathbf{R})(\mathbf{a},\mathbf{v})\,\mathbf{w} = 0\,.}$$

For completeness we report the proof of the previous properties following the treatment in [143].

- The antisymmetry of $\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a})$ in the first pair of arguments derives immediately from the definition.

- The antisymmetry in the second pair of arguments can be deduced from the fact that the quadratic form associated with the tensor $\mathbf{R}(\mathbf{v},\mathbf{u})$ does vanish. This follows from the formula

$$\nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\, f - \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\, f = \nabla_{[\mathbf{v}\,,\mathbf{u}]}\, f\,.$$

Indeed the vanishing of the covariant derivative of the metric implies that

$$\begin{aligned}\mathbf{g}\,(\mathbf{R}(\mathbf{v},\mathbf{u})\,\mathbf{w},\mathbf{w}) &= \mathbf{g}\,(\nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\,\mathbf{w},\mathbf{w}) - \mathbf{g}\,(\nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\,\mathbf{w},\mathbf{w}) - \mathbf{g}\,(\nabla_{[\mathbf{v}\,,\mathbf{u}]}\,\mathbf{w},\mathbf{w}) \\ &= \nabla_{\mathbf{v}}(\mathbf{g}\,(\nabla_{\mathbf{u}}\,\mathbf{w},\mathbf{w})) - \mathbf{g}\,(\nabla_{\mathbf{u}}\,\mathbf{w},\nabla_{\mathbf{v}}\,\mathbf{w}) \\ &\quad -\nabla_{\mathbf{u}}(\mathbf{g}\,(\nabla_{\mathbf{v}}\,\mathbf{w},\mathbf{w})) + \mathbf{g}\,(\nabla_{\mathbf{v}}\,\mathbf{w},\nabla_{\mathbf{u}}\,\mathbf{w}) \\ &\quad -\tfrac{1}{2}\,\nabla_{[\mathbf{v}\,,\mathbf{u}]}(\mathbf{g}\,(\mathbf{w},\mathbf{w})) \\ &= \tfrac{1}{2}\,(\nabla_{\mathbf{v}}\,\nabla_{\mathbf{u}} - \nabla_{\mathbf{u}}\,\nabla_{\mathbf{v}} - \nabla_{(\mathbf{v},\mathbf{u})})(\mathbf{g}\,(\mathbf{w},\mathbf{w})) \\ &= \tfrac{1}{2}\,\mathbf{R}(\mathbf{v},\mathbf{u})(\mathbf{g}\,(\mathbf{w},\mathbf{w})) = 0\,.\end{aligned}$$

- The fulfillment of the first BIANCHI identity is proven by observing that by virtue of the tensoriality property, the computations are independent of the extension of the vector arguments to vector fields. By proposition

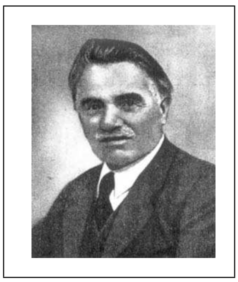

Figure 1.46: Luigi Bianchi (1856 - 1928)

1.3.9 we may assume that the vector fields $\mathbf{v},\mathbf{u},\mathbf{w}$ commute pairwise, that is

$$[\mathbf{v}\,,\mathbf{u}] = [\mathbf{u}\,,\mathbf{w}] = [\mathbf{w}\,,\mathbf{v}] = 0\,.$$

Since the torsion of the connection vanishes, we have that

$$\begin{aligned}\mathbf{R}[\mathbf{v}\,,\mathbf{u}]\,\mathbf{w} + \mathbf{R}[\mathbf{w}\,,\mathbf{v}]\,\mathbf{u} + \mathbf{R}[\mathbf{u}\,,\mathbf{w}]\,\mathbf{v} = \\ &= \nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\,\mathbf{w} - \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\,\mathbf{w} \\ &+ \nabla_{\mathbf{w}}\nabla_{\mathbf{v}}\,\mathbf{u} - \nabla_{\mathbf{v}}\nabla_{\mathbf{w}}\,\mathbf{u} \\ &+ \nabla_{\mathbf{u}}\nabla_{\mathbf{w}}\,\mathbf{v} - \nabla_{\mathbf{w}}\nabla_{\mathbf{u}}\,\mathbf{v} \\ &= \nabla_{\mathbf{v}}\,(\nabla_{\mathbf{u}}\,\mathbf{w} - \nabla_{\mathbf{w}}\,\mathbf{u}) \\ &+ \nabla_{\mathbf{w}}\,(\nabla_{\mathbf{v}}\,\mathbf{u} - \nabla_{\mathbf{u}}\,\mathbf{v}) \\ &+ \nabla_{\mathbf{u}}\,(\nabla_{\mathbf{w}}\,\mathbf{v} - \nabla_{\mathbf{v}}\,\mathbf{w}) \\ &= \nabla_{\mathbf{v}}\,[\mathbf{u}\,,\mathbf{w}] + \nabla_{\mathbf{w}}\,[\mathbf{v}\,,\mathbf{u}] + \nabla_{\mathbf{u}}\,[\mathbf{w}\,,\mathbf{v}] = 0\,.\end{aligned}$$

- The property of symmetry with respect to an exchange between the first and the second pair of arguments is proven by a direct computation. Indeed we have that

$$\begin{aligned}\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a}) &= -\mathbf{R}(\mathbf{u},\mathbf{w},\mathbf{v},\mathbf{a}) - \mathbf{R}(\mathbf{w},\mathbf{v},\mathbf{u},\mathbf{a}) \\ &= \mathbf{R}(\mathbf{u},\mathbf{w},\mathbf{a},\mathbf{v}) + \mathbf{R}(\mathbf{w},\mathbf{v},\mathbf{a},\mathbf{u}) \\ &= -\mathbf{R}(\mathbf{w},\mathbf{a},\mathbf{u},\mathbf{v}) - \mathbf{R}(\mathbf{a},\mathbf{u},\mathbf{w},\mathbf{v}) \\ &\quad -\mathbf{R}(\mathbf{v},\mathbf{a},\mathbf{w},\mathbf{u}) - \mathbf{R}(\mathbf{a},\mathbf{w},\mathbf{v},\mathbf{u}) \\ &= 2\,\mathbf{R}(\mathbf{w},\mathbf{a},\mathbf{v},\mathbf{u}) + \mathbf{R}(\mathbf{a},\mathbf{u},\mathbf{v},\mathbf{w}) + \mathbf{R}(\mathbf{v},\mathbf{a},\mathbf{u},\mathbf{w}) \\ &= 2\,\mathbf{R}(\mathbf{w},\mathbf{a},\mathbf{v},\mathbf{u}) - \mathbf{R}(\mathbf{u},\mathbf{v},\mathbf{a},\mathbf{w}) \\ &= 2\,\mathbf{R}(\mathbf{w},\mathbf{a},\mathbf{v},\mathbf{u}) - \mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a})\,.\end{aligned}$$

Then $2\,\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{w},\mathbf{a}) = 2\,\mathbf{R}(\mathbf{w},\mathbf{a},\mathbf{v},\mathbf{u})$.

- In order to establish the fulfillment of the second BIANCHI identity let us assume again that $[\mathbf{v}\,,\mathbf{u}] = [\mathbf{u}\,,\mathbf{w}] = [\mathbf{w}\,,\mathbf{v}] = 0$.

Then we have that

$$\mathbf{R}(\mathbf{v},\mathbf{u})\,\mathbf{w} = [\nabla_{\mathbf{v}}\,,\nabla_{\mathbf{u}}]\,\mathbf{w} - \nabla_{[\mathbf{v}\,,\mathbf{u}]}\,\mathbf{w} = [\nabla_{\mathbf{v}}\,,\nabla_{\mathbf{u}}]\,\mathbf{w}\,.$$

It must be noted that

$$\begin{aligned}(\nabla_{\mathbf{w}}\mathbf{R})(\mathbf{v},\mathbf{u})\,\mathbf{a} &= \nabla_{\mathbf{w}}(\mathbf{R}(\mathbf{v},\mathbf{u})\,\mathbf{a}) - \mathbf{R}(\nabla_{\mathbf{w}}\mathbf{v},\mathbf{u})\,\mathbf{a}\\ &-\mathbf{R}(\mathbf{v},\nabla_{\mathbf{w}}\mathbf{u})\,\mathbf{a} - \mathbf{R}(\mathbf{v},\mathbf{u})\,\nabla_{\mathbf{w}}\mathbf{a}\\ &= [\nabla_{\mathbf{w}}\,,\mathbf{R}(\mathbf{v},\mathbf{u})]\,\mathbf{a} - \mathbf{R}(\nabla_{\mathbf{w}}\mathbf{v},\mathbf{u})\,\mathbf{a} - \mathbf{R}(\mathbf{v},\nabla_{\mathbf{w}}\mathbf{u})\,\mathbf{a}\,.\end{aligned}$$

Therefore, by JACOBI identity for the commutator, it follows that

$$\begin{aligned}&(\nabla_{\mathbf{w}}\mathbf{R})(\mathbf{v},\mathbf{u})\,\mathbf{a} + (\nabla_{\mathbf{v}}\mathbf{R})(\mathbf{u},\mathbf{w})\,\mathbf{a} + (\nabla_{\mathbf{u}}\mathbf{R})(\mathbf{w},\mathbf{v})\,\mathbf{a}\\ &= [\nabla_{\mathbf{w}}\,,\mathbf{R}(\mathbf{v},\mathbf{u})]\,\mathbf{a} + [\nabla_{\mathbf{v}}\,,\mathbf{R}(\mathbf{u},\mathbf{w})]\,\mathbf{a} + [\nabla_{\mathbf{u}}\,,\mathbf{R}(\mathbf{w},\mathbf{v})]\,\mathbf{a}\\ &-\mathbf{R}(\nabla_{\mathbf{w}}\mathbf{v},\mathbf{u})\,\mathbf{a} - \mathbf{R}(\mathbf{v},\nabla_{\mathbf{w}}\mathbf{u})\,\mathbf{a}\\ &-\mathbf{R}(\nabla_{\mathbf{v}}\mathbf{u},\mathbf{w})\,\mathbf{a} - \mathbf{R}(\mathbf{u},\nabla_{\mathbf{v}}\mathbf{w})\,\mathbf{a}\\ &-\mathbf{R}(\nabla_{\mathbf{u}}\mathbf{w},\mathbf{v})\,\mathbf{a} - \mathbf{R}(\mathbf{w},\nabla_{\mathbf{u}}\mathbf{v})\,\mathbf{a}\\ &= [\nabla_{\mathbf{w}}\,,\mathbf{R}(\mathbf{v},\mathbf{u})]\,\mathbf{a} + [\nabla_{\mathbf{v}}\,,\mathbf{R}(\mathbf{u},\mathbf{w})]\,\mathbf{a} + [\nabla_{\mathbf{u}}\,,\mathbf{R}(\mathbf{w},\mathbf{v})]\,\mathbf{a}\\ &+\mathbf{R}([\mathbf{v}\,,\mathbf{w}],\mathbf{u})\,\mathbf{a} + \mathbf{R}([\mathbf{w}\,,\mathbf{u}],\mathbf{v})\,\mathbf{a} + \mathbf{R}([\mathbf{u}\,,\mathbf{v}],\mathbf{w})\,\mathbf{a}\\ &= \big([\nabla_{\mathbf{w}}\,,[\nabla_{\mathbf{v}}\,,\nabla_{\mathbf{u}}]] + [\nabla_{\mathbf{v}}\,,[\nabla_{\mathbf{u}}\,,\nabla_{\mathbf{w}}]] + [\nabla_{\mathbf{u}}\,,[\nabla_{\mathbf{w}}\,,\nabla_{\mathbf{v}}]]\big)\,\mathbf{a} = 0\,.\end{aligned}$$

1.9.7 Directional and sectional curvature

- The directional curvature operator $\mathbf{R}_{\mathbf{t}} \in BL(\mathbb{TM}\,;\mathbb{TM})$ defined by

$$\mathbf{R}_{\mathbf{t}}(\mathbf{a}) := \mathbf{R}(\mathbf{a},\mathbf{t},\mathbf{t})\,,\quad \forall\,\mathbf{a}\in\mathbb{TM}\,,\quad \mathbf{t}\in \mathrm{C}^2(\mathbb{M}\,;\mathbb{TM})\,,$$

is a symmetric operator since

$$\begin{aligned}\mathbf{g}_{\mathbb{S}}\,(\mathbf{R}_{\mathbf{t}}(\mathbf{a}),\mathbf{h}) &= \mathbf{g}_{\mathbb{S}}\,(\mathbf{R}(\mathbf{a},\mathbf{t},\mathbf{t}),\mathbf{h}) = \mathbf{R}(\mathbf{a},\mathbf{t},\mathbf{t},\mathbf{h})\\ &= \mathbf{R}(\mathbf{t},\mathbf{h},\mathbf{a},\mathbf{t}) = \mathbf{R}(\mathbf{h},\mathbf{t},\mathbf{t},\mathbf{a}) = \mathbf{g}_{\mathbb{S}}\,(\mathbf{R}_{\mathbf{t}}(\mathbf{h}),\mathbf{a})\,.\end{aligned}$$

Further $\mathbf{t}\in\mathbb{TM}$ is in the kernel of $\mathbf{R}_{\mathbf{t}}$.

We further define:

- The *canonical 2-tensor* is the symmetric tensor $\mathbf{R}_2(\mathbf{v},\mathbf{u}) := \mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{u},\mathbf{v})$.

The canonical 2-tensor provides a complete information on the curvature, due to the formula:

$$\begin{aligned}&\partial_{t=0}\,\partial_{s=0}\;\mathbf{R}(\mathbf{v}+t\mathbf{a},\mathbf{u}+s\mathbf{b},\mathbf{v}+t\mathbf{a},\mathbf{u}+s\mathbf{b})\\ -\,&\partial_{t=0}\,\partial_{s=0}\;\mathbf{R}(\mathbf{v}+t\mathbf{b},\mathbf{u}+s\mathbf{a},\mathbf{v}+t\mathbf{b},\mathbf{u}+s\mathbf{a}) = 6\,\mathbf{R}(\mathbf{v},\mathbf{u},\mathbf{a},\mathbf{b})\,.\end{aligned}$$

- The *sectional curvature* operator $sec \in BL(\mathbb{TM}^2\,;\Re)$ is defined by

  $$sec(\mathbf{a},\mathbf{b}) := \frac{\mathbf{g}_{\mathbb{S}}\,(\mathbf{R}_{\mathbf{a}}(\mathbf{b}),\mathbf{b})}{\boldsymbol{\mu}_{\mathbb{M}}(\mathbf{a},\mathbf{b})^2} = \frac{\mathbf{R}(\mathbf{b},\mathbf{a},\mathbf{a},\mathbf{b})}{\boldsymbol{\mu}_{\mathbb{M}}(\mathbf{a},\mathbf{b})^2}\,,$$

  for any pair of linearly independent vectors $\mathbf{a},\mathbf{b}\in\mathbb{TM}$. Here $\mathbf{n}$ is the normal versor to the middle surface $\mathbb{M}$ and $\boldsymbol{\mu}_{\mathbb{M}} := \boldsymbol{\mu}_{\mathbb{S}}\,\mathbf{n}$ is the volume form on $\mathbb{M}$ induced on $\mathbb{M}$ by the volume form $\boldsymbol{\mu}_{\mathbb{S}}$ on $\mathbb{S}$ according to the relation

  $$\boldsymbol{\mu}_{\mathbb{M}}(\mathbf{a},\mathbf{b})\,\mathbf{g}_{\mathbb{S}}\,(\mathbf{n},\mathbf{h}) = (\boldsymbol{\mu}_{\mathbb{S}}\,\mathbf{h})(\mathbf{a},\mathbf{b}) = \boldsymbol{\mu}_{\mathbb{S}}(\mathbf{h},\mathbf{a},\mathbf{b})\,,\quad \forall\,\mathbf{a},\mathbf{b},\mathbf{h}\in\mathbb{TM}\,,$$

  which states that the volume of the oriented parallelepiped with sides $\mathbf{a},\mathbf{b},\mathbf{h}$ in $\mathbb{TM}$ is equal to the product of the area of the oriented base parallelogram with sides $\mathbf{a},\mathbf{b}$ in $\mathbb{TM}$ times the relative height $\mathbf{g}_{\mathbb{S}}\,(\mathbf{n},\mathbf{h})$.

  It is easy to check that $sec(\mathbf{a},\mathbf{b})$ depends only on the plane spanned by the linearly independent vectors $\mathbf{a},\mathbf{b}\in\mathbb{TM}$.

1.9.8 Riemannian isometries

Let $\{\mathbb{M}\,,\mathbf{g}\}$ be a riemannian manifold endowed with the Levi-Civita connection and let $\mathbb{N} = \boldsymbol{\varphi}(\mathbb{M})$ be a differentiable manifold which is diffeomorphic to $\{\mathbb{M}\,,\mathbf{g}\}$ thru the diffeomorphism $\boldsymbol{\varphi}:\mathbb{M}\mapsto\mathbb{N}$. The riemannian manifold $\{\boldsymbol{\varphi}(\mathbb{M})\,,\boldsymbol{\varphi}\!\uparrow\!\mathbf{g}\}$ is said to be *isometric* to the riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$.

Proposition 1.9.8 (Riemannian connections and isometries) *Let us consider two isometric riemannian manifolds* $\{\mathbb{M}\,,\mathbf{g}\}$ *and* $\{\boldsymbol{\varphi}(\mathbb{M})\,,\boldsymbol{\varphi}\!\uparrow\!\mathbf{g}\}$. *The covariant derivatives* ∇ *on* $\{\mathbb{M}\,,\mathbf{g}\}$ *and* $\boldsymbol{\varphi}\!\uparrow\!\nabla$ *on* $\{\boldsymbol{\varphi}(\mathbb{M})\,,\boldsymbol{\varphi}\!\uparrow\!\mathbf{g}\}$, *defined by the corresponding* Levi-Civita *connections, are natural with respect to the push:*

a) $\quad \boldsymbol{\varphi}\!\uparrow\!(\nabla_{\mathbf{u}}\,\mathbf{v}) = (\boldsymbol{\varphi}\!\uparrow\!\nabla)_{(\boldsymbol{\varphi}\uparrow\mathbf{u})}\,(\boldsymbol{\varphi}\!\uparrow\!\mathbf{v})\,,\quad \forall\,\mathbf{v}\in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})\,,\quad \forall\,\mathbf{u}:\mathbb{M}\mapsto\mathbb{TM}\,.$

In other terms, the Levi-Civita *connection on the manifold* $\{\varphi(\mathbb{M}), \varphi\!\uparrow\!\mathbf{g}\}$ *is the connection induced by the diffeomorphism* $\varphi : \mathbb{M} \mapsto \mathbb{N} = \varphi(\mathbb{M})$ *, as defined in section 1.4.17.*

Proof. Formula $a)$ to be proven may alternatively be written as

$$b) \quad (\varphi\!\uparrow\!\mathbf{g})\,((\varphi\!\uparrow\!\nabla)_{(\varphi\uparrow\mathbf{u})}\,(\varphi\!\uparrow\!\mathbf{v}), \varphi\!\uparrow\!\mathbf{w}) = \varphi\!\uparrow\!\big(\mathbf{g}\,(\nabla_{\mathbf{u}}\,\mathbf{v}, \mathbf{w})\big)\,,$$

for all $\mathbf{v}, \mathbf{u}, \mathbf{w} : \mathbb{M} \mapsto \mathbb{TM}$. Indeed from the definition of the push of a tensor we have that

$$c) \quad (\varphi\!\uparrow\!\mathbf{g})\,((\varphi\!\uparrow\!\nabla)_{(\varphi\uparrow\mathbf{u})}\,(\varphi\!\uparrow\!\mathbf{v}), (\varphi\!\uparrow\!\mathbf{w})) = \varphi\!\uparrow\!\big(\mathbf{g}\,(\varphi\!\downarrow\!((\varphi\!\uparrow\!\nabla)_{(\varphi\uparrow\mathbf{u})}\,(\varphi\!\uparrow\!\mathbf{v})), \mathbf{w})\big)\,.$$

By equating the r.h.s. terms of expressions $b)$ and $c)$ we obtain the result $a)$ and viceversa. In order to demonstrate relation $b)$ we recall that the metric tensor of the riemannian manifold $\mathbb{N}$ is $\varphi\!\uparrow\!\mathbf{g}$. Hence, by applying Koszul formula, the l.h.s. term in $b)$ may be rewritten as the sum of the terms

$$d) \quad d_{(\varphi\uparrow\mathbf{u})}\big((\varphi\!\uparrow\!\mathbf{g})\,(\varphi\!\uparrow\!\mathbf{v}, \varphi\!\uparrow\!\mathbf{w})\big)\,, \qquad e) \quad (\varphi\!\uparrow\!\mathbf{g})\,([\varphi\!\uparrow\!\mathbf{u}\,, \varphi\!\uparrow\!\mathbf{v}], \varphi\!\uparrow\!\mathbf{w})\,.$$

From the definition of push we have that

$$\begin{aligned} d) \qquad & d_{(\varphi\uparrow\mathbf{u})}\big((\varphi\!\uparrow\!\mathbf{g})\,(\varphi\!\uparrow\!\mathbf{v}, \varphi\!\uparrow\!\mathbf{w})\big) = d_{(\varphi\uparrow\mathbf{u})}\big(\varphi\!\uparrow\!\big(\mathbf{g}\,(\mathbf{v}, \mathbf{w})\big)\big)\,, \\ e) \qquad & (\varphi\!\uparrow\!\mathbf{g})\,([\varphi\!\uparrow\!\mathbf{u}\,, \varphi\!\uparrow\!\mathbf{v}], \varphi\!\uparrow\!\mathbf{w}) = \varphi\!\uparrow\!\big(\mathbf{g}\,(\varphi\!\downarrow\![\varphi\!\uparrow\!\mathbf{u}\,, \varphi\!\uparrow\!\mathbf{v}], \mathbf{w})\big)\,. \end{aligned}$$

Propositions 1.2.4 and 1.3.4 ensure that the directional derivative and the Lie bracket are natural with respect to the push. The following equalities then hold

$$\begin{aligned} d) \qquad & d_{(\varphi\uparrow\mathbf{u})}\big(\varphi\!\uparrow\!\big(\mathbf{g}\,(\mathbf{v}, \mathbf{w})\big)\big) = \varphi\!\uparrow\!\big(d_{\mathbf{u}}\big(\mathbf{g}\,(\mathbf{v}, \mathbf{w})\big)\big)\,, \\ e) \qquad & \varphi\!\uparrow\!\big(\mathbf{g}\,(\varphi\!\downarrow\![\varphi\!\uparrow\!\mathbf{u}\,, \varphi\!\uparrow\!\mathbf{v}], \mathbf{w})\big) = \varphi\!\uparrow\!\big(\mathbf{g}\,([\mathbf{u}\,, \mathbf{v}], \mathbf{w})\big)\,. \end{aligned}$$

By applying again Koszul formula we obtain the equality in $b)$. ■

Proposition 1.9.8 tells us that the Levi-Civita covariant derivative is natural with respect to riemannian isometries.

Let now $\{\mathbb{M}, \mathbf{g}\}$ and $\{\mathbb{N}, \varphi\!\uparrow\!\mathbf{g}\}$ be two isometric riemannian manifolds related by the diffeomorphism $\varphi : \mathbb{M} \mapsto \mathbb{N}$. In force of Proposition 1.9.8 the curvature tensor fields $\mathbf{R}_{\mathbb{M}}$ and $\mathbf{R}_{\mathbb{N}}$ are related by the formula

$$\mathbf{R}_{\mathbb{N}} = \varphi\!\uparrow\!\mathbf{R}_{\mathbb{M}}\,.$$

Indeed it is sufficient to observe that

$$\varphi\!\uparrow\!(\mathbf{R}_{\mathbb{M}}(\mathbf{v},\mathbf{u})\,\mathbf{w}) = \varphi\!\uparrow\!(\nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\,\mathbf{w}) - \varphi\!\uparrow\!(\nabla_{\mathbf{u}}\nabla_{\mathbf{v}}\,\mathbf{w}) - \varphi\!\uparrow\!(\nabla_{(\mathbf{v},\mathbf{u})}\,\mathbf{w})\,,$$

and that

$$\varphi\!\uparrow\!(\nabla_{\mathbf{v}}\nabla_{\mathbf{u}}\,\mathbf{w}) = \nabla_{\varphi\uparrow\mathbf{v}}\varphi\!\uparrow\!(\nabla_{\mathbf{u}}\,\mathbf{w}) = \nabla_{\varphi\uparrow\mathbf{v}}\,\nabla_{\varphi\uparrow\mathbf{u}}\,(\varphi\!\uparrow\!\mathbf{w})\,,$$

$$\varphi\!\uparrow\!(\nabla_{[\mathbf{v}\,,\mathbf{u}]}\,\mathbf{w}) = \nabla_{\varphi\uparrow[\mathbf{v}\,,\mathbf{u}]}(\varphi\!\uparrow\!\mathbf{w}) = \nabla_{[\varphi\uparrow\mathbf{v}\,,\varphi\uparrow\mathbf{u}]}(\varphi\!\uparrow\!\mathbf{w})\,.$$

Hence

$$\varphi\!\uparrow\!(\mathbf{R}_{\mathbb{M}}\,[\mathbf{v}\,,\mathbf{u}]\,\mathbf{w}) = \mathbf{R}_{\mathbb{N}}\,[\varphi\!\uparrow\!\mathbf{v}\,,\varphi\!\uparrow\!\mathbf{u}]\,(\varphi\!\uparrow\!\mathbf{w})\,.$$

Since

$$\varphi\!\uparrow\!(\mathbf{R}_{\mathbb{M}}\,[\mathbf{v}\,,\mathbf{u}]\,\mathbf{w}) = (\varphi\!\uparrow\!\mathbf{R}_{\mathbb{M}})\,[\varphi\!\uparrow\!\mathbf{v}\,,\varphi\!\uparrow\!\mathbf{u}]\,(\varphi\!\uparrow\!\mathbf{w})\,,$$

we deduce that $\mathbf{R}_{\mathbb{N}} = \varphi\!\uparrow\!\mathbf{R}_{\mathbb{M}}$.

A riemannian manifold with an identically vanishing curvature tensor field is called a *flat manifold*. From the previous formula it follows that a riemannian manifold which is isometric to a flat riemannian manifold is flat too.

1.9.9 Euclidean spaces

In a euclidean space the translation defines a distant parallel transport and the related standard connection. It follows that the curvature tensor field vanishes identically. Moreover the torsion also vanishes since, by tensoriality, we may extend the vector fields by translation, so that

$$\textsc{tors}(\mathbf{v},\mathbf{u}) := \nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v} - [\mathbf{v}\,,\mathbf{u}] = -[\mathbf{v}\,,\mathbf{u}] = 0\,,$$

since flows with constant velocities in a euclidean space pairwise commute. Then:

- The euclidean space $\{\mathbb{S}\,,\textsc{can}\}$, endowed with the canonical metric tensor, is flat and torsion-free.

Let $\varphi \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ be a diffeomorphism between the $3D$-submanifolds $\mathbb{M}$ and $\mathbb{N}$ of the euclidean space $\mathbb{S}$. If the manifold $\mathbb{N}$ is equipped with the euclidean canonical metric, then the riemannian manifold $\{\mathbb{M}\,,\varphi\!\downarrow\!\textsc{can}\}$ is flat since it is isometric to $\{\mathbb{N}\,,\textsc{can}\}$ which is flat. The Green metric tensor field $\varphi\!\downarrow\!\textsc{can}$ is defined by

$$(\varphi\!\downarrow\!\textsc{can})(\mathbf{v},\mathbf{u}) = \varphi\!\downarrow\!(\textsc{can}\,(\varphi\!\uparrow\!\mathbf{v},\varphi\!\uparrow\!\mathbf{u}))\,,$$

for any pair of vector fields $\mathbf{v},\mathbf{u} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$.

The Green metric tensor field is associated in $\{\mathbb{S}\,,\text{CAN}\}$ with the field of linear operators $\boldsymbol{\varphi}\!\uparrow^T \circ \boldsymbol{\varphi}\!\uparrow \in \mathrm{C}^1(\mathbb{M}\,;BL\,(\mathbb{TM}\,;\mathbb{TM}))$. Indeed we have that

$$\boldsymbol{\varphi}\!\uparrow((\boldsymbol{\varphi}\!\downarrow\text{CAN})(\mathbf{v},\mathbf{u})) = \text{CAN}\,(\boldsymbol{\varphi}\!\uparrow \mathbf{v}, \boldsymbol{\varphi}\!\uparrow \mathbf{u}) = \text{CAN}\,(\boldsymbol{\varphi}\!\uparrow^T \boldsymbol{\varphi}\!\uparrow \mathbf{v}, \mathbf{u})\,, \quad \forall\, \mathbf{v},\mathbf{u} \in \mathbb{T}_{\mathbb{M}}\,.$$

We may now discuss an integrability condition which plays an important role in continuum mechanics. If the support of a riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$ is an open connected subset of an euclidean space $\{\mathbb{S}\,,\text{CAN}\}$, it is physically important to ask whether the metric tensor field $\mathbf{g}$ could be obtained as the Green metric tensor field $\boldsymbol{\varphi}\!\downarrow\text{CAN}$ associated with a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$ between $\mathbb{M}$ and a submanifold $\{\mathbb{N}\,,\text{CAN}\}$ of the euclidean space $\{\mathbb{S}\,,\text{CAN}\}$. The necessary and sufficient condition is that the riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$ be flat:

Proposition 1.9.9 (Isometries in an euclidean space) *A nD* Riemann *manifold* $\{\mathbb{M}\,,\mathbf{g}\}$ *with a flat* Levi-Civita *connection is locally isometrically diffeomorphic to a submanifold* $\{\mathbb{N}\,,\text{CAN}\}$ *of the nD euclidean space.*

Proof. Flatness means that the horizontal subbundle is locally integrable. Each vector $\mathbf{e}_i$ of a frame at a point $\mathbf{x}_0 \in \mathbb{M}$ can then be extended in an open connected neighborhood $U_{\mathbb{M}}(\mathbf{x}_0) \subset \mathbb{M}$ to a vector field $\mathbf{v}_i \in \mathrm{C}^1(U_{\mathbb{M}}(\mathbf{x}_0)\,;\mathbb{TM})$ tangent to the leaf of the horizontal foliation passing through $\mathbf{v}_i(\mathbf{x}_0) = \mathbf{e}_i$. The values taken in the horizontal leaf by the vector field $\mathbf{v}_i \in \mathrm{C}^1(U_{\mathbb{M}}(\mathbf{x}_0)\,;\mathbb{TM})$ along a curve $\mathbf{c} = \mathbf{Fl}^{\mathbf{v}} \in \mathrm{C}^1(I\,;\mathbb{M})$ in the base manifold, with $\mathbf{c}(0) = \mathbf{x}_0$, are generated by parallel transport along that curve:

$$(\mathbf{v}_i \circ \mathbf{Fl}^{\mathbf{v}}_\lambda)(\mathbf{x}_0) = \mathbf{Fl}^{\mathbf{H}_{\mathbf{v}}}_\lambda(\mathbf{e}_i)\,, \quad \forall\, \lambda \in I\,.$$

By definition, the covariant derivative $\nabla_{\mathbf{v}}\mathbf{v}_i$ vanishes identically. By the vanishing of the torsion, the Lie derivative of any pair of these vector fields vanishes too:

$$\text{TORS}(\mathbf{v}_i, \mathbf{v}_j) := \nabla_{\mathbf{v}_i}\mathbf{v}_j - \nabla_{\mathbf{v}_j}\mathbf{v}_i - [\mathbf{v}_i\,,\mathbf{v}_j] = -[\mathbf{v}_i\,,\mathbf{v}_j] = 0\,.$$

The local frame is then a local coordinate system defined by a chart $\xi \in \mathrm{C}^1(U_{\mathbb{M}}(\mathbf{x}_0)\,;\Re^m)$ with $\xi\!\uparrow\!\mathbf{v}_i(\mathbf{x}) = \sigma_i$, the standard basis vector in $\Re^m$. We denote by STD the standard metric in $\Re^m$. Since the covariant derivative is metric preserving and the vector fields $\mathbf{v}_i \in \mathrm{C}^1(U_{\mathbb{M}}(\mathbf{x}_0)\,;\mathbb{TM})$ are parallel transported along any curve, from the formula:

$$\nabla_{\mathbf{v}}[\mathbf{g}(\mathbf{v}_i, \mathbf{v}_j)] = (\nabla_{\mathbf{v}}\mathbf{g})(\mathbf{v}_i, \mathbf{v}_j) + \mathbf{g}(\nabla_{\mathbf{v}}\mathbf{v}_i, \mathbf{v}_j) + \mathbf{g}(\mathbf{v}_i, \nabla_{\mathbf{v}}\mathbf{v}_j) = 0\,,$$

we infer that the value of the metric evaluated on each pair of vector fields of the local frame is constant along any curve and hence, by connectedness:

$$\mathbf{g}_{\mathbf{x}}(\mathbf{v}_i(\mathbf{x}), \mathbf{v}_j(\mathbf{x})) = \mathbf{g}_{\mathbf{x}_0}(\mathbf{e}_i, \mathbf{e}_j)\,, \quad \forall\, \mathbf{x} \in U(\mathbf{x}_0)\,.$$

Then

$$(\xi{\uparrow}\mathbf{g})_{\xi(\mathbf{x})}(\sigma_i, \sigma_j) = \mathbf{g}_{\mathbf{x}}(\xi{\downarrow}\sigma_i, \xi{\downarrow}\sigma_j) = \mathbf{g}_{\mathbf{x}}(\mathbf{v}_i(\mathbf{x}), \mathbf{v}_j(\mathbf{x})) = \mathbf{g}_{\mathbf{x}_0}(\mathbf{e}_i, \mathbf{e}_j)\,.$$

If the basis $\{\mathbf{e}_i \mid i = 1, \dots, n\}$ is $\mathbf{g}$-orthonormal, we get: $\xi{\uparrow}\mathbf{g} = \textsc{std}$. Since the euclidean nD space is flat and torsion-free when endowed woth the Levi-Civita connnection associated with the canonical metric $\textsc{can}$, we may also consider a submanifold $\{\mathbb{N}\,, \textsc{can}\}$ locally mapped to $\xi(U_{\mathbb{M}}(\mathbf{x}_0) \subset \Re^n)$ by a chart $\zeta \in \mathrm{C}^1(U_{\mathbb{N}}(\mathbf{z}_0)\,; \xi(U_{\mathbb{M}}(\mathbf{x}_0)) \subset \Re^n)$ with $\zeta(\mathbf{z}_0) = \xi(\mathbf{x}_0)$ and

$$(\zeta{\uparrow}\textsc{can})_{\zeta(\mathbf{z})}(\sigma_i, \sigma_j) = \textsc{can}_{\mathbf{z}}(\xi{\downarrow}\sigma_i, \xi{\downarrow}\sigma_j) = \textsc{can}_{\mathbf{z}}(\mathbf{w}_i(\mathbf{z}), \mathbf{w}_j(\mathbf{z})) = \textsc{can}_{\mathbf{z}_0}(\mathbf{h}_i, \mathbf{h}_j)\,,$$

with the basis $\{\mathbf{h}_i \mid i = 1, \dots, n\}$ $\textsc{can}$-orthonormal so that $\zeta{\uparrow}\textsc{can} = \textsc{std}$. The diffeomorphism $\boldsymbol{\varphi} = \zeta^{-1} \circ \xi \in \mathrm{C}^1(U_{\mathbb{M}}(\mathbf{x}_0)\,; U_{\mathbb{N}}(\mathbf{z}_0))$ is then such that $\boldsymbol{\varphi}{\downarrow}\textsc{can} = \mathbf{g}$. ■

In continuum mechanics the stretching of a body is defined as the change in lenght of any curve drawn in it. A suitable measure of the stretching is provided by a field $\mathbf{g}$ of metric tensors on the initial placement $\mathbb{M}$ of the body, which evaluates the lenght of a curve $\mathbf{c} \in \mathrm{C}^1(I\,; \mathbb{M})$ by means of the formula

$$\int_I \mathbf{g}\,(\partial_{\tau=t}\,\mathbf{c}(\tau), \partial_{\tau=t}\,\mathbf{c}(\tau))^{\frac{1}{2}}\; d\tau\,,$$

whose value is independent of the chosen parametrization.

A field of metric tensors which is not equal to the canonical metric field $\textsc{can}$ provides a pointwise measure of the stretching and the difference

$$\tfrac{1}{2}(\mathbf{g} - \textsc{can})$$

is called the Green's strain field. The scalar factor $\frac{1}{2}$ is inserted for convenience in order to get for the stretching rate the expression $\frac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g} = \mathrm{sym}\,\nabla\mathbf{v}$ for a Levi-Civita connection, see section 1.9.4. This choice eventually leads to define by duality a stress field whose flux is a force per unit surface area. The strain field is said to be kinematically compatible if $\boldsymbol{\varphi}{\downarrow}\textsc{can} = \mathbf{g}$ with $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{N})$ is a diffeomorphism describing a change of placement of the body in the 3D euclidean space $\{\mathbb{S}\,, \textsc{can}\}$.

1.10 Hypersurfaces

Let us consider a $(n-1)$-dimensional submanifold $\mathbb{M}$ of a n-dimensional riemannian manifold $\{\mathbb{S}, \mathbf{g}_{\mathbb{S}}\}$ where $\mathbf{g}_{\mathbb{S}} \in BL(\mathbb{TS}, \mathbb{TS}; \Re)$ is the metric tensor field on $\mathbb{S}$ and $\mathbb{TS}$ is the tangent bundle to $\mathbb{S}$.

Let us denote by $\mathbb{TS}, \mathbb{TM}$ the tangent bundles to the manifold $\mathbb{S}$ and to the submanifold $\mathbb{M}$ respectively, and by $\mathbb{TS}(\mathbb{M})$ the restriction of the tangent bundle $\mathbb{TS}$ to the submanifold $\mathbb{M}$. The elements of the linear space $\mathbb{TS}(\mathbb{M})$ are the applied vectors $\{\mathbf{x}, \mathbf{v}\}$ with $\mathbf{v} \in \mathbb{TS}$ and base point $\mathbf{x} \in \mathbb{M}$.

1.10.1 Distance function and shape operator

The tangent bundle $\mathbb{TS}(\mathbb{M})$ is n-dimensional and the bundle $\mathbb{TM}$ is $(n-1)$-dimensional. A *distance function* from a $(n-1)$-dimensional submanifold $\mathbb{M} \subset \mathbb{S}$ is a scalar valued map $f \in C^2(\mathcal{O}; \Re)$ which is twice continuously differentiable in an open neighborhood $\mathcal{O} \subset \mathbb{S}$ and such that its gradient is a vector field with unitary norm. A distance function is then a solution the non-linear *eikonal equation* or Hamilton-Jacobi equation:

$$\|\nabla f(\mathbf{x})\| = 1, \quad \forall \mathbf{x} \in \mathcal{O},$$

where ∇ is the gradient operator on $\{\mathbb{S}, \mathbf{g}_{\mathbb{S}}\}$ according to the Levi-Civita connection.

To provide a constructive example of a distance function we may consider an open strip $U_{\mathbb{M}} \subset \mathbb{S}$ including $\mathbb{M}$ whose thickness is suitably small so that every point $\mathbf{x} \in U_{\mathbb{M}}$ can be orthogonally projected in an unique fashion onto a point $P_{\mathrm{V}}(\mathbf{x}) \in \mathbb{M}$ according to the metric of the riemannian manifold $\{\mathbb{S}, \mathbf{g}_{\mathbb{S}}\}$.

For any fixed $r \in \Re$ we denote by $\mathbb{M}^r$ the r-level set of $f \in C^2(\mathcal{O}; \Re)$ which is an hypersurface parallel to $\mathbb{M}$. The hypersurface $\mathbb{M}^r$ is the r-level folium and the family $U_{\mathbb{M}}$ of all admissible folii is the foliation of $\mathbb{M}$.

We denote by $\mathbb{TS}(U_{\mathbb{M}})$ the family of all vectors $\{\mathbf{x}, \mathbf{h}\}$ of $\mathbb{TS}$ with base point $\mathbf{x} \in U_{\mathbb{M}}$ and by $\mathbb{TM}(U_{\mathbb{M}}) \subset \mathbb{TS}(U_{\mathbb{M}})$ those which are tangent to the folium passing through $\mathbf{x} \in U_{\mathbb{M}}$. Let us then consider the nonlinear projector $P_{\mathrm{V}} \in C^1(U_{\mathbb{M}}; \mathbb{M})$ which maps any point $\mathbf{x} \in \mathbb{M}^r \subset U_{\mathbb{M}}$ of the r-level folium to the unique point $P_{\mathrm{V}}(\mathbf{x}) \in \mathbb{M}$ which minimizes the distance between $\mathbf{x} \in U_{\mathbb{M}}$ and $\mathbb{M}$ according to the metric of $\{\mathbb{S}, \mathbf{g}_{\mathbb{S}}\}$.

We denote by $\mathbf{n}(\mathbf{x})$ the normal versor to $\mathbb{M}^r$ at $\mathbf{x} \in \mathbb{M}^r$ and by $\mathbf{n}(P_{\mathrm{V}}(\mathbf{x}))$ the normal versor to $\mathbb{M}$ at $P_{\mathrm{V}}(\mathbf{x}) \in \mathbb{M}$. Observing that $\mathbf{n}(\mathbf{x}) = \mathbf{n}(P_{\mathrm{V}}(\mathbf{x}))$ we

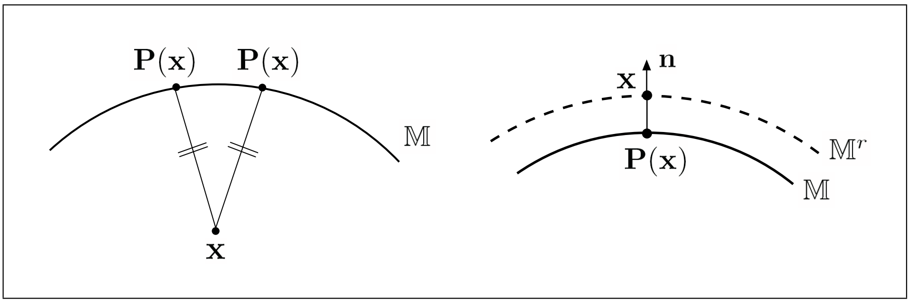

Figure 1.47: Distance function

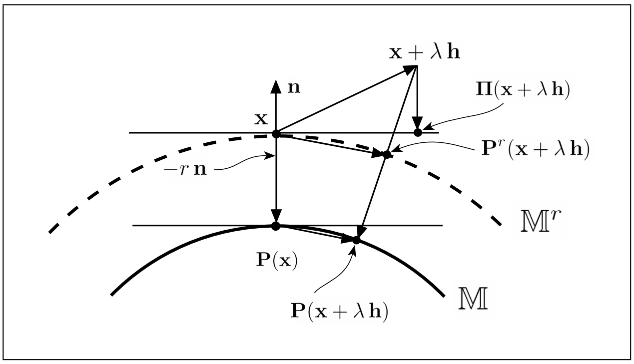

Figure 1.48: Projectors

may write

$$\mathbf{x} = P_{\mathrm{V}}(\mathbf{x}) + r(\mathbf{x})\,\mathbf{n}(\mathbf{x}) = P_{\mathrm{V}}(\mathbf{x}) + r(\mathbf{x})\,\mathbf{n}(P_{\mathrm{V}}(\mathbf{x}))\,,$$

and then define the signed-distance function

$$f(\mathbf{x}) = r(\mathbf{x})\,, \quad \forall\, \mathbf{x} \in \mathcal{O} = U_{\mathbb{M}}\,,$$

which is differentiable in the open set $U_{\mathbb{M}}$, and also the distance function

$$f(\mathbf{x}) = |r(\mathbf{x})|\,, \quad \forall\, \mathbf{x} \in \mathcal{O} = U_{\mathbb{M}} \backslash \mathbb{M}\,.$$

By construction we have that

$$\mathbf{n}(\mathbf{x}) = \nabla f(\mathbf{x})\,, \quad \forall\, \mathbf{x} \in \mathcal{O}\,.$$

The hessian of the distance function is the tensor field of type $(1,1)$:

$$\nabla^2 f(\mathbf{x}) := \nabla(\nabla f)(\mathbf{x}) \in BL\,(\mathbb{TM}(\mathbf{x})\,; \mathbb{TM}(\mathbf{x}))\,, \quad \forall\, \mathbf{x} \in \mathcal{O}\,,$$

It provides a description of the variation of the normal to the manifold $\mathbb{M}$ at each point, since

$$\nabla \mathbf{n}(\mathbf{x}) = \nabla(\nabla f)(\mathbf{x})\,, \quad \forall\, \mathbf{x} \in \mathcal{O}\,,$$

and is therefore called the *shape operator* of $\mathbb{M}$.

Two basic properties are proven in the next Lemmas.

Lemma 1.10.1 *The shape operator* $\mathbf{S}(\mathbf{x}) := \nabla \mathbf{n}(\mathbf{x}) \in BL\,(\mathbb{TS}(\mathbf{x})\,; \mathbb{TS}(\mathbf{x}))$ *is symmetric.*

Proof. The symmetry of $\mathbf{S}(\mathbf{x}) = \nabla \mathbf{n}(\mathbf{x}) = \nabla^2 f(\mathbf{x})$ is a direct consequence of the symmetry of the riemannian connection of $\{\mathbb{S}\,, \mathbf{g}_{\mathbb{S}}\}$.

Lemma 1.10.2 *At any point* $\mathbf{x} \in \mathbb{M}^r$ *the normal versor* $\mathbf{n} \in \mathbb{TS}(\mathbf{x})$ *belongs to the kernel of the shape operator* $\mathbf{S}(\mathbf{x}) \in BL\,(\mathbb{TS}(\mathbf{x})\,; \mathbb{TS}(\mathbf{x}))$*:*

$$\mathbf{S}(\mathbf{x})\,\mathbf{n}(\mathbf{x}) = \nabla^2 f(\mathbf{x})\,\mathbf{n}(\mathbf{x}) = 0\,,$$

hence $Im\,\mathbf{S}(\mathbf{x}) = (Ker\,\mathbf{S}(\mathbf{x}))^{\perp} \subseteq \mathbb{T}_{\mathbb{M}^r}(\mathbf{x})$.

Proof. By the symmetry of $\mathbf{S}(\mathbf{x}) \in BL\,(\mathbb{TS}(\mathbf{x})\,; \mathbb{TS}(\mathbf{x}))$ we have that

$$\mathbf{g}_{\mathbb{S}}\,(\mathbf{Sn}, \mathbf{h}) = \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sh}, \mathbf{n}) = \mathbf{g}_{\mathbb{S}}\,(d_{\mathbf{h}}\mathbf{n}, \mathbf{n}) = \frac{1}{2}\, d_{\mathbf{h}}\, \mathbf{g}_{\mathbb{S}}\,(\mathbf{n}, \mathbf{n}) = 0\,,$$

for any $\mathbf{h} \in \mathbb{TS}(\mathbf{x})$. The last statement follows from the well-known property that the kernel and the image of a symmetric operator are mutual orthogonal complements. ■

In the sequel we will be interested in the case where the n-dimensional riemannian manifold $\{\mathbb{S}\,, \mathbf{g}_{\mathbb{S}}\}$ is an n-dimensional euclidean space endowed with the canonical connection corresponding to the parallel transport defined by the translations in the euclidean space. The covariant derivative in $\{\mathbb{S}\,, \mathbf{g}_{\mathbb{S}}\}$ is then the usual directional derivative in $\{\mathbb{S}\,, \mathbf{g}_{\mathbb{S}}\}$ and will be denoted by d.

1.10.2 Nonlinear projector

Let us derive here for subsequent use a noteworthy formula concerning the derivative of the nonlinear projector $P_{\mathrm{V}} \in \mathrm{C}^1(U_{\mathbb{M}}\,;\mathbb{M})$ on an oriented $(n-1)$-dimensional submanifold $\mathbb{M}$ of the n-dimensional euclidean space $\{\mathbb{S}\,,\mathbf{g}_{\mathbb{S}}\}$.

Lemma 1.10.3 *Let $P_{\mathrm{V}} \in \mathrm{C}^1(U_{\mathbb{M}}\,;\mathbb{M})$ be the nonlinear projector of the points of the foliation $U_{\mathbb{M}} \subset \mathbb{S}$ on the hypersurface $\mathbb{M}$. Its derivative at $\mathbf{x} \in \mathbb{M}^r \subset U_{\mathbb{M}}$ is a linear operator $dP_{\mathrm{V}}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(P_{\mathrm{V}}(\mathbf{x})))$ which is related to the linear projector $\boldsymbol{\pi}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(\mathbf{x}))$ of the vectors $\mathbf{h} \in \mathbb{TS}(\mathbf{x})$ on the tangent plane at $\mathbf{x} \in \mathbb{M}^r$ to the r-level folium $\mathbb{M}^r$, by the formulas*

$$\boldsymbol{\pi}(\mathbf{x}) = dP_{\mathrm{V}}(\mathbf{x}) + r(\mathbf{x})\,\mathbf{S}(\mathbf{x}) = (\mathbf{I} + r(\mathbf{x})\,\mathbf{S}(P_{\mathrm{V}}(\mathbf{x})))\,dP_{\mathrm{V}}(\mathbf{x}) = \boldsymbol{\pi}(P_{\mathrm{V}}(\mathbf{x}))\,,$$

where $\mathbf{S}(P_{\mathrm{V}}(\mathbf{x}))$ is the shape operator of $\mathbb{M}$ at the point $P_{\mathrm{V}}(\mathbf{x}) \in \mathbb{M}$ and $\mathbf{S}(\mathbf{x})$ is the shape operator of $\mathbb{M}^r$ at the point $\mathbf{x} \in \mathbb{M}^r$.

Proof. Taking the directional derivative along any $\mathbf{a} \in \mathbb{T}_{\mathbb{M}^r}(\mathbf{x})$ in the formula $\mathbf{x} = P_{\mathrm{V}}(\mathbf{x}) + r(\mathbf{x})\,\mathbf{n}(\mathbf{x}) = P_{\mathrm{V}}(\mathbf{x}) + r(\mathbf{x})\,\mathbf{n}(P_{\mathrm{V}}(\mathbf{x}))$ and observing that

$$\begin{aligned} d\mathbf{n}(\mathbf{x}) &= \mathbf{S}(\mathbf{x})\,, \\ d(\mathbf{n} \circ P_{\mathrm{V}})(\mathbf{x}) &= d\mathbf{n}(P_{\mathrm{V}}(\mathbf{x}))\,dP_{\mathrm{V}}(\mathbf{x}) = \mathbf{S}(P_{\mathrm{V}}(\mathbf{x}))\,dP_{\mathrm{V}}(\mathbf{x})\,, \end{aligned}$$

we get the relation

$$\mathbf{a} = dP_{\mathrm{V}}(\mathbf{x})\mathbf{a} + r(\mathbf{x})\,\mathbf{S}(\mathbf{x})\mathbf{a}\,, \quad \forall\, \mathbf{a} \in \mathbb{T}_{\mathbb{M}^r}(\mathbf{x})\,.$$

Since $dP_{\mathrm{V}}(\mathbf{x})\mathbf{n} = 0$ and $\mathbf{S}(\mathbf{x})\mathbf{n} = 0$ also

$$(dP_{\mathrm{V}}(\mathbf{x}) + r(\mathbf{x})\,\mathbf{S}(\mathbf{x}))\mathbf{n} = 0\,.$$

Hence the operator $dP_{\mathrm{V}}(\mathbf{x}) + r(\mathbf{x})\,\mathbf{S}(\mathbf{x})$ maps any vector $\mathbf{h} \in \mathbb{TS}(\mathbf{x})$ into its projection onto $\mathbb{T}_{\mathbb{M}^r}(\mathbf{x})$ and the formula is proved. ■

Remark 1.10.1 *The ranges of the linear operators $\boldsymbol{\pi}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(\mathbf{x}))$ and of $dP_{\mathrm{V}}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(P_{\mathrm{V}}(\mathbf{x})))$ at $\mathbf{x} \in \mathbb{M}^r$ are $\mathbb{T}_{\mathbb{M}^r}(\mathbf{x})$ and $\mathbb{TM}(P_{\mathrm{V}}(\mathbf{x}))$ respectively. In an euclidean space the linear subspaces $\mathbb{TS}(\mathbf{x})$ and $\mathbb{TS}(P_{\mathrm{V}}(\mathbf{x}))$ are identified by means of the parallel translation defined by the translation operation. Accordingly also the subspaces $\mathbb{T}_{\mathbb{M}^r}(\mathbf{x})$ and $\mathbb{TM}(P_{\mathrm{V}}(\mathbf{x}))$ will be identified and considered as subspaces of the linear space $\mathbb{TS}(\mathbf{x})$.*

Lemma 1.10.4 *For any* $\mathbf{x} \in U_{\mathbb{M}}$ *the operator* $dP_{\mathrm{V}}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(\mathbf{x}))$ *is symmetric and is related to the linear projector* $\boldsymbol{\pi}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(\mathbf{x}))$ *by the formulas*

$$dP_{\mathrm{V}}(\mathbf{x}) = \boldsymbol{\pi}(\mathbf{x}) - r(\mathbf{x})\,\mathbf{S}(\mathbf{x}) = (\mathbf{I} + r(\mathbf{x})\,\mathbf{S}(P_{\mathrm{V}}(\mathbf{x})))^{-1}\,\boldsymbol{\pi}(\mathbf{x})\,.$$

Moreover $Ker\, dP_{\mathrm{V}}(\mathbf{x}) = Ker\, \boldsymbol{\pi}(\mathbf{x}) = Span(\mathbf{n}(\mathbf{x}))$.

Proof. The formulas for $dP_{\mathrm{V}}(\mathbf{x})$ both follow directly from Lemma 1.10.3. The symmetry of $dP_{\mathrm{V}}(\mathbf{x})$ is apparent from the first formula since both terms on the r.h.s are symmetric. Indeed the shape operator is symmetric by Lemma 1.10.1 and the linear operator $\boldsymbol{\pi}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(\mathbf{x}))$ is symmetric being a linear orthogonal projector since (see e.g. [192]):

$$\mathbf{g}_{\mathbb{S}}(\boldsymbol{\pi}\mathbf{a}, \mathbf{b}) = \mathbf{g}_{\mathbb{S}}(\boldsymbol{\pi}\mathbf{a}, \boldsymbol{\pi}\mathbf{b})\,, \qquad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{TS}(\mathbf{x})\,.$$

To establish the second formula we preliminarily observe that the linear operator $\mathbf{I} + r(\mathbf{x})\,\mathbf{S}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(\mathbf{x}))$ is symmetric and positive definite, and hence invertible, due to the suitably small value of the thickness of the shell choosen to ensure that the nonlinear projector $P_{\mathrm{V}} \in \mathrm{C}^1(U_{\mathbb{M}}\,;\mathbb{M})$ be well-defined. ■

The symmetry of $dP_{\mathrm{V}}(\mathbf{x}) \in BL(\mathbb{TS}(\mathbf{x})\,;\mathbb{TS}(\mathbf{x}))$ may also be inferred from the second formula since the symmetric operators $\boldsymbol{\pi}(\mathbf{x})$ and $\mathbf{I} + r(\mathbf{x})\,\mathbf{S}(\mathbf{x})$ commute. Indeed at $\mathbf{x} \in \mathbb{M}^r$ the eigenspaces of the former, which are the tangent spaces $\mathbb{T}_{\mathbb{M}^r}(\mathbf{x})$ to $\mathbb{M}^r$ and the linear span of $\mathbf{n}(P_{\mathrm{V}}(\mathbf{x}))$, are invariant subspaces for the latter [69]. The same is true if the latter operator is replaced by its inverse.

1.10.3 First and second fundamental forms

- The first fundamental form on $\mathbb{M}$ is the twice covariant tensor field

  $$\mathbf{g}_{\mathbb{M}} \in BL(\mathbb{TM}^2\,;\Re)\,,$$

  defined on $\mathbb{M}$ as the restriction of the metric $\mathbf{g}_{\mathbb{S}} \in BL(\mathbb{TS}^2\,;\Re)$ of $\{\mathbb{S}\,, \mathbf{g}_{\mathbb{S}}\}$ to the vectors $\{\mathbf{x}\,, \mathbf{h}\} \in \mathbb{TM}$.

An analogous definition may be given for $\mathbf{g}_{\mathbb{M}^r} \in BL(\mathbb{T}^2_{\mathbb{M}^r}\,;\Re)$ on each folium $\mathbb{M}^r$. These fundamental forms $\mathbf{g}_{\mathbb{M}}$ induce, on each folium $\mathbb{M}^r$ of the foliation $U_{\mathbb{M}}$, a riemannian metric. The resulting riemannian manifold is $\{\mathbb{M}^r\,, \mathbf{g}_{\mathbb{M}^r}\}$.

- The second fundamental form on $\mathbb{M}$ is the twice covariant tensor field $\mathbf{s}_{\mathbb{M}} \in BL(\mathbb{TM}^2 ; \Re)$ defined by

$$\mathbf{s}_{\mathbb{M}}(\mathbf{a}, \mathbf{b}) = \mathbf{g}_{\mathbb{S}}(\mathbf{S}\mathbf{a}, \mathbf{b}) = \mathbf{g}_{\mathbb{M}}(\mathbf{S}_{\mathbb{M}}\mathbf{a}, \mathbf{b}), \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{TM}.$$

The mixed tensor $\mathbf{S}_{\mathbb{M}} \in BL(\mathbb{TM} ; \mathbb{TM})$, which picks up the two-dimensional essential part of the shape operator, is dubbed the WEINGARTEN operator.

Lemma 1.10.5 *The* WEINGARTEN *operator* $\mathbf{S}_{\mathbb{M}} \in BL(\mathbb{TM} ; \mathbb{TM})$ *meets the identity*

$$\mathbf{g}_{\mathbb{M}}(\mathbf{S}_{\mathbb{M}}\mathbf{a}, \mathbf{b}) = -\mathbf{g}_{\mathbb{S}}(\mathbf{n}, d_{\mathbf{a}}\mathbf{b}), \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{TM}.$$

Proof. Since for all $\mathbf{b} \in \mathbb{TM}$ the inner product $\mathbf{g}_{\mathbb{S}}(\mathbf{n}, \mathbf{b})$ vanishes identically on $\mathbb{M}$, we have that

$$0 = d_{\mathbf{a}}\, \mathbf{g}_{\mathbb{S}}(\mathbf{n}, \mathbf{b}) = \mathbf{g}_{\mathbb{S}}(\mathbf{S}\mathbf{a}, \mathbf{b}) + \mathbf{g}_{\mathbb{S}}(\mathbf{n}, d_{\mathbf{a}}\mathbf{b}), \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{TM},$$

which is the result. ■

Remark 1.10.2 *The identity in Lemma 1.10.5 is usually taken as the definition of the* WEINGARTEN *operator. It is important to highlight the surprising tensoriality property of the bilinear form* $\mathbf{g}_{\mathbb{S}}(\mathbf{n}, d_{\mathbf{a}}\mathbf{b})$ *in spite of the apparent dependence of the derivative* $d_{\mathbf{a}}\mathbf{b}$ *from the local behavior of the field* $\mathbf{b} \in \mathrm{C}^1(\mathbb{M} ; \mathbb{TM})$.

By applying the tensoriality criterion of Lemma 1.2.1:

$$\mathbf{g}_{\mathbb{S}}(\mathbf{n}, d_{\mathbf{a}}(f\,\mathbf{b})) = (d_{\mathbf{a}}f)\, \mathbf{g}_{\mathbb{S}}(\mathbf{n}, \mathbf{b})) + f\, \mathbf{g}_{\mathbb{S}}(\mathbf{n}, d_{\mathbf{a}}\mathbf{b}) = f\, \mathbf{g}_{\mathbb{S}}(\mathbf{n}, d_{\mathbf{a}}\mathbf{b}),$$

we realize that the tensoriality property is due to the orthogonality between the vector $\mathbf{n}$ *and the vectors* $\mathbf{b} \in \mathbb{TM}(\mathbf{x})$ *at any point* $\mathbf{x} \in \mathbb{M}$.

1.10.4 Gauss and Mainardi-Codazzi formulas

Let $\{\mathbb{S}, \mathbf{g}_{\mathbb{S}}\}$ be a riemannian manifold and $\{\mathbb{M}^r, \mathbf{g}_{\mathbb{M}^r}\}$ be the r-level set of a foliation $U_{\mathbb{M}}$ of $\{\mathbb{S}, \mathbf{g}_{\mathbb{S}}\}$. Let us denote by ∇^r the riemannian connection induced on $\mathbb{M}^r$ by the riemannian connection ∇ on $\mathbb{S}$. The associated covariant derivative is defined by the orthogonal projection formula:

$$\nabla^r_{\mathbf{a}}\mathbf{b} := \boldsymbol{\pi}\, \nabla_{\mathbf{a}}\mathbf{b}, \quad \forall\, \mathbf{a} \in \mathbb{T}_{\mathbb{M}^r}, \quad \mathbf{b} \in \mathrm{C}^1(\mathbb{M}^r ; \mathbb{T}_{\mathbb{M}^r}).$$

Hence we have that

$$\mathbf{g}_{\mathbb{S}}(\nabla_{\mathbf{a}}\mathbf{t}, \mathbf{b}) = \mathbf{g}_{\mathbb{S}}(\nabla^r_{\mathbf{a}}\mathbf{t}, \mathbf{b}), \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{T}_{\mathbb{M}^r}, \quad \mathbf{t} \in \mathrm{C}^1(\mathbb{M}^r ; \mathbb{T}_{\mathbb{M}^r}).$$

We recall (see section 1.4.8) that the second covariant derivative of a tangent vector field is defined by:

$$\nabla^2_{\mathbf{ab}}\mathbf{t} := \nabla_{\mathbf{a}}\nabla_{\mathbf{b}}\mathbf{t} - \nabla_{\nabla_{\mathbf{a}}\mathbf{b}}\mathbf{t} \quad \forall\, \mathbf{a} \in \mathbb{TS}\,, \quad \mathbf{b} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})\,, \quad \mathbf{t} \in \mathrm{C}^2(\mathbb{S}\,;\mathbb{TS})\,,$$

and that the RIEMANN-CHRISTOFFEL curvature of a riemannian manifold $\{\mathbb{S}\,, \mathbf{g}_{\mathbb{S}}\}$ is the fourth order tensor field $\mathbf{R} \in BL\,(\mathbb{TS}^3\,;\mathbb{TS})$ which provides the skew part of the second covariant derivative of a tangent vector field:

$$\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) := \nabla_{\mathbf{a}}\nabla_{\mathbf{b}}\mathbf{t} - \nabla_{\mathbf{b}}\nabla_{\mathbf{a}}\mathbf{t} - \nabla_{[\mathbf{a}\,,\mathbf{b}]}\mathbf{t} = \nabla^2_{\mathbf{ab}}\mathbf{t} - \nabla^2_{\mathbf{ba}}\mathbf{t}\,.$$

for all $\mathbf{a},\mathbf{b} \in \mathbb{TS}$ and $\mathbf{t} \in \mathrm{C}^2(\mathbb{S}\,;\mathbb{TS})$.

Defining the four times covariant curvature tensor $\mathbf{R} \in BL\,(\mathbb{TS}^3\,;\Re)$:

$$\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t},\mathbf{h}) := \mathbf{g}_{\mathbb{S}}\,(\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}),\mathbf{h})\,, \quad \forall\, \mathbf{a},\mathbf{b},\mathbf{h} \in \mathbb{TS}\,, \quad \mathbf{t} \in \mathrm{C}^2(\mathbb{S}\,;\mathbb{TS})\,.$$

we recall also the symmetry property

$$\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t},\mathbf{h}) = \mathbf{R}(\mathbf{t},\mathbf{h},\mathbf{a},\mathbf{b})\,,$$

and the skew-symmetry properties

$$\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t},\mathbf{h}) = -\mathbf{R}(\mathbf{b},\mathbf{a},\mathbf{t},\mathbf{h}) = \mathbf{R}(\mathbf{b},\mathbf{a},\mathbf{h},\mathbf{t})\,.$$

It is natural to look for the relation between the RIEMANN-CHRISTOFFEL curvature of the ambient manifold $\{\mathbb{S}\,, \mathbf{g}_{\mathbb{S}}\}$ and the one of the embedded manifold $\{\mathbb{M}^r\,, \mathbf{g}_{\mathbb{M}^r}\}$.

The answer is provided by a direct computation (see e.g. [143]) which leads to the following formulas for the tangent and the normal components of the curvature vector $\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t})$, respectively named after GAUSS and MAINARDI-CODAZZI:

$$\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) = \mathit{tan}\,\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) + \mathit{nor}\,\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t})\,,$$

for all $\mathbf{a},\mathbf{b} \in \mathbb{T}_{\mathbb{M}^r}\,, \mathbf{t} \in \mathrm{C}^2(\mathbb{M}^r\,;\mathbb{T}_{\mathbb{M}^r})$, where

$$\mathit{tan}\,\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) = \mathbf{R}^r(\mathbf{a},\mathbf{b},\mathbf{t}) + \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sa},\mathbf{t})\,\mathbf{Sb} - \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sb},\mathbf{t})\,\mathbf{Sa}\,,$$

$$\mathit{nor}\,\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) = \mathbf{g}_{\mathbb{S}}\,((\nabla_{\mathbf{b}}\mathbf{S})\mathbf{a},\mathbf{t})\,\mathbf{n} - \mathbf{g}_{\mathbb{S}}\,((\nabla_{\mathbf{a}}\mathbf{S})\mathbf{b},\mathbf{t})\,\mathbf{n}\,.$$

Remark 1.10.3 *We may give an alternative form to the* MAINARDI-CODAZZI *formula by observing that the covariant derivative of the shape operator is still a symmetric operator. Indeed we have that*

$$\begin{aligned}\nabla_{\mathbf{t}}(\mathbf{g}_{\mathbb{S}}\,(\mathbf{Sa},\mathbf{b})) &= \mathbf{g}_{\mathbb{S}}\,(\nabla_{\mathbf{t}}(\mathbf{Sa}),\mathbf{b}) + \mathbf{g}_{\mathbb{S}}\,((\mathbf{Sa}),\nabla_{\mathbf{t}}\mathbf{b}) = \\ &= \mathbf{g}_{\mathbb{S}}\,((\nabla_{\mathbf{t}}\mathbf{S})\mathbf{a},\mathbf{b}) + \mathbf{g}_{\mathbb{S}}\,(\mathbf{S}\nabla_{\mathbf{t}}\mathbf{a},\mathbf{b}) + \mathbf{g}_{\mathbb{S}}\,(\mathbf{S}\nabla_{\mathbf{t}}\mathbf{b},\mathbf{a})\,.\end{aligned}$$

The symmetry of the first term on the r.h.s. follows from the symmetry of the term on l.h.s and the symmetry of the sum of the last two terms on the r.h.s. Hence the MAINARDI-CODAZZI *formula may be written as*

$$nor\,\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) = \mathbf{g}_{\mathbb{S}}\,((\nabla_{\mathbf{b}}\mathbf{S})\mathbf{t},\mathbf{a})\,\mathbf{n} - \mathbf{g}_{\mathbb{S}}\,((\nabla_{\mathbf{a}}\mathbf{S})\mathbf{t},\mathbf{b})\,\mathbf{n}\,.$$

The RIEMANN-CHRISTOFFEL curvature $\mathbf{R} \in BL\,(\mathbb{TS}^4\,;\Re)$ vanishes on the manifold $\mathbb{S}$ if and only if the manifold is euclidean.

In this case the GAUSS and MAINARDI-CODAZZI formulas yield two integrability conditions:

$$tan\,\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) = \mathbf{R}^r(\mathbf{a},\mathbf{b},\mathbf{t}) + \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sa},\mathbf{t})\,\mathbf{Sb} - \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sb},\mathbf{t})\,\mathbf{Sa} = 0\,,$$

$$nor\,\mathbf{R}(\mathbf{a},\mathbf{b},\mathbf{t}) = \mathbf{g}_{\mathbb{S}}\,((\nabla_{\mathbf{b}}\mathbf{S})\mathbf{a},\mathbf{t})\,\mathbf{n} - \mathbf{g}_{\mathbb{S}}\,((\nabla_{\mathbf{a}}\mathbf{S})\mathbf{b},\mathbf{t})\,\mathbf{n} = 0\,,$$

for all $\mathbf{a},\mathbf{b} \in \mathbb{T}_{\mathbb{M}^r}\,,\mathbf{t} \in \mathrm{C}^2(\mathbb{M}\,;\mathbb{T}_{\mathbb{M}^r})$.

In the euclidean space the curvature of an embedded hypersurface $\{\mathbb{M}^r\,,\mathbf{g}_{\mathbb{M}^r}\}$ vanishes if the shape operator vanishes.

Let us recall that:

- The *sectional curvature* $sec \in BL\,(\mathbb{TM}^2\,;\Re)$ is defined by

  $$sec(\mathbf{a},\mathbf{b}) := \frac{\mathbf{g}_{\mathbb{S}}\,(\mathbf{R}_{\mathbf{a}}(\mathbf{b}),\mathbf{b})}{\boldsymbol{\mu}_{\mathbb{M}}(\mathbf{a},\mathbf{b})^2} = \frac{\mathbf{R}(\mathbf{b},\mathbf{a},\mathbf{a},\mathbf{b})}{\boldsymbol{\mu}_{\mathbb{M}}(\mathbf{a},\mathbf{b})^2}\,,$$

  for any pair of linearly independent vectors $\mathbf{a},\mathbf{b} \in \mathbb{TM}$, where $\mathbf{n}$ is the normal versor to the middle surface $\mathbb{M}$ and $\boldsymbol{\mu}_{\mathbb{M}} := \boldsymbol{\mu}_{\mathbb{S}}\,\mathbf{n}$ is the volume form on $\mathbb{M}$ induced on $\mathbb{M}$ by the volume form $\boldsymbol{\mu}_{\mathbb{S}}$ on $\mathbb{S}$.

Lemma 1.10.6 (Theorema egregium) *Let* $\mathbb{S}$ *be the* 3*-dimensional euclidean space and* $\mathbb{M}$ *a regular surface in* $\mathbb{S}$*. Then we have that*

$$sec(\mathbf{a},\mathbf{b}) = \det\mathbf{S}\,,$$

Proof. By GAUSS formula for the tangential curvature we have that

$$\mathbf{R}^r(\mathbf{a},\mathbf{b},\mathbf{t}) = \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sb},\mathbf{t})\,\mathbf{Sa} - \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sa},\mathbf{t})\,\mathbf{Sb}\,,$$

and hence

$$\mathbf{R}^r(\mathbf{a},\mathbf{b},\mathbf{b},\mathbf{a}) = \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sa},\mathbf{a})\,\mathbf{g}_{\mathbb{S}}\,(\mathbf{Sb},\mathbf{b}) - \mathbf{g}_{\mathbb{S}}\,(\mathbf{Sa},\mathbf{b})^2\,.$$

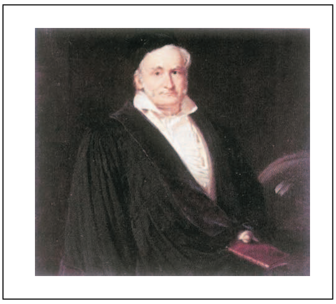

Figure 1.49: Karl Friederich Gauss (1777 - 1855)

If $\{\mathbf{a}, \mathbf{b}\}$ is an orthonormal basis of $\mathbb{TM}$ we get the relation

$$\mathbf{R}^r(\mathbf{a}, \mathbf{b}, \mathbf{b}, \mathbf{a}) = sec(\mathbf{a}, \mathbf{b}) = \det \mathbf{S}\,,$$

which is a formal expression of the celebrated *theorema egregium* of GAUSS: although $\mathbf{S}$ is an extrinsic quantity, its determinant $\det \mathbf{S}$ is an intrinsic quantity, i.e. it depends only on the metric tensor of the surface. ■

The eigenvalues of $\mathbf{S}$ are called the principal curvatures while the determinant of $\mathbf{S}$ (i.e. the product of the the principal curvatures) is called the gaussian curvature of the surface.

If a sheet of paper is bent without any stretching, the principal curvatures do change while the gaussian curvature remains invariant.

A direct computation provides a third fundamental equation which relates the directional curvature operator to the shape operator:

$$\nabla_{\mathbf{n}}\mathbf{S} + \mathbf{S}^2 = -\mathbf{R}_{\mathbf{n}}\,,$$

and is referred to as the *radial curvature equation* (see e.g. [143]).

Indeed, by LEIBNIZ rule:

$$(\nabla_{\mathbf{n}}\mathbf{S})\mathbf{a} = \nabla_{\mathbf{n}}(\mathbf{S}\mathbf{a}) - \mathbf{S}(\nabla_{\mathbf{n}}\mathbf{a}) = \nabla_{\mathbf{n}}(\nabla_{\mathbf{a}}\mathbf{n}) - \nabla_{\nabla_{\mathbf{n}}\mathbf{a}}\mathbf{n}$$
$$\mathbf{S}^2\mathbf{a} = \mathbf{S}\,\mathbf{S}\mathbf{a} = \mathbf{S}(\nabla_{\mathbf{a}}\mathbf{n}) = \nabla_{\nabla_{\mathbf{a}}\mathbf{n}}\mathbf{n}\,.$$

Then, being $\nabla_{\mathbf{n}}\mathbf{n} = 0$ we get

$$\begin{aligned}(\nabla_{\mathbf{n}}\mathbf{S})\mathbf{a} + \mathbf{S}^2\mathbf{a} &= \nabla_{\mathbf{n}}(\mathbf{S}\mathbf{a}) - \mathbf{S}(\nabla_{\mathbf{n}}\mathbf{a}) = \nabla_{\mathbf{n}}(\nabla_{\mathbf{a}}\mathbf{n}) - \nabla_{\nabla_{\mathbf{n}}\mathbf{a}}\mathbf{n} + \nabla_{\nabla_{\mathbf{a}}\mathbf{n}}\mathbf{n} \\ &= \nabla_{\mathbf{n}}(\nabla_{\mathbf{a}}\mathbf{n}) - \nabla_{\mathbf{a}}(\nabla_{\mathbf{n}}\mathbf{n}) - (\nabla_{\nabla_{\mathbf{n}}\mathbf{a}} - \nabla_{\nabla_{\mathbf{a}}\mathbf{n}})\mathbf{n} = \\ &= \nabla_{\mathbf{n}}(\nabla_{\mathbf{a}}\mathbf{n}) - \nabla_{\mathbf{a}}(\nabla_{\mathbf{n}}\mathbf{n}) - \nabla_{[\mathbf{n}\,,\mathbf{a}]}\,\mathbf{n} = \mathbf{R}(\mathbf{n}, \mathbf{a}, \mathbf{n}) = -\mathbf{R}_{\mathbf{n}}\mathbf{a}\,.\end{aligned}$$

1.10.5 Flowing hypersurfaces

Let us begin by stating the abstract context we are dealing with.

We consider in a nD riemannian manifold $\{\mathbb{M}\,,\mathbf{g}\}$ with standard volume-form $\boldsymbol{\mu}$ and riemannian connection ∇:

- a $(n-1)$D submanifold Σ (an hypersurface) with boundary $\partial\Sigma$,
- a flow $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ with velocity field $\mathbf{v}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ and the nD flow-tube $J_{\mathbf{v}}(\Sigma)$ traced by Σ.

Denoting by $\mathbf{n}_\Sigma(\mathbf{x}) \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ the unit normal to Σ at $\mathbf{x} \in \Sigma$, the flow generates on the tube $J_{\mathbf{v}}(\Sigma)$ a vector field $\mathbf{n}_\Sigma$ of unit normals to the dragged hypersurfaces $\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)$ by setting:

$$(\mathbf{n}_\Sigma \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t})(\mathbf{x}) := \mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}(\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\mathbf{x}))\,, \quad \mathbf{x} \in \Sigma\,.$$

Accordingly a $(n-1)$-form-valued field $\boldsymbol{\mu}_\Sigma$ is generated on $J_{\mathbf{v}}(\Sigma)$ by the contraction:

$$(\boldsymbol{\mu}_\Sigma \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t})(\mathbf{x}) := \boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}(\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\mathbf{x})) = \boldsymbol{\mu}\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}(\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\mathbf{x}))\,, \quad \mathbf{x} \in \Sigma\,,$$

whose restriction to the tangent bundle of $\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)$ defines a field of $(n-1)$-dimensional volume-forms on the dragged hypersurfaces which we shall call the *area-form* of the hypersurfaces.

1.10.6 Transport theorem for a flowing hypersurface

If we consider a $(n-1)$D submanifold Σ flowing in a nD ambient riemannian manifold $\mathbb{M}$, the transport formula may be given a peculiar form.

The following two preliminary results are interesting *di per se* and will be referred to in the proof of Proposition 1.10.1.

Lemma 1.10.7 *In a riemannian manifold* $\{\mathbb{M}\,,\mathbf{g}\}$ *the* LIE*-derivative of the hypersurface area-form along the flow generated by the field of normals* $\mathbf{n}_\Sigma$ *is equal to the surface mean-curvature times the area-form:*

$$\mathcal{L}_{\mathbf{n}_\Sigma}\,\boldsymbol{\mu}_\Sigma = (\mathrm{tr}\mathbf{S}_\Sigma)\,\boldsymbol{\mu}_\Sigma\,.$$

Proof. We proceed as in Lemma 1.9.3, noting in addition that, by Lemma 1.10.2, we have that $\nabla_{\mathbf{n}_\Sigma}\mathbf{n}_\Sigma = 0$ and $\mathbf{g}(\nabla_{\mathbf{a}}\mathbf{n}_\Sigma, \mathbf{n}_\Sigma) = 0$ so that $\nabla_{\mathbf{a}}\mathbf{n}_\Sigma \in \mathbb{T}\Sigma$ for all $\mathbf{a} \in \mathbb{T}\Sigma$. Then, being

$$\mathcal{L}_{\mathbf{a}}\mathbf{b} = \nabla_{\mathbf{a}}\mathbf{b} - \nabla_{\mathbf{b}}\mathbf{a}\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathrm{C}^1(\Sigma\,; \mathbb{T}\Sigma)\,,$$

and $\nabla\,\boldsymbol{\mu} = 0$ we have that:

$$\begin{aligned}
&(\mathcal{L}_{\mathbf{n}_\Sigma}\,\boldsymbol{\mu}_\Sigma)(\mathbf{a}, \mathbf{b}) = (\mathcal{L}_{\mathbf{n}_\Sigma}\,(\boldsymbol{\mu}\mathbf{n}_\Sigma))(\mathbf{a}, \mathbf{b}) = \\
&\mathcal{L}_{\mathbf{n}_\Sigma}\,(\boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b})) - \boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathcal{L}_{\mathbf{n}_\Sigma}\mathbf{a}, \mathbf{b})) - \boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathbf{a}, \mathcal{L}_{\mathbf{n}_\Sigma}\mathbf{b})) = \\
&\nabla_{\mathbf{n}_\Sigma}\,(\boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b})) - \boldsymbol{\mu}(\nabla_{\mathbf{n}_\Sigma}\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b}) - \boldsymbol{\mu}(\mathbf{n}_\Sigma, \nabla_{\mathbf{n}_\Sigma}\mathbf{a}, \mathbf{b}) - \boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathbf{a}, \nabla_{\mathbf{n}_\Sigma}\mathbf{b}) \\
&\qquad + \boldsymbol{\mu}(\nabla_{\mathbf{n}_\Sigma}\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b}) + \boldsymbol{\mu}(\mathbf{n}_\Sigma, \nabla_{\mathbf{a}}\mathbf{n}_\Sigma, \mathbf{b}) + \boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathbf{a}, \nabla_{\mathbf{b}}\mathbf{n}_\Sigma) = \\
&(\nabla_{\mathbf{n}_\Sigma}\,\boldsymbol{\mu})(\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b}) + \boldsymbol{\mu}(\nabla_{\mathbf{n}_\Sigma}\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b}) + \boldsymbol{\mu}(\mathbf{n}_\Sigma, \nabla_{\mathbf{a}}\mathbf{n}_\Sigma, \mathbf{b}) + \boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathbf{a}, \nabla_{\mathbf{b}}\mathbf{n}_\Sigma) = \\
&\qquad = \boldsymbol{\mu}_\Sigma(\mathbf{S}_\Sigma\mathbf{a}, \mathbf{b}) + \boldsymbol{\mu}_\Sigma(\mathbf{a}, \mathbf{S}_\Sigma\mathbf{b}) = \mathrm{tr}(\mathbf{S}_\Sigma)\,\boldsymbol{\mu}_\Sigma(\mathbf{a}, \mathbf{b})\,,
\end{aligned}$$

and the assertion is proved. ■

Lemma 1.10.8 *The* Lie*-derivative of the hypersurfaces area-forms along the flow generated by the field of normal velocities* $v_{\mathbf{n}}\mathbf{n}_\Sigma$ *is equal to the* Lie*-derivative of volume-form along the flow generated by the field of normals* $\mathbf{n}_\Sigma$ *times the normal component* $v_{\mathbf{n}}$ *of the velocity :*

$$\mathcal{L}_{(v_{\mathbf{n}}\mathbf{n}_\Sigma)}\,\boldsymbol{\mu}_\Sigma = v_{\mathbf{n}}\,\mathcal{L}_{\mathbf{n}_\Sigma}\,\boldsymbol{\mu}_\Sigma\,.$$

Proof. By applying twice the homotopy formula, we have that:

$$\begin{aligned}
\mathcal{L}_{(v_{\mathbf{n}}\mathbf{n}_\Sigma)}\,\boldsymbol{\mu}_\Sigma &= d(v_{\mathbf{n}}\boldsymbol{\mu}_\Sigma\mathbf{n}_\Sigma) + v_{\mathbf{n}}(d\boldsymbol{\mu}_\Sigma)\mathbf{n}_\Sigma \\
&= d(v_{\mathbf{n}}\boldsymbol{\mu}_\Sigma\mathbf{n}_\Sigma) + v_{\mathbf{n}}\,(\mathcal{L}_{\mathbf{n}_\Sigma}\,\boldsymbol{\mu}_\Sigma) - d(\boldsymbol{\mu}_\Sigma\mathbf{n}_\Sigma)\,,
\end{aligned}$$

and hence, being $\boldsymbol{\mu}_\Sigma\mathbf{n}_\Sigma = \boldsymbol{\mu}\mathbf{n}_\Sigma\mathbf{n}_\Sigma = 0$, we get the result. ■

Proposition 1.10.1 (Flowing hypersurface) *Let* $\{\mathbb{M}\,, \mathbf{g}\}$ *be a* nD *riemannian manifold with standard volume-form* $\boldsymbol{\mu}$ *and connection* ∇ *and let* Σ *be a* $(n-1)D$ *submanifold with boundary* $\partial\Sigma$. *Given a flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{M})$

with velocity field $\mathbf{v}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *and a time-dependent* $(n-1)$*-form* $\boldsymbol{\omega}_t^{n-1}$ *on the* n*D flow tube* $J_{\mathbf{v}}(\Sigma)$, *the transport formula*

$$\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\omega}_\tau^{n-1} = \int_\Sigma \mathcal{L}_{t,\mathbf{v}}\, \boldsymbol{\omega}_t^{n-1}\,,$$

takes the expression

$$\begin{aligned}\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} f_\tau\, \boldsymbol{\mu}_\Sigma &= \int_\Sigma \mathcal{L}_{t,\mathbf{v}}\,(f_t\, \boldsymbol{\mu}_\Sigma) = \int_\Sigma (\mathcal{L}_{t,\mathbf{v}}\, f_t)\, \boldsymbol{\mu}_\Sigma + f_t\,(\mathcal{L}_{\mathbf{v}}\, \boldsymbol{\mu}_\Sigma) \\ &= \int_\Sigma (\mathcal{L}_{t,\mathbf{v}}\, f_t + f_t\,(v_{\mathbf{n}}\,\mathrm{tr}\mathbf{S}_\Sigma + \mathrm{div}_\Sigma \mathbf{v}^{||}))\, \boldsymbol{\mu}_\Sigma \\ &= \int_\Sigma (\partial_{\tau=t}\ f_\tau + v_{\mathbf{n}}\, \mathcal{L}_{\mathbf{n}_\Sigma} f_t + f_t\, v_{\mathbf{n}}\,\mathrm{tr}\mathbf{S}_\Sigma + \mathrm{div}_\Sigma(f_t\, \mathbf{v}^{||}))\, \boldsymbol{\mu}_\Sigma \\ &= \int_\Sigma (\partial_{\tau=t}\ f_\tau + v_{\mathbf{n}}\, \mathcal{L}_{\mathbf{n}_\Sigma} f_t + f_t\, v_{\mathbf{n}}\,\mathrm{tr}\mathbf{S}_\Sigma)\, \boldsymbol{\mu}_\Sigma + \int_{\partial\Sigma} f_t\, v_{\partial\Sigma}\, \boldsymbol{\mu}_{\partial\Sigma}\,,\end{aligned}$$

where

- $\mathbf{n}_{\partial\Sigma} \in \mathbb{T}\Sigma$ *unit normal field to* $\partial\Sigma$
- $\boldsymbol{\mu}_{\partial\Sigma} = \boldsymbol{\mu}_\Sigma\, \mathbf{n}_{\partial\Sigma}$ *induced boundary-volume-form on* $\partial\Sigma$
- $\mathbf{S}_\Sigma = \nabla \mathbf{n}_\Sigma$ *shape operator of* Σ
- $\mathbf{v} = v_{\mathbf{n}}\, \mathbf{n}_\Sigma + \mathbf{v}^{||}\,, \quad \mathbf{v}^{||} \in \mathbb{T}\Sigma$
- $v_{\partial\Sigma} = \mathbf{g}(\mathbf{v}^{||}, \mathbf{n}_{\partial\Sigma})$ *normal speed of* $\partial\Sigma$
- $\mathcal{L}_{\mathbf{v}}$ LIE *derivative along the flow*
- $\mathcal{L}_{t,\mathbf{v}}$ *convective time-derivative along the flow.*

Proof. Let us write $\boldsymbol{\alpha}\Sigma$ for the integral of a area-form $\boldsymbol{\alpha}$ on Σ. Since area-forms are proportional one-another we may set $\boldsymbol{\omega}_t^{n-1} = f_t\, \boldsymbol{\mu}_\Sigma$ on the $(n-1)$D submanifold Σ, so that

$$\mathcal{L}_{t,\mathbf{v}}\, \boldsymbol{\omega}_t^{n-1} = \mathcal{L}_{t,\mathbf{v}}\,(f_t\, \boldsymbol{\mu}_\Sigma) = (\mathcal{L}_{t,\mathbf{v}}\, f_t)\, \boldsymbol{\mu}_\Sigma + f_t\, \mathcal{L}_{\mathbf{v}}\, \boldsymbol{\mu}_\Sigma\,.$$

To explicit the dependence on the shape of the hypersurface, the velocity is decomposed into its normal and tangential components to Σ: $\mathbf{v} = v_{\mathbf{n}}\,\mathbf{n}_\Sigma + \mathbf{v}^{||}$. Substituting, and recalling the formulas in Lemmas 1.10.7 and 1.10.8, we get

$$\begin{aligned}\mathcal{L}_{\mathbf{v}}\,\boldsymbol{\mu}_\Sigma &= \mathcal{L}_{(v_{\mathbf{n}}\,\mathbf{n}_\Sigma)}\,\boldsymbol{\mu}_\Sigma + \mathcal{L}_{\mathbf{v}^{||}}\,\boldsymbol{\mu}_\Sigma \\ &= v_{\mathbf{n}}\,\mathcal{L}_{\mathbf{n}_\Sigma}\,\boldsymbol{\mu}_\Sigma + \mathcal{L}_{\mathbf{v}^{||}}\,\boldsymbol{\mu}_\Sigma \\ &= v_{\mathbf{n}}\,(\mathrm{tr}\mathbf{S}_\Sigma)\,\boldsymbol{\mu}_\Sigma + (\mathrm{div}_\Sigma\mathbf{v}^{||})\,\boldsymbol{\mu}_\Sigma\,.\end{aligned}$$

The alternative expression of the transport formula may be obtained by setting

$$\mathcal{L}_{t,\mathbf{v}}\,\boldsymbol{\omega}_t^{n-1} = \mathcal{L}_{t,\mathbf{v}}\,(f_t\,\boldsymbol{\mu}_\Sigma) = \partial_{\tau=t}\;f_\tau\,\boldsymbol{\mu}_\Sigma + \mathcal{L}_{t,\mathbf{v}}\,(f_t\,\boldsymbol{\mu}_\Sigma)\,,$$

and noting that

$$\begin{aligned}\mathcal{L}_{\mathbf{v}}\,(f_t\,\boldsymbol{\mu}_\Sigma) &= \mathcal{L}_{\mathbf{v}^{||}}(f_t\,\boldsymbol{\mu}_\Sigma) + \mathcal{L}_{v_{\mathbf{n}}\mathbf{n}_\Sigma}(f_t\,\boldsymbol{\mu}_\Sigma) \\ &= \mathcal{L}_{(f_t\mathbf{v}^{||})}\boldsymbol{\mu}_\Sigma + (\mathcal{L}_{v_{\mathbf{n}}\mathbf{n}_\Sigma}f_t)\boldsymbol{\mu}_\Sigma + f_t\,(\mathcal{L}_{(v_{\mathbf{n}}\mathbf{n}_\Sigma)}\boldsymbol{\mu}_\Sigma) \\ &= (\mathrm{div}_\Sigma(f_t\mathbf{v}^{||}) + v_{\mathbf{n}}\,(\mathcal{L}_{\mathbf{n}_\Sigma}f_t) + f_t\,v_{\mathbf{n}}\,(\mathrm{tr}\mathbf{S}_\Sigma))\,\boldsymbol{\mu}_\Sigma\,,\end{aligned}$$

and then the result is proved. ■

In particular, form the transport theorem of Proposition 1.10.1, we get the following formula for the rate of change of the total area of the flowing hypersurface

$$\begin{aligned}\partial_{\tau=t}\int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}\boldsymbol{\mu}_\Sigma &= \int_\Sigma (v_{\mathbf{n}}\,\mathrm{tr}\mathbf{S}_\Sigma + \mathrm{div}_\Sigma\mathbf{v}^{||})\,\boldsymbol{\mu}_\Sigma \\ &= \int_\Sigma v_{\mathbf{n}}\,\mathrm{tr}\mathbf{S}_\Sigma\,\boldsymbol{\mu}_\Sigma + \int_{\partial\Sigma} v_{\partial\Sigma}\,\boldsymbol{\mu}_{\partial\Sigma}\,,\end{aligned}$$

which tells us that:

- The specific rate of change of the area of the flowing hypersurface is the sum between the normal velocity of the flow times the mean curvature of the hypersurface and the divergence of the parallel velocity on the surface.
- Alternatively the latter contribution may be globally interpreted as the flux of the parallel velocity thru the boundary of the surface. It vanishes if the hypersurface is closed (no-boundary).

1.10.7 Piola's transform and Nanson's formula

PIOLA's transform $P_{\boldsymbol{\varphi}} \in BL(\mathbb{TM}\,;\mathbb{TM})$ answers to the following question: which is the vector whose flux is equal to the pull back of the flux of a given vector? The flux of a vector $\mathbf{v}(\mathbf{x}) \in \mathbb{T}_{\mathbf{x}}\mathbb{M}$ is given by the contraction with the volume-form $\boldsymbol{\mu} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{M}^n\,;\Re)$. Hence, if the pull back is performed according to a diffeomorphic map $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$, the PIOLA's transform $P_{\boldsymbol{\varphi}}(\mathbf{v})$ is pointwise defined by the formula:

$$\boldsymbol{\mu}\,(P_{\boldsymbol{\varphi}}(\mathbf{v})) := \boldsymbol{\varphi}\!\downarrow\!(\boldsymbol{\mu}\mathbf{v})\,.$$

Proposition 1.10.2 (Piola's formula and identity) PIOLA*'s transform a vector field* $\mathbf{v} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is pointwise defined by the equivalent formula*

$$P_{\boldsymbol{\varphi}}(\mathbf{v}) = J_{\boldsymbol{\varphi}}\,\boldsymbol{\varphi}\!\downarrow\!\mathbf{v}\,,$$

and fulfills the differential property

$$\operatorname{div} P_{\boldsymbol{\varphi}}(\mathbf{v}) = J_{\boldsymbol{\varphi}}\,(\boldsymbol{\varphi}\!\downarrow\!\operatorname{div}\mathbf{v}) = J_{\boldsymbol{\varphi}}\,(\operatorname{div}\mathbf{v} \circ \boldsymbol{\varphi})\,.$$

In the literature the former is usually referred to as PIOLA*'s formula and the latter as* PIOLA*'s identity, see e.g. [106].*

Proof. The equivalence of the former formula follows from the definition $\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\mu} = J_{\boldsymbol{\varphi}}\,\boldsymbol{\mu}$ of the jacobian $J_{\boldsymbol{\varphi}} = \det(d\boldsymbol{\varphi})$ and from the non-degeneracy of the volume form: $\boldsymbol{\mu}(\mathbf{v}, \mathbf{a}, \mathbf{b}) = 0\,, \quad \forall\,\mathbf{a}, \mathbf{b} \implies \mathbf{v} = 0$. Indeed, by the formula for the pull-back of a contraction:

$$\boldsymbol{\varphi}\!\downarrow\!(\boldsymbol{\mu}\mathbf{v}) = (\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\mu})(\boldsymbol{\varphi}\!\downarrow\!\mathbf{v}) \ = J_{\boldsymbol{\varphi}}\,\boldsymbol{\mu}\,(\boldsymbol{\varphi}\!\downarrow\!\mathbf{v})\,.$$

The latter property is simply the equality between the exterior derivatives of the equality in PIOLA's formula. Indeed the equality

$$\begin{aligned}\operatorname{div}(J_{\boldsymbol{\varphi}}\,\boldsymbol{\varphi}\!\downarrow\!\mathbf{v})\,\boldsymbol{\mu} &= d(\boldsymbol{\mu}\,(J_{\boldsymbol{\varphi}}\,\boldsymbol{\varphi}\!\downarrow\!\mathbf{v})) = d(\boldsymbol{\varphi}\!\downarrow\!(\boldsymbol{\mu}\,\mathbf{v})) = \boldsymbol{\varphi}\!\downarrow\! d(\boldsymbol{\mu}\,\mathbf{v}) \\ &= \boldsymbol{\varphi}\!\downarrow\!((\operatorname{div}\mathbf{v})\,\boldsymbol{\mu}) = (\boldsymbol{\varphi}\!\downarrow\!\operatorname{div}\mathbf{v})\,\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\mu} = J_{\boldsymbol{\varphi}}\,(\operatorname{div}\mathbf{v} \circ \boldsymbol{\varphi})\,\boldsymbol{\mu}\,,\end{aligned}$$

is implied by the commutativity between exterior derivative and pull-back. ■

PIOLA's formula may be expressed in an equivalent way which is known as NANSON's formula concerning the changes of the volume-form of a hypersurface under a diffeomorphism $\boldsymbol{\varphi} \in C^1(\mathbb{M};\mathbb{M})$.

Proposition 1.10.3 (Nanson's formula) *The volume-form* $\boldsymbol{\mu}_\Sigma$ *of a hypersurface* $\Sigma \subset \mathbb{M}$ *in a riemannian manifold* $\{\mathbb{M}, \mathbf{g}\}$ *subject to a diffeomorphism* $\boldsymbol{\varphi} \in C^1(\mathbb{M};\mathbb{M})$ *transforms according to the equivalent relations:*

$$\boldsymbol{\varphi}\!\downarrow\!(\mathbf{gn}_{\boldsymbol{\varphi}(\Sigma)} \otimes \boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)}) = \boldsymbol{\varphi}\!\downarrow\!(\mathbf{gn}_{\boldsymbol{\varphi}(\Sigma)}) \otimes \boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)} = J_{\boldsymbol{\varphi}}\, \mathbf{gn}_\Sigma \otimes \boldsymbol{\mu}_\Sigma\,,$$

$$(\mathbf{gn}_{\boldsymbol{\varphi}(\Sigma)} \circ \boldsymbol{\varphi}) \otimes \boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)} = J_{\boldsymbol{\varphi}}\, \boldsymbol{\varphi}\!\uparrow\!(\mathbf{gn}_\Sigma) \otimes \boldsymbol{\mu}_\Sigma = J_{\boldsymbol{\varphi}}\, \mathbf{g}(d\boldsymbol{\varphi}^{-T}\mathbf{n}_\Sigma) \otimes \boldsymbol{\mu}_\Sigma\,,$$

$$(\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)} \circ \boldsymbol{\varphi}) \otimes \boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)} = J_{\boldsymbol{\varphi}}\, (d\boldsymbol{\varphi}^{-T}\mathbf{n}_\Sigma) \otimes \boldsymbol{\mu}_\Sigma\,,$$

where $\mathbf{n}_\Sigma$ *and* $\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}$ *are the unit normals to the hypersurfaces* Σ *and* $\boldsymbol{\varphi}(\Sigma)$. *The last equality is often referred to as* NANSON*'s formula in the literature.*

Proof. If $\{\mathbf{a}, \mathbf{b}\}$ is a frame at $\mathbb{T}_{\mathbf{x}}\Sigma$, then $\{\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b}\}$ is a frame of $\mathbb{T}_{\mathbf{x}}\mathbb{M}$, so that

$$\boldsymbol{\mu}(\mathbf{v}, \mathbf{a}, \mathbf{b}) = \langle \mathbf{gn}_\Sigma, \mathbf{v} \rangle\, \boldsymbol{\mu}(\mathbf{n}_\Sigma, \mathbf{a}, \mathbf{b}) = (\mathbf{gn}_\Sigma \otimes \boldsymbol{\mu}_\Sigma)(\mathbf{v}, \mathbf{a}, \mathbf{b})\,, \quad \forall\, \mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{M}\,.$$

Moreover, $\{\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{a}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{b}\}$ is a frame at $\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{M}$, so that

$$\begin{aligned}
\boldsymbol{\varphi}\!\downarrow\!(\mathbf{gn}_{\boldsymbol{\varphi}(\Sigma)} \otimes \boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)})(\mathbf{v}, \mathbf{a}, \mathbf{b}) &= (\mathbf{gn}_{\boldsymbol{\varphi}(\Sigma)} \otimes \boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)})(\boldsymbol{\varphi}\!\uparrow\!\mathbf{v}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{a}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{b}) \\
&= \mathbf{g}(\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{v})\, \boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)}(\boldsymbol{\varphi}\!\uparrow\!\mathbf{a}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{b}) \\
&= \mathbf{g}(\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{v})\, \boldsymbol{\mu}(\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{a}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{b}) \\
&= \boldsymbol{\mu}(\boldsymbol{\varphi}\!\uparrow\!\mathbf{v}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{a}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{b}) = J_{\boldsymbol{\varphi}}\, \boldsymbol{\mu}(\mathbf{v}, \mathbf{a}, \mathbf{b})\,.
\end{aligned}$$

The other formulas may be readily obtained by relying on the property that an equality between tensor products of the same kind still holds if alteration or push operations are performed on each of its members, and by recalling the formula $\boldsymbol{\varphi}\!\uparrow\!(\mathbf{gn}_\Sigma) = \mathbf{g}(d\boldsymbol{\varphi}^{-T}\mathbf{n}_\Sigma)$ provided in section 1.2.6. ■

By acting both sides of the second of NANSON's formulas on the normal $\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}$ we get the ratio between the hypersurface volume-form and its pull-back:

$$\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)} = J_{\boldsymbol{\varphi}}\, \mathbf{g}(d\boldsymbol{\varphi}^{-T}\mathbf{n}_\Sigma, \mathbf{n}_{\boldsymbol{\varphi}(\Sigma)})\, \boldsymbol{\mu}_\Sigma = J_{\boldsymbol{\varphi}}\, \mathbf{g}(d\boldsymbol{\varphi}^{-1}\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}, \mathbf{n}_\Sigma)\, \boldsymbol{\mu}_\Sigma\,.$$

which, rewritten as

$$\boldsymbol{\varphi}\downarrow(\boldsymbol{\mu}\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}) = \boldsymbol{\varphi}\downarrow\boldsymbol{\mu}_{\boldsymbol{\varphi}(\Sigma)} = J_{\boldsymbol{\varphi}}\,\mathbf{g}(\boldsymbol{\varphi}\downarrow\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}, \mathbf{n}_\Sigma)\,\boldsymbol{\mu}_\Sigma = \boldsymbol{\mu}(J_{\boldsymbol{\varphi}}\boldsymbol{\varphi}\downarrow\mathbf{n}_{\boldsymbol{\varphi}(\Sigma)})\,,$$

is PIOLA's formula with $\mathbf{v} = \mathbf{n}_{\boldsymbol{\varphi}(\Sigma)}$.

The equivalence between PIOLA's and NANSON's formulas is apparent from the fact that both stem from the very definition of the Jacobian.

1.10.8 Lamb's formula

By taking the time-derivative of NANSON's formula we get a well-known formula, due to LAMB, which provides a tool for the evaluation of the rate of change of the flux of a time-dependent vector field thru a hypersurface flowing in a euclidean space. LAMB's formula will be derived in the next Proposition 1.10.4 and the consequent surface transport formula is contributed in Proposition 1.10.8.

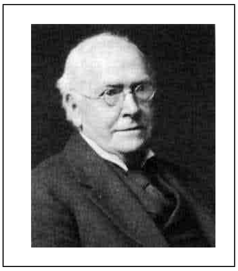

Figure 1.50: Horace Lamb (1849 - 1934)

A more general transport formula for a hypersurface flowing on a riemannian manifold will be provided in Proposition 1.10.8. It may be adopted as an alternative to the one provided in Proposition 1.10.1.

Proposition 1.10.4 (Lamb's formula) *Let us consider a hypersurface* $\Sigma \subset \mathbb{M}$ *flowing in a euclidean space* $\{\mathbb{M}\,,\mathbf{g}\}$ *dragged by flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$. *Then*

$$\partial_{\tau=t}\,(\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t}) \otimes \mathbf{Fl}^{\mathbf{v}}_{\tau,t}\downarrow\boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} = ((\operatorname{div}\mathbf{v}_t)\,\mathbf{I} - d\mathbf{v}_t^T)\,\mathbf{n}_\Sigma \otimes \boldsymbol{\mu}_\Sigma\,.$$

Proof. By differentiating NANSON's formula with respect to time we get

$$\begin{aligned}\partial_{\tau=t}\ ((\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t}) \otimes \mathbf{Fl}^{\mathbf{v}}_{\tau,t}\downarrow \boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}) &= \partial_{\tau=t}\ (J_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}}\, d\mathbf{Fl}^{\mathbf{v}}_{\tau,t}{}^{-T}\mathbf{n}_\Sigma \otimes \boldsymbol{\mu}_\Sigma) \\ &= ((\operatorname{div}\mathbf{v}_t)\,\mathbf{n}_\Sigma + \partial_{\tau=t}\ d\mathbf{Fl}^{\mathbf{v}}_{\tau,t}{}^{-T}\,\mathbf{n}_\Sigma) \otimes \boldsymbol{\mu}_\Sigma \\ &= ((\operatorname{div}\mathbf{v}_t)\,\mathbf{I} - d\mathbf{v}_t^T)\,\mathbf{n}_\Sigma \otimes \boldsymbol{\mu}_\Sigma\,,\end{aligned}$$

which is the result. ■

A direct application of LAMB's formula yields the surface transport formula.

Proposition 1.10.5 (Surface transport formula) *Let us consider a hypersurface* $\Sigma \subset \mathbb{M}$ *flowing in a euclidean space* $\{\mathbb{M}\,,\mathbf{g}\}$ *dragged by a flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *. Then, for any time-dependent vector field* $\mathbf{a}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *:*

$$\partial_{\tau=t}\ \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \mathbf{g}(\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}, \mathbf{a}_\tau)\boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} = \int_\Sigma \mathbf{g}(\dot{\mathbf{a}}_t + ((\operatorname{div}\mathbf{v}_t)\,\mathbf{I} - d\mathbf{v}_t)\,\mathbf{a}_t, \mathbf{n}_\Sigma)\,\boldsymbol{\mu}_\Sigma\,,$$

where $\dot{\mathbf{a}}_t := \partial_{\tau=t}\ (\mathbf{a}_\tau \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t}) = \partial_{\tau=t}\ \mathbf{a}_\tau + d\mathbf{a}_t \cdot \mathbf{v}_t$ *.*

Proof. From the formula

$$\int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \mathbf{g}(\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}, \mathbf{a}_\tau)\,\boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} = \int_\Sigma \mathbf{g}(\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t}, \mathbf{a}_\tau \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t})\,\mathbf{Fl}^{\mathbf{v}}_{\tau,t}\downarrow\boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}$$

$$= \int_\Sigma ((\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t}) \otimes \mathbf{Fl}^{\mathbf{v}}_{\tau,t}\downarrow\boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)})(\mathbf{g}\mathbf{a}_\tau \circ \mathbf{Fl}^{\mathbf{v}}_{\tau,t})\,,$$

taking the time-derivative, we get

$$\begin{aligned}\partial_{\tau=t}\ \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \mathbf{g}(\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}, \mathbf{a}_\tau)\,\boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} &= \int_\Sigma (((\operatorname{div}\mathbf{v}_t)\,\mathbf{I} - d\mathbf{v}_t^T)\,\mathbf{n}_\Sigma \otimes \boldsymbol{\mu}_\Sigma)\,\mathbf{g}\mathbf{a}_t \\ &+ \int_\Sigma (\mathbf{n}_\Sigma \otimes \boldsymbol{\mu}_\Sigma)\,\mathbf{g}\dot{\mathbf{a}}_t\,,\end{aligned}$$

and the result follows since $\mathbf{g}(\mathbf{a}_t, d\mathbf{v}_t^T \cdot \mathbf{n}_\Sigma) = \mathbf{g}(d\mathbf{v}_t \cdot \mathbf{a}_t, \mathbf{n}_\Sigma)$. ■

From the surface transport formula of Proposition 1.10.5, setting $\mathbf{a}_\tau = \mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}$ and hence $\mathbf{a}_t = \mathbf{n}_\Sigma$, and observing that

$$\mathbf{g}(\mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}, \mathbf{n}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)}) = \mathbf{g}(\mathbf{n}_\Sigma, \mathbf{n}_\Sigma) = 1,$$

$$\mathbf{g}(\mathbf{n}_\Sigma, \dot{\mathbf{n}}_\Sigma) = 0\,,$$

we get the following formula for the rate of change of the global area of the flowing hypersurface

$$\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\mu}_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} = \int_{\Sigma} (\operatorname{div} \mathbf{v}_t - \mathbf{g}(d\mathbf{v}_t \cdot \mathbf{n}_\Sigma, \mathbf{n}_\Sigma))\, \boldsymbol{\mu}_\Sigma \,,$$

which tells us that

- The rate of change of the area of the flowing hypersurface is the integral over the hypersurface of the difference between the volumetric dilatation-rate induced by the flow and the dilatation-rate in the direction normal to the hypersurface.

1.10.9 Hypersurface transport

Proposition 1.10.6 (Hypersurface transport formula) *Let us consider a hypersurface* $\Sigma \subset \mathbb{M}$ *flowing in a manifold* $\{\mathbb{M}, \mathbf{g}\}$ *dragged by a flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *. Then the rate of change of the flux thru* Σ *of any time-dependent vector field* $\mathbf{a}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is given by the formulas*

$$\begin{aligned}
\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\mu}\mathbf{a}_\tau &= \int_\Sigma \boldsymbol{\mu}(\mathbf{a}'_t + (\operatorname{div} \mathbf{v})\, \mathbf{a} + \mathcal{L}_{\mathbf{v}}\mathbf{a})\,, \\
&= \int_\Sigma \boldsymbol{\mu}\mathbf{a}'_t + \int_\Sigma d(\boldsymbol{\mu}\mathbf{a})\mathbf{v} + \int_{\partial\Sigma} \boldsymbol{\mu}\mathbf{a}\mathbf{v} \\
&= \int_\Sigma \mathbf{g}(\mathbf{a}'_t + (\operatorname{div} \mathbf{v})\, \mathbf{a} + \mathcal{L}_{\mathbf{v}}\mathbf{a}\,, \mathbf{n}_\Sigma)\, \boldsymbol{\mu}_\Sigma \\
&= \int_\Sigma \mathbf{g}(\dot{\mathbf{a}}_t + (\operatorname{div} \mathbf{v})\, \mathbf{a} - \nabla_{\mathbf{a}}\mathbf{v}\,, \mathbf{n}_\Sigma)\, \boldsymbol{\mu}_\Sigma \,,
\end{aligned}$$

where $\mathbf{a}'_t = \partial_{\tau=t}\, \mathbf{a}_\tau$ *is the partial time-derivative and* $\dot{\mathbf{a}}_t := \mathbf{a}' + \nabla_{\mathbf{v}}\mathbf{a}$ *is the covariant time-derivative with respect to a torsion-free connection.*

Proof. By the transport formula we get

$$\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\mu}\mathbf{a}_\tau = \int_\Sigma \boldsymbol{\mu}\mathbf{a}'_t + \mathcal{L}_{\mathbf{v}}(\boldsymbol{\mu}\mathbf{a}) = \int_\Sigma \boldsymbol{\mu}\mathbf{a}'_t + (\mathcal{L}_{\mathbf{v}}\boldsymbol{\mu})\mathbf{a} + \boldsymbol{\mu}(\mathcal{L}_{\mathbf{v}}\mathbf{a})\,.$$

Then the first formula follows by the definition of divergence: $\mathcal{L}_{\mathbf{v}}\boldsymbol{\mu} = (\operatorname{div} \mathbf{v})\boldsymbol{\mu}$. The second formula stems from the homotopy formula:

$$\mathcal{L}_{\mathbf{v}}(\boldsymbol{\mu}\mathbf{a}) = d(\boldsymbol{\mu}\mathbf{a}\mathbf{v}) + d(\boldsymbol{\mu}\mathbf{a})\mathbf{v}\,,$$

and Stokes' theorem. The third formula is based on the equalities

$$\int_{\Sigma} \boldsymbol{\mu}\mathbf{a}'_t = \int_{\Sigma} \mathbf{g}(\mathbf{a}'_t, \mathbf{n}_{\Sigma})\, \boldsymbol{\mu}\mathbf{n}_{\Sigma}$$

$$\int_{\Sigma} \boldsymbol{\mu}(\mathcal{L}_{\mathbf{v}}\mathbf{a}) = \int_{\Sigma} \mathbf{g}(\mathcal{L}_{\mathbf{v}}\mathbf{a}, \mathbf{n}_{\Sigma})\, \boldsymbol{\mu}\mathbf{n}_{\Sigma}\,,$$

and the fourth, being

$$\mathcal{L}_{\mathbf{v}}\mathbf{a} = \nabla_{\mathbf{v}}\mathbf{a} - \nabla_{\mathbf{a}}\mathbf{v}\,.$$

is valid for a torsion-free connection ■

Setting $\mathbf{a}_t = \mathbf{n}_{\Sigma}$ we infer the following result which generalizes LAMB's formula to riemannian manifolds.

Proposition 1.10.7 (Hypersurface area change) *Let us consider a hypersurface* $\Sigma \subset \mathbb{M}$ *flowing in a riemannian manifold* $\{\mathbb{M}\,, \mathbf{g}\}$ *dragged by a flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$. *Then the rate of change of the global hypersurface area is given by*

$$\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\mu}_{\Sigma} = \int_{\Sigma} (\mathrm{div}\,\mathbf{v} - \tfrac{1}{2}(\mathcal{L}_{\mathbf{v}}\mathbf{g})(\mathbf{n}_{\Sigma}, \mathbf{n}_{\Sigma}))\, \boldsymbol{\mu}_{\Sigma}$$

$$= \int_{\Sigma} (\mathrm{tr}(\nabla\mathbf{v}) - \mathbf{g}((\mathrm{sym}\,\nabla\mathbf{v})\mathbf{n}_{\Sigma}, \mathbf{n}_{\Sigma}))\, \boldsymbol{\mu}_{\Sigma}\,,$$

Proof. Since $\mathbf{n}_{\Sigma}$ doen't depend explicitly on time, we have that $\mathbf{n}'_{\Sigma} = 0$. Moreover

$$2\,\mathbf{g}(\mathcal{L}_{\mathbf{v}}\mathbf{n}_{\Sigma}, \mathbf{n}_{\Sigma}) = \mathcal{L}_{\mathbf{v}}(\mathbf{g}(\mathbf{n}_{\Sigma}, \mathbf{n}_{\Sigma})) - (\mathcal{L}_{\mathbf{v}}\mathbf{g})(\mathbf{n}_{\Sigma}, \mathbf{n}_{\Sigma}) = -(\mathcal{L}_{\mathbf{v}}\mathbf{g})(\mathbf{n}_{\Sigma}, \mathbf{n}_{\Sigma})\,,$$

since $\mathbf{g}(\mathbf{n}_{\Sigma}, \mathbf{n}_{\Sigma}) = 1$, and the formula follows from Proposition 1.10.6. ■

1.10.10 Surface transport

Proposition 1.10.8 (Surface transport formula) *Let us consider a 2D surface* $\Sigma \subset \mathbb{M}$ *flowing in a 3D riemannian manifold* $\{\mathbb{M}\,, \mathbf{g}\}$ *dragged by a flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$. *Then the rate of change of the flux thru* Σ *of a time-dependent vector field* $\mathbf{a}_t \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *is given by the formula*

$$\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\mu}\cdot\mathbf{a}_{\tau} = \int_{\Sigma} \boldsymbol{\mu}\cdot(\mathbf{a}'_t + \mathrm{rot}\,(\mathbf{a}\times\mathbf{v}) + (\mathrm{div}\,\mathbf{a})\,\mathbf{v})\,.$$

Proof. Recalling that

$$\begin{aligned}
&\boldsymbol{\mu} \cdot \mathbf{a} \cdot \mathbf{v} = \mathbf{g} \cdot (\mathbf{a} \times \mathbf{v}) \,, \\
&d(\boldsymbol{\mu} \cdot \mathbf{a}) = (\operatorname{div} \mathbf{a}) \, \boldsymbol{\mu} \,, \\
&d(\mathbf{g} \cdot \mathbf{w}) = \boldsymbol{\mu} \cdot (\operatorname{rot} \mathbf{w}) \,,
\end{aligned}$$

setting $\mathbf{w} = \mathbf{a} \times \mathbf{v}$ and substituting in the transport formula of Proposition 1.10.6 rewritten as

$$\begin{aligned}
\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\mu} \cdot \mathbf{a}_\tau &= \int_\Sigma \boldsymbol{\mu} \cdot \mathbf{a}'_t + \int_\Sigma d(\boldsymbol{\mu} \cdot \mathbf{a}) \cdot \mathbf{v} + \int_{\partial\Sigma} \boldsymbol{\mu} \cdot \mathbf{a} \cdot \mathbf{v} \\
&= \int_\Sigma \boldsymbol{\mu} \cdot \mathbf{a}'_t + \int_\Sigma d(\boldsymbol{\mu} \cdot \mathbf{a}) \cdot \mathbf{v} + \int_\Sigma d(\boldsymbol{\mu} \cdot \mathbf{a} \cdot \mathbf{v}) \,,
\end{aligned}$$

we get the result. ■

Chapter 2

Dynamics

2.1 Introduction

Classical dynamics may be conventionally considered to be born about 1687 with NEWTON's *Principia* and grew up to a well-established theory in the fundamental works on the subject by EULER, LAGRANGE, HAMILTON and JACOBI during the XVIII century and the first half of the XIX century. EULER's law refers to motions of arbitrary bodies in the eucliden space but, in most modern textbooks, the presentation of the foundations of dynamics is still developed in the spirit of rigid body dynamics and with reference to finite dimensional systems [6]. Extensions to continuous systems are illustrated in [2], [106], [3] by assuming that the configuration manifold is modeled on a BANACH space. Anyway these treatments essentially reproduce the formal structure of the dynamics of finite dimensional systems with suitable technical changes required by functional analysis. The point of view followed in the present treament is described hereafter. The EULER-CAUCHY model of continua, which is the worldwide standard, requires to build up the axiomatics of classical dynamics as a discipline which investigates the motions of continuous bodies in the eucliden space, possibly subject to kinematical constraints which are assumed to describe a fibered manifold of admissible states. In the definition of dynamical equilibrium the test fields are isometric fields of virtual velocities of the body, according to the point of view expressed by JOHANN BERNOULLI in a famous letter to PIERRE VARIGNON dated 1717. The assumption concerning the isometry of test fields of virtual velocities expresses the basic physical idea that equilibrium is independent of the material which a body is made of, and the virtuality of the test fields means that equilibrium does not take into account the time dependency of the constraints defining the manifold of admissible states. It is then apparent that, to comply with the original ideas of the old masters, it is compelling to express the condition of dynamical equilibrium in variational form. We choose to take HAMILTON's action principle, inspired by earlier ideas by FERMAT and HUYGENS in optics, as the basic axiom of dynamics since it has the pleasant flavour of an extremality property and, much more than this, it leads in a natural and direct way to the most general formulation of Lagrangian dynamics. In this respect we quote from [1], Part II *Analytical Dynamics*, section 3.8 *Variational Principles in Mechanics*, the following opinion: *Historically, variational principles have played a fundamental role in the evolution of mathematical models in mechanics but in the last few sections we have obtained the bulk of classical mechanics without a single reference to the calculus of variations. In principle, we may envision two equivalent models for mechanics. In the first, we may take the Hamiltonian or Lagrangian equations as an axiom and, if we wish, obtain variational principles as theorems. In the second, we may assume variational principles and derive the Hamiltonian or Lagrangian equations as theorems. We prefer the first because it is quite difficult to be rigorous in the calculus of variations, and in practise, the variational principles are not necessary to the prediction of the model. In fact, in the model-theoretic view, we consider the variational principles important primarily to the inductive formation of the theory. After this most basic function, they do not have a crucial role within the theory.* Probably, this point of view, shared by other authoritative authors, has contributed to the wide acceptance of the classical LAGRANGE's and HAMILTON's equations of motion as the starting point for subsequent developments of the theory. But, apart from personal tastes, there is a drawback shared by most usual presentations of the fundamentals of dynamics. In fact, LAGRANGE's equations, either assumed as axioms or derived as differential conditions equivalent to HAMILTON's action principle, are always formulated in coordinates since their expression involve partial derivatives which are well-defined in a linear space. On the other hand, HAMILTON's equations have been translated in invariant form on a manifold [2], [6], [106], but their explicit expression is always given in coordinates. Moreover, LAGRANGE's and HAMILTON's equations are both written in a non-variational form so that their validity is restricted to rigid body dynamics or more generally to perfect dynamical systems (see section 2.1.8 for the definition). Our original plan was to find an explicit expression of the fundamental one-form appearing in HAMILTON's equations without any recourse to coordinates. This goal has been achieved by a recourse to concepts of calculus on manifolds, the suitable mathematical tool for dynamics of continuous systems undergoing motions in a nonlinear configuration manifold. A detailed account of the basic concepts, due to Marius Sophus Lie, Henri Poincaré and Elie Cartan has been provided in chapter 1 and may be found in [2], [27], [6], [180], [106], [177].

In section 2.1.2 we provide an abstract statement of the action principle as a stationarity condition for a signed-lenght of a path in which the variations are left free to move the end points. In deriving the differential condition of stationarity the Reynolds transport theorem, the Ampère-Hankel-Kelvin transform, usually dubbed Stokes's formula, and its expression in terms of differential forms due to Poincaré, the Cartan's magic formula and the Palais' formula for the exterior derivative of a differential one-form, are the playmates. In the remaining sections the abstract theory is applied to continuum dynamics. Hamilton's action principle is restated in variational terms by introducing a lagrangian one-form in the velocity-time phase-space, and the stationarity condition is expressed in terms of its exterior derivative. The general variational law of dynamics is derived by providing an explicit expression of the exterior derivative of the one-form in terms of the Lagrangian of the system. To get the result, the key property is the tensoriality of the exterior derivative so that Palais' formula [138] may be applied by envisaging an expedient extension of the time-speed of the trajectory at the actual configuration-velocity point in the velocity phase-space. The new form of the law of dynamics provides the most general formulation of the governing rules in terms of the Lagrangian of the system and, to the author's knowledge, is not quoted in the literature. A generalized version of the celebrated Noether's theorem [129] on symmetry of the Lagrangian and invariance along the trajectory is implied as an immediate, simple corollary. Remarkably, the expression of the general law of dynamics requires no special connection to be defined on the configuration manifold. In section 2.3 we show that, if the configuration manifold is endowed with an affine connection, the general law of dynamics may be rewritten in terms of the *fiber derivative* and the *base derivative* of the Lagrangian and that the standard Lagrange's form is recovered, if the torsion of the connection vanishes. The proof of this result is enlighting since it reveals that the steps of reasoning could be followed backwards to get the general law of dynamics from the classical Lagrange's expression. However a direct discovery of the right back-steps appear to be much harder to envisage than the opposite direct-step reasoning. This is likely the reason why this track has not been followed before. The definition of the *base derivative* of a functional on the tangent bundle according to a given parallel transport, is an original idea and provides the key tool to get results independent of coordinates. It is thus possible to prove the equivalence between LAGRANGE's and HAMILTON's formulations by showing that the sum of the base derivatives of the Lagrangian and the Hamiltonian vanishes for any chosen connection.

Only after having obtained these results, I realized that the general law of dynamics can be reached, in a by far simpler way, by a skillful reformulation of HAMILTON's action principle in which the assumption of fixed initial and final configurations is substituted by a proper boundary term. The way to such a reformulation is however revealed by a less direct analysis performed by the tools of calculus on manifolds. This fact could explain why a simple proof of a generalized NOETHER's theorem was not envisaged before. A main innovative feature of the analysis developed in the present paper is the explicit introduction of the rigidity constraint from the very beginning. This is in the spirit of the basic definition of dynamical equilibrium. To take account of the rigidity constraint, it is compelling to state principles and laws of dynamics in variational form and this leads, in addition, to develop a completely general and coordinate-free theory.

2.1.1 Tools from calculus on manifolds

When dealing with a nonlinear manifold, most usual rules of calculus in linear spaces are no more available and the general concepts and methods of calculus on manifolds must be resorted to [2], [27], [36], [180], [106], [3], [143], [80]. Manifolds are nonlinear geometrical entities which are locally linear. This means that they admit a covering made of intesecting open subsets which are mapped by diffeomorphic charts onto open sets of a BANACH space, a complete normed linear space. Each local chart endows the related open subset of the manifold of the induced topology and an atlas of compatible charts provides a topology for the whole manifold. Anyway, physically meaningful concepts and results must be independent of the recourse to a particular description by means of charts. A theoretical approach which does not make reference to charts is then appealing to get directly physically significant results. The first issue to be stressed is that at each point of a nonlinear manifold there is an attached tangent space. Since linear operations are only defined on vectors of the same tangent space, vectors belonging to tangent spaces at distinct points cannot be compared one-another, unless a special way of connecting vectors in distinct tangent spaces is defined. This is the concept of connection or parallel transport on a manifold. Each local chart induces on the manifold a distant parallel transport which is inherited by the translational transport in the linear model space. The trouble is that there is not a unique way to endow a manifold with a connection. As a significant example, we quote the dynamical notion of acceleration which makes sense in a euclidean space since it is tacitly understood that the connection is provided by the standard translation operation. In a nonlinear configuration manifold the notion of acceleration depends on the choosen connection. To get rid of this choice, we have to consider a parametrized curve of velocities, which are pairs of base points and vectors of the relevant tangent space, and to compute the tangent vector at each point of the curve. Another issue concerns integration over nonlinear manifolds which is properly defined for volume-forms on compact subsets of a finite dimensional submanifold. Volume forms are alternating k-linear scalar-valued functions defined on the tangent spaces to a kD submanifold. Since fundamental concepts of continnum dynamics are defined in terms of integrals over finite dimensional submanifols of the ambient manifold, a variational approch is compelling and natural because it leads to integration of volume forms. Integration of tensor fields, which do not take point-values on a given linear space when evaluated on a basis of tangent vectors, is meaningless, being addition of their values not defined. A variational approach leads naturally to integration of volume forms. After these general premises, we summarize hereafter concepts, results and notations of calculus on manifolds which will be referred to in the sequel. We consider a non-finite dimensional differentiable manifold $\mathbb{M}$ modeled on a linear Banach's space E. The tangent bundle $\mathbb{TM}$ is the collection of the tangent spaces at the points of $\mathbb{M}$ and the dual cotangent bundle $\mathbb{T}^*\mathbb{M}$ is the collection of the cotangent spaces, i.e. of the linear spaces of bounded linear forms on the tangent spaces. Push-forward and its inverse, the pull-back, of scalar, vector and tensor fields due to a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{M};\mathbb{M})$ are respectively denoted by $\boldsymbol{\varphi}\!\uparrow$ and $\boldsymbol{\varphi}\!\downarrow$. The usual notation in differential geometry is $\boldsymbol{\varphi}_* = \boldsymbol{\varphi}\!\uparrow$ and $\boldsymbol{\varphi}^* = \boldsymbol{\varphi}\!\downarrow$ but then too many stars appear in the geometrical sky (duality, Hodge operator). A dot $\cdot$ denotes linear dependence on subsequent arguments and the crochet $\langle \, , \rangle$ denotes a duality pairing. The variational analysis performed in this paper is mainly based on the following tools of calculus on manifolds which have been illustrated in chapter 1. The first tool is the Poincaré-Stokes' formula which states that the integral of a differential $(k-1)$-form $\boldsymbol{\omega}^{k-1}$ on the boundary chain $\partial\Sigma$ of a kD submanifold Σ of $\mathbb{M}$ is equal to the integral of its exterior derivative $d\boldsymbol{\omega}^{k-1}$, a differential k-form, on Σ i.e.

$$\int_{\Sigma} d\boldsymbol{\omega}^{k-1} = \oint_{\partial\Sigma} \boldsymbol{\omega}^{k-1} \,.$$

The second tool is LIE's derivative of a vector field $\mathbf{w} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$ along a flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M};\mathbb{M})$ with velocity $\mathbf{v} = \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$:

$$\mathcal{L}_{\mathbf{v}}\mathbf{w} = \partial_{\lambda=0}\,(\boldsymbol{\varphi}_\lambda\!\downarrow\!\mathbf{w})\,,$$

which is equal to the antisymmetric LIE-bracket: $\mathcal{L}_{\mathbf{v}}\mathbf{w} = [\mathbf{v},\mathbf{w}] = -[\mathbf{w},\mathbf{v}]$ defined by: $d_{[\mathbf{v},\mathbf{w}]}f = d_{\mathbf{v}}d_{\mathbf{w}}f - d_{\mathbf{w}}d_{\mathbf{v}}f$, for any $f \in \mathrm{C}^2(\mathbb{M};\Re)$.

The LIE derivative of a differential form $\boldsymbol{\omega}^k \in \mathrm{C}^1(\mathbb{M};\Lambda^k(\mathbb{TM}))$ is similarly defined by $\mathcal{L}_{\mathbf{v}}\boldsymbol{\omega}^k = \partial_{\lambda=0}\,(\boldsymbol{\varphi}_\lambda\!\downarrow\!\boldsymbol{\omega}^k)$. The third tool is REYNOLDS' transport formula:

$$\int_{\boldsymbol{\varphi}_\lambda(\Sigma)} \boldsymbol{\omega}^k = \int_\Sigma \boldsymbol{\varphi}_\lambda\!\downarrow\!\boldsymbol{\omega}^k \implies \partial_{\lambda=0}\int_{\boldsymbol{\varphi}_\lambda(\Sigma)} \boldsymbol{\omega}^k = \int_\Sigma \mathcal{L}_{\mathbf{v}}\,\boldsymbol{\omega}^k\,,$$

and the fourth tool is the *extrusion formula*

$$\partial_{\lambda=0}\int_{\boldsymbol{\varphi}_\lambda(\Sigma)} \boldsymbol{\omega}^k = \int_\Sigma (d\boldsymbol{\omega}^k)\mathbf{v} + \int_{\partial\Sigma} \boldsymbol{\omega}^k\mathbf{v}\,,$$

and the related CARTAN's magic formula (or *homotopy formula*):

$$\mathcal{L}_{\mathbf{v}}\,\boldsymbol{\omega}^k = (d\boldsymbol{\omega}^k)\mathbf{v} + d(\boldsymbol{\omega}^k\mathbf{v})\,,$$

where the $(k-1)$-form $\boldsymbol{\omega}^k\mathbf{v} = \boldsymbol{\omega}^k \cdot \mathbf{v}$ is the contraction performed by taking $\mathbf{v}$ as the first argument of the form $\boldsymbol{\omega}^k$. The homotopy formula may be readily inverted to get PALAIS formula for the exterior derivative. Indeed, by LEIBNIZ rule for the LIE derivative, we have that, for any two vector fields $\mathbf{v},\mathbf{w} \in \mathrm{C}^1(\mathbb{M};\mathbb{TM})$:

$$\begin{aligned} d\boldsymbol{\omega}^1 \cdot \mathbf{v} \cdot \mathbf{w} &= (\mathcal{L}_{\mathbf{v}}\,\boldsymbol{\omega}^1) \cdot \mathbf{w} - d(\boldsymbol{\omega}^1\mathbf{v}) \cdot \mathbf{w} \\ &= d_{\mathbf{v}}\,(\boldsymbol{\omega}^1\mathbf{w}) - \boldsymbol{\omega}^1 \cdot [\mathbf{v},\mathbf{w}] - d_{\mathbf{w}}\,(\boldsymbol{\omega}^1\mathbf{v})\,. \end{aligned}$$

The expression at the r.h.s. of PALAIS formula fulfills the tensoriality criterion, as quoted in Lemma 1.2.1 on page 26. A proof may be found in [180], [80]. The exterior derivative of a differential one-form is thus well-defined as a differential two-form, since its value at a point depends only on the values of the argument vector fields at that point.

The same algebra may be repeatedly applied to deduce PALAIS formula for the exterior derivative of a k-form.

2.1.2 Abstract results

Let a status of the system be described by a point of a manifold $\mathbb{M}$, the *state space*. In both theory and applications, there are many instances in which it is compelling to consider fields which are only piecewise regular on $\mathbb{M}$. To this end, we give the following

Definition 2.1.1 *A* ***patchwork*** $\mathrm{PAT}(\mathbb{M})$ *on* $\mathbb{M}$ *is a finite family of disjoint open subsets of* $\mathbb{M}$ *such that the union of their closures is a covering of* $\mathbb{M}$. *The closure of each subset in the family is called an* ***element*** *of the patchwork.*

The disjoint union of the boundaries of the elements, deprived of the boundary of $\mathbb{M}$, is the set of *singularity interfaces* $\mathrm{SING}(\mathbb{M})$ associated with the patchwork $\mathrm{PAT}(\mathbb{M})$. A field is said to be *piecewise regular* on $\mathbb{M}$ if it is regular, say C^1, on each element of a patchwork on $\mathbb{M}$ which is called a *regularity patchwork*.

In the family of all patchworks on $\mathbb{M}$ we may define a *partial ordering* by saying that a patchwork PAT_1 is *finer than* a patchwork PAT_2 if every element of PAT_1 is included in an element of PAT_2.

Given two patchworks it is always possible to find a patchwork *finer than* both by taking as elements the nonempty pairwise intersections of their elements. This property is expressed by saying that the family of all patchworks on $\mathbb{M}$ is an inductive set.

Then, let $\mathrm{PAT}(I)$ be a time-patchwork, that is a patchwork of a time interval I. The evolution of the system along a piecewise regular trajectory $\Gamma \in \mathrm{C}^1(\mathrm{PAT}(I)\,;\mathbb{M})$ is assumed to be governed by a variational condition on its signed-length, evaluated according to the piecewise regular differential *action one-form* $\boldsymbol{\omega}^1 \in \Lambda^1(\mathrm{PAT}(\mathbb{M})\,;\mathbb{T}^*\mathbb{M})$, with $\mathrm{PAT}(\mathbb{M})$ a regularity patchwork.

We assume, without loss in generality, that the trajectory $\Gamma \in \mathrm{C}^1(\mathrm{PAT}(I)\,;\mathbb{M})$ is regular in each element of the time-patchwork $\mathrm{PAT}(I)$. The test fields for the variational condition are vector fields with values in a subbundle $\mathrm{TEST}\mathbb{M} \subset \mathbb{T}\mathbb{M}$, called the *test-subbundle*. The restriction of the test-subbundle $\mathrm{TEST}\mathbb{M} \subset \mathbb{T}\mathbb{M}$ to $\boldsymbol{\Gamma} := \Gamma(I)$ is denoted by $\mathrm{TEST}\mathbb{M}(\boldsymbol{\Gamma})$.

2.1.3 Action principle and Euler conditions

Definition 2.1.2 (Action integral) *The action integral, of a piecewise regular path* $\Gamma \in \mathrm{C}^1(\mathrm{PAT}(I)\,;\mathbb{M})$ *in the state-space, is the signed-length of the* $1D$ *oriented submanifold* $\boldsymbol{\Gamma} := \Gamma(I)$, *evaluated according to the action one-form:*

$$\int_{\boldsymbol{\Gamma}} \boldsymbol{\omega}^1 \,.$$

A general statement of the action principle requires to define properly the *virtual flows* along which the trajectory is assumed to be varied.

To this end we denote by $\mathbb{T}_{\mathbf{\Gamma}}\mathbb{M}$ the bundle which is the restriction of the tangent bundle $\mathbb{T}\mathbb{M}$ to the path $\mathbf{\Gamma}$.

Definition 2.1.3 (Virtual flows) *The virtual flows of* $\mathbf{\Gamma}$ *are flows* $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathrm{PAT}(\mathbb{M})\,;\mathrm{PAT}(\mathbb{M}))$ *whose velocities* $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathbb{T}_{\mathbf{\Gamma}}\mathbb{M}$ *are tangent to interelement boundaries of the patchwork* $\mathrm{PAT}(\mathbb{M})$.

Velocities of the virtual flows are called *virtual velocities.* The linear space of virtual velocities at $\mathbf{\Gamma}$ will be denoted by $\mathrm{VIRT}(\mathbf{\Gamma})$. The linear space of virtual velocities at $\mathbf{\Gamma}$ taking value in the test subbundle $\mathrm{TEST}\mathbb{M}(\mathbf{\Gamma})$ will be denoted by $\mathrm{VTEST}(\mathbf{\Gamma})$.

Definition 2.1.4 (Action principle) *A trajectory of the system governed by a piecewise regular differential one-form* $\boldsymbol{\omega}^1$ *on* $\mathbb{M}$, *is a piecewise regular path* $\Gamma \in \mathrm{C}^1(\mathrm{PAT}(I)\,;\mathbb{M})$ *such that the action integral meets the variational condition:*

$$\partial_{\lambda=0} \int_{\boldsymbol{\varphi}_\lambda \circ \mathbf{\Gamma}} \boldsymbol{\omega}^1 = \oint_{\partial \mathbf{\Gamma}} \boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} \,,$$

for all virtual flows $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *whose velocities* $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda \in \mathrm{VTEST}(\Gamma)$ *take values in the test subbundle.*

This means that the initial rate of increase of the $\boldsymbol{\omega}^1$-length of the trajectory Γ along a virtual flow is equal to the outward flux of virtual velocities at end points. Denoting by $\mathbf{x}_1$ and $\mathbf{x}_2$ the initial and final end points of $\mathbf{\Gamma}$, we have that $\partial\mathbf{\Gamma} = \mathbf{x}_2 - \mathbf{x}_1$ (a 0-chain) and the boundary integral may be written as

$$\oint_{\partial \mathbf{\Gamma}} \boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} = (\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}})(\mathbf{x}_2) - (\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}})(\mathbf{x}_1)\,.$$

The action principle is purely geometrical since it characterizes the trajectory Γ to within an arbitrary reparametrization.

In geometrical optics the action principle is FERMAT principle, the action functional is the *eikonal* functional and its level sets are light *wave-fronts. Elementary waves* are wave fronts emerging from a single point at a given instant. HUYGENS *theorem* in optics states that wave-fronts can be obtained as the envelopes of the elementary waves issuing, at one instant, from each point of a given wave-front (see e.g. [76],[6]). A detailed account of the variational approach to geometrical optics will be given in section 2.5. The translation of these concepts to mechanics is due to HAMILTON.

We owe to Jacobi the observation that this point of view provides an effective tool in determining the evolution of a mechanical system and the development of what is still considered the most powerful method of solution of dynamical problems [6]. The stationarity of the action integral is a problem of *calculus of variations* on a nonlinear manifold. A necessary and sufficient differential condition for a path to be a trajectory is provided by the next theorem and will be called the Euler's condition. The classical result of Euler deals with regular paths and fixed end points and is formulated in coordinates.

The new statement, introduced below, deals with the more general case of non-fixed end points and piecewise regular paths, so that stationarity is expressed in terms of differential and jump conditions. Moreover, the formulation is coordinate-free and relies on the notion of exterior derivative [36].

The author became recently aware of a 1938 paper by P. Weiss where non-fixed end points in the action principle were considered [190]. Our development was independently performed before Weiss' treatment was brought to our attention through the quotation in [50]. Weiss' approach deals with regular dynamics in finite dimensions and considers arbitrary test fields.

Theorem 2.1.1 (Euler's conditions) *A path* $\Gamma \in \mathrm{C}^1(\mathrm{Pat}(I)\,;\mathbb{M})$ *is a trajectory if and only if the tangent vector field* $\mathbf{v}_{\boldsymbol{\Gamma}} \in \mathrm{C}^1(\mathrm{Pat}(\boldsymbol{\Gamma})\,;\mathbb{T}\boldsymbol{\Gamma})$ *meets, in each element of a regularity patchwork* $\mathrm{Pat}(\boldsymbol{\Gamma})$*, the differential condition*

$$d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}} \cdot \mathbf{v}_{\boldsymbol{\varphi}} = 0\,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{Vtest}(\boldsymbol{\Gamma})\,,$$

and, at the singularity interfaces $\mathrm{If}(\boldsymbol{\Gamma})$*, the jump conditions*

$$[[\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}]] = 0\,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{Vtest}(\boldsymbol{\Gamma})\,.$$

Proof. By applying the extrusion formula in each element of the regularity partition we get

$$\partial_{\lambda=0} \int_{\boldsymbol{\varphi}_\lambda \circ \boldsymbol{\Gamma}} \boldsymbol{\omega}^1 - \oint_{\partial \boldsymbol{\Gamma}} \boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} = \int_{\mathrm{Pat}(\boldsymbol{\Gamma})} d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} - \int_{\mathrm{Sing}(\boldsymbol{\Gamma})} [[\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}]]\,,$$

so that the action principle writes

$$\int_{\mathrm{Pat}(\boldsymbol{\Gamma})} d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} = \int_{\mathrm{Sing}(\boldsymbol{\Gamma})} [[\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}]]\,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{Vtest}(\boldsymbol{\Gamma})\,.$$

Then, by the fundamental theorem of the calculus of variations, we get the result. Indeed, let us assume that the path $\boldsymbol{\Gamma}$ be parametrized by $s \in I$ and let $\mathbf{v}_{\mathbf{\Gamma}} \in \mathrm{C}^1(\mathbf{\Gamma}\,;\mathbb{T}\mathbf{\Gamma})$ be the velocity field along the path, so that:

$$\int_{\mathrm{PAT}(\mathbf{\Gamma})} d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} - \int_{\mathrm{SING}(\mathbf{\Gamma})} [[\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}]] = \int_{\mathrm{PAT}(I)} d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} \cdot \mathbf{v}_{\mathbf{\Gamma}}\, ds - \int_{\mathrm{SING}(\mathbf{\Gamma})} [[\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}]]\,.$$

If the differential and jump conditions are fulfilled, the action principle holds. Conversely, if $d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\mathbf{\Gamma}} \cdot \mathbf{v}_{\boldsymbol{\varphi}} \neq 0$ at a point inside an element $\mathcal{P} \in \mathrm{PAT}(\mathbf{\Gamma})$ of the regularity partition, by continuity of $d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\mathbf{\Gamma}} \cdot \mathbf{v}_{\boldsymbol{\varphi}}$, we could take $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{VTEST}(\mathbf{\Gamma})$ such that $d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} \cdot \mathbf{v}_{\mathbf{\Gamma}} > 0$ on an open segment $U \subset \mathbf{\Gamma}$ around that point and $d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} \cdot \mathbf{v}_{\mathbf{\Gamma}} = 0$ on $\mathcal{P}\backslash U$. Hence $\int_{\mathrm{PAT}(\mathbf{\Gamma})}(d\boldsymbol{\omega}^1) \cdot \mathbf{v}_{\boldsymbol{\varphi}} > 0$, contrary to the assumption. The vanishing of the jumps follows by a simple argument. ■

EULER's conditions show that the geometry of the trajectory is uniquely determinate if the exact two-form $d\boldsymbol{\omega}^1$ has a 1D kernel at each point. This is the basic assumption to ensure local existence and uniqueness of the trajectory through a point of the state-space.

The next proposition states that the action principle and the EULER's conditions are preserved if the state-space is changed into another one by a diffeomorphic transformation.

Proposition 2.1.1 (Invariance under a diffeomorphism) *If the manifolds* $\mathbb{M}$ *and* $\mathbb{N}$ *are related by a diffeomorphic tranformation* $\boldsymbol{\psi} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{N})$, *then the action principle and the related* EULER *condition for the trajectory* $\mathbf{\Gamma} \subset \mathbb{M}$*:*

$$\partial_{\lambda=0} \int_{\boldsymbol{\varphi}_\lambda(\mathbf{\Gamma})} \boldsymbol{\omega}^1 = \int_{\partial\mathbf{\Gamma}} \boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} \iff d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\mathbf{\Gamma}} \cdot \mathbf{v}_{\boldsymbol{\varphi}} = 0\,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{VTEST}(\mathbf{\Gamma})\,,$$

are identical to the action principle and the related EULER *condition for the trajectory* $\boldsymbol{\psi}(\mathbf{\Gamma}) \subset \mathbb{N}$*:*

$$\partial_{\lambda=0} \int_{(\boldsymbol{\psi}\circ\boldsymbol{\varphi}_\lambda)(\mathbf{\Gamma})} \boldsymbol{\psi}{\uparrow}\boldsymbol{\omega}^1 = \int_{\partial\boldsymbol{\psi}(\mathbf{\Gamma})} \boldsymbol{\psi}{\uparrow}\boldsymbol{\omega}^1 \cdot \boldsymbol{\psi}{\uparrow}\mathbf{v}_{\boldsymbol{\varphi}} \iff d(\boldsymbol{\psi}{\uparrow}\boldsymbol{\omega}^1) \cdot \boldsymbol{\psi}{\uparrow}\mathbf{v}_{\mathbf{\Gamma}} \cdot \boldsymbol{\psi}{\uparrow}\mathbf{v}_{\boldsymbol{\varphi}} = 0\,.$$

Proof. The equality of the integrals follows from the formula for the change of variables since $\boldsymbol{\psi}(\partial\mathbf{\Gamma}) = \partial\boldsymbol{\psi}(\mathbf{\Gamma})$. Moreover, by the naturality of the exterior derivative with respect to the push, we have that:

$$\begin{aligned} d(\boldsymbol{\psi}{\uparrow}\boldsymbol{\omega}^1) \cdot \boldsymbol{\psi}{\uparrow}\mathbf{v}_{\mathbf{\Gamma}} \cdot \boldsymbol{\psi}{\uparrow}\mathbf{v}_{\boldsymbol{\varphi}} &= \boldsymbol{\psi}{\uparrow}(d\boldsymbol{\omega}^1) \cdot \boldsymbol{\psi}{\uparrow}\mathbf{v}_{\mathbf{\Gamma}} \cdot \boldsymbol{\psi}{\uparrow}\mathbf{v}_{\boldsymbol{\varphi}} \\ &= \boldsymbol{\psi}{\uparrow}(d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\mathbf{\Gamma}} \cdot \mathbf{v}_{\boldsymbol{\varphi}})\,, \end{aligned}$$

and this proves the equivalence. Alternatively the proof could be carried out in terms of integrals by the formula for the change of domain of integration. ■

Remark 2.1.1 *The local conditions are necessary and sufficient for the fulfilment of the action principle under various boundary conditions. Indeed the equivalence*

$$\int_{\mathbf{\Gamma}} d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} = 0 \iff d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\mathbf{\Gamma}} \cdot \mathbf{v}_{\boldsymbol{\varphi}} = 0 \,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{V\small{TEST}}(\mathbf{\Gamma}) \,,$$

still holds when the space $\mathrm{V\small{TEST}}(\mathbf{\Gamma})$ *is substituted by any linear subspace which contains the space* $\mathrm{C}_0^\infty(\mathbf{\Gamma}\,; \mathrm{\small{TEST}}\mathbb{M}(\mathbf{\Gamma}))$ *of indefinitely differentiable test vector fields vanishing in a neighbourhood of the end points.*

However, the assumption that the field $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{V\small{TEST}}(\mathbf{\Gamma})$ *vanishes at each endpoint of* $\mathbf{\Gamma}$ *, usually made in stating the action principle on a manifold (see e.g. [36]), is unmotivatedly special, unsatisfactory from the epistemological point of view (see remark 2.2.3) and is not adequate to deal with singular points in the trajectory.*

Figure 2.1: Emmy Amalie Noether (1882 - 1935)

The next two results, which are direct consequences of Theorem 2.1.1, are original in their presentation.

Proposition 2.1.2 (A symmetry condition) *The differential condition fulfilled by a trajectory* $\mathbf{\Gamma} \subset \mathbb{M}$ *may be equivalently written as*

$$d_{\mathbf{v}_{\mathbf{\Gamma}}}(\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}) = d_{\mathbf{v}_{\boldsymbol{\varphi}}}(\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\mathbf{\Gamma}}) \,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{V\small{TEST}}(\mathbf{\Gamma}) \,,$$

where $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^0(\mathbb{M}\,; \mathbb{TM})$ *is an extension of the virtual velocity* $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{V\small{TEST}}(\mathbf{\Gamma})$ *and* $\mathbf{v}_{\mathbf{\Gamma}} \in \mathrm{C}^0(\mathbb{M}\,; \mathbb{TM})$ *is the extension of* $\mathbf{v}_{\mathbf{\Gamma}} \in \mathrm{C}^1(\mathrm{P\small{AT}}(\mathbf{\Gamma})\,; \mathbb{T}\mathbf{\Gamma})$ *performed by pushing it along the flow* $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{M})$ *generated by* $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^0(\mathbb{M}\,; \mathbb{TM})$ *.*

Proof. The result follows from Theorem 2.1.1 by a direct application of PALAIS formula:

$$d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} \cdot \mathbf{v}_{\boldsymbol{\Gamma}} = d_{\mathbf{v}_{\boldsymbol{\varphi}}} (\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}}) - d_{\mathbf{v}_{\boldsymbol{\Gamma}}} (\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}) - \boldsymbol{\omega}^1 \cdot [\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_{\boldsymbol{\Gamma}}] \,.$$

Indeed, by tensoriality of the exterior derivative, the r.h.s. is independent of the extensions of $\mathbf{v}_{\boldsymbol{\varphi}}$ and $\mathbf{v}_{\boldsymbol{\Gamma}}$ and the special extension of $\mathbf{v}_{\boldsymbol{\Gamma}}$ implies that $[\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_{\boldsymbol{\Gamma}}] = 0$. ■

As a simple corollary we infer that:

Proposition 2.1.3 (Abstract Noether's theorem) *If the functional $\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}}$ enjoys the stationarity property: $d_{\mathbf{v}_{\boldsymbol{\varphi}}} (\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}}) = 0$, then the functional $\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}$ is constant along the trajectory $\boldsymbol{\Gamma} \subset \mathbb{M}$.*

2.1.4 Abstract force forms

Let us consider the bundle $\mathbb{T}_{\boldsymbol{\Gamma}}\mathbb{M}$ which is the restriction of the tangent bundle $\mathbb{T}\mathbb{M}$ to the path $\boldsymbol{\Gamma}$ and a differential two-form $\boldsymbol{\alpha}^2$ on $\mathbb{T}_{\boldsymbol{\Gamma}}\mathbb{M}$, the *regular-force-form*, which provides an abstract description of a possibly non-potential system of forces acting along the trajectory.

The force-form $\boldsymbol{\alpha}^2$ is said to be *potential* if it is defined on a neighbourhood $U(\boldsymbol{\Gamma}) \subset \mathbb{M}$ of the path and there is exact.

This amounts to assume that there exists a differential one-form $\boldsymbol{\beta}^1 \in \mathrm{C}^1(\mathbb{M}\,; U(\boldsymbol{\Gamma}))$ such that $\boldsymbol{\alpha}^2 = d\boldsymbol{\beta}^1$, where d is the exterior differentiation.

We consider also a differential one-form $\boldsymbol{\alpha}^1$ on $\mathbb{T}_{\mathrm{SING}(\boldsymbol{\Gamma})}\mathbb{M}$, the *impulsive-force-form*, which provides an abstract description of an impulsive system of forces acting at singular points on the trajectory.

The expression of the *force-forms* for mechanical system are provided in section 2.2.8.

Definition 2.1.5 (Action principle) *A trajectory $\boldsymbol{\Gamma} \subset \mathbb{M}$ of the system is a piecewise regular path $\boldsymbol{\Gamma} \in \mathrm{C}^1(\mathrm{PAT}(I)\,; \mathbb{M})$ such that the action integral meets the variational condition:*

$$\partial_{\lambda=0} \int_{\boldsymbol{\varphi}_\lambda \circ \boldsymbol{\Gamma}} \boldsymbol{\omega}^1 = \int_{\partial \boldsymbol{\Gamma}} \boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} + \int_{\boldsymbol{\Gamma}} \boldsymbol{\alpha}^2 \cdot \mathbf{v}_{\boldsymbol{\varphi}} + \int_{\mathrm{SING}(\boldsymbol{\Gamma})} \boldsymbol{\alpha}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}} \,,$$

for all flows $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{M})$ with initial velocity $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{VTEST}(\boldsymbol{\Gamma})$.

Theorem 2.1.2 (Euler's conditions) *A path* $\mathbf{\Gamma} \subset \mathbb{M}$ *is a trajectory if and only if the tangent vector field* $\mathbf{v}_{\mathbf{\Gamma}} \in \mathrm{C}^1(\mathrm{PAT}(\mathbf{\Gamma})\,; \mathbb{T}\mathbf{\Gamma})$ *meets, in each element of a regularity partition* $\mathrm{PAT}(\mathbf{\Gamma})$ *, the differential condition*

$$(d\boldsymbol{\omega}^1 - \boldsymbol{\alpha}^2) \cdot \mathbf{v}_{\mathbf{\Gamma}} \cdot \mathbf{v}_{\boldsymbol{\varphi}} = 0\,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{VTEST}(\mathbf{\Gamma})\,.$$

and, at the singularity interfaces $\mathrm{SING}(\mathbf{\Gamma})$ *, the jump conditions*

$$[[\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}]] = \boldsymbol{\alpha}^1 \cdot \mathbf{v}_{\boldsymbol{\varphi}}\,, \quad \forall\, \mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{VTEST}(\mathbf{\Gamma})\,.$$

2.1.5 Continuum vs rigid-body dynamics

The abstract theory concerning the action principle may be applied to continuum mechanics by envisaging a suitable phase-space to describe motions.

A continuous body is identified with an open, connected, reference domain $\mathbb{B} \subset \mathbb{S}$ embedded in the euclidean space $\{\mathbb{S}\,, \mathbf{g}\}$.

A configuration $\boldsymbol{\chi} \in \mathrm{C}^1(\mathbb{B}; \mathbb{S})$ of a continuous body $\mathbb{B} \subset \mathbb{S}$ is an injective map with the property of being a diffeomorphic transformation onto its range. The *configuration-space* $\mathbb{C}$ is assumed to be a differentiable manifold modeled on a BANACH space.

The *velocity phase-space* is the tangent bundle $\mathbb{TC}$ and the *covelocity phase-space* is the cotangent bundle $\mathbb{T}^*\mathbb{C}$.

The *velocity-time state-space* is $\mathbb{TC} \times I$, is the cartesian product of the velocity-space $\mathbb{TC}$ and an open time interval I , and the *covelocity-time state-space* is $\mathbb{T}^*\mathbb{C} \times I$.

These two state-spaces are respectively adopted in the Lagrangian and the Hamiltonian descriptions of dynamics. Vectors tangent to the velocity-time state-space $\mathbb{TC} \times I$ are in the bundle $\mathbb{TTC} \times \mathbb{T}I$ whose elements are pairs $\{\mathbf{X}(\mathbf{v})\,, \Theta(t)\} \in \mathbb{T}_{\mathbf{v}}\mathbb{TC} \times \mathbb{T}_t I$.

Denoting by $\boldsymbol{\tau}_{\mathbb{C}} \in \mathrm{C}^1(\mathbb{TC}\,; \mathbb{C})$ the projector on the base manifold, the velocity of the configuration $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \in \mathbb{C}$, corresponding to a tangent vector $\mathbf{X}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{TC}$ is found by acting on it with the differential $T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \in BL(\mathbb{T}_{\mathbf{v}}\mathbb{TC}\,; \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C})$ of the projector, to get: $T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}$.

A section $\mathbf{X} \in \mathrm{C}^1(\mathbb{TC}\,; \mathbb{TTC})$ of $\boldsymbol{\pi}_{\mathbb{TC}} \in \mathrm{C}^1(\mathbb{TTC}\,; \mathbb{TC})$, is such that $\boldsymbol{\pi}_{\mathbb{TC}} \circ \mathbf{X} = \mathbf{id}_{\mathbb{TC}}$. The tangent map $T\boldsymbol{\tau}_{\mathbb{C}} \in \mathrm{C}^1(\mathbb{TTC}\,; \mathbb{TC})$, defined by $(T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{X})(\mathbf{v}) = T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v})$ maps each vector $\mathbf{X}(\mathbf{v})$ into the velocity of the configuration $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \in \mathbb{C}$.

2.1.6 Holonomic vs non-holonomic constraints

A dynamical system is said to be subject to *ideal constraints* if the admissibile velocities are imposed to belong to a vector sub-bundle $\mathcal{A}$ of $\mathbb{TC}$, that is, a bundle with base manifold $\mathbb{C}$ and fibers which are linear subspaces of the tangent spaces to $\mathbb{C}$.

The subbundle $\mathcal{A}$ is integrable if for any $\mathbf{x} \in \mathbb{C}$ there exists a (local) submanifold (the integral manifold) $\mathbb{I}_{\mathcal{A}} \subset \mathbb{C}$ thru $\mathbf{x}$ such that $\mathbb{TI}_{\mathcal{A}}$ is $\mathcal{A}$ restricted to $\mathbb{I}_{\mathcal{A}}$.

If the subbundle $\mathcal{A}$ is integrable, the ideal constraints are said *holonomic*.

FROBENIUS theorem states that integrability holds if and only if the subbundle $\mathcal{A}$ is involutive, that is for any pair of vector fields $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathbb{C}; \mathcal{A})$ in the vector sub-bundle $\mathcal{A}$ of $\mathbb{TC}$ we have that $[\mathbf{u}, \mathbf{v}] = \mathcal{L}_{\mathbf{u}} \mathbf{v} \in \mathrm{C}^1(\mathbb{C}; \mathcal{A})$.

2.1.7 Rigidity constraint

Two configurations $\boldsymbol{\chi}_1 \in \mathrm{C}^1(\mathbb{B}; \mathbb{S})$ and $\boldsymbol{\chi}_2 \in \mathrm{C}^1(\mathbb{B}; \mathbb{S})$ are metric-equivalent if $\boldsymbol{\varphi}_2 \downarrow \mathbf{g} = \boldsymbol{\varphi}_1 \downarrow \mathbf{g}$. Then the diffeomorphic map $\boldsymbol{\chi}_2 \circ \boldsymbol{\chi}_1^{-1} \in \mathrm{C}^1(\boldsymbol{\varphi}_1(\mathbb{B}); \boldsymbol{\varphi}_2(\mathbb{B}))$ is a metric-preserving (or rigid) transformation of the configuration $\boldsymbol{\chi}_1 \in \mathrm{C}^1(\mathbb{B}; \mathbb{S})$ into the configuration $\boldsymbol{\chi}_2 \in \mathrm{C}^1(\mathbb{B}; \mathbb{S})$.

By the metric-equivalence relation so introduced, the manifold $\mathbb{C}$ is partitioned into a family of disjoint connected rigidity-classes $\mathbb{C}_R$ which are submanifolds of $\mathbb{C}$.

The elements of the tangent space $\mathbb{T}_{\boldsymbol{\chi}} \mathbb{C}_R$ to a rigidity-class $\mathbb{C}_R$ at the configuration $\boldsymbol{\chi} \in \mathbb{C}_R$ are the *infinitesimal isometries* $\mathbf{v} \in$ VTEST, that is, the vector fields $\mathbf{v} \in \mathrm{C}^1(\boldsymbol{\chi}(\mathbb{B}); \mathbb{S})$ fulfilling the EULER-KILLING condition:

$$\mathcal{L}_{\mathbf{v}} \mathbf{g} = \mathbf{g} \circ (2 \operatorname{sym} \nabla \mathbf{v}) = 0 .$$

The LIE derivative of the metric tensor is defined by:

$$\mathcal{L}_{\mathbf{v}} \mathbf{g} := \partial_{\lambda=0} \, \boldsymbol{\varphi}_\lambda \downarrow \mathbf{g} ,$$

where $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\boldsymbol{\chi}(\mathbb{B}); \mathbb{S})$ is the flow generated by $\mathbf{v} = \partial_{\lambda=0} \, \boldsymbol{\varphi}_\lambda$ and $\boldsymbol{\varphi}_\lambda \downarrow \mathbf{g}$ is the pull back along $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\boldsymbol{\chi}(\mathbb{B}); \mathbb{S})$ of the metric tensor:

$$(\boldsymbol{\varphi}_\lambda \downarrow \mathbf{g})(\mathbf{a}, \mathbf{b}) = \mathbf{g}(T\boldsymbol{\varphi}_\lambda \cdot \mathbf{a}, T\boldsymbol{\varphi}_\lambda \cdot \mathbf{b}) , \quad \forall \, \mathbf{a}, \mathbf{b} \in \mathbb{T}_{\boldsymbol{\chi}(\mathbb{B})} \mathbb{S} .$$

2.1.8 Perfect dynamical systems

Let the phase-space $\mathbb{M}$ be the velocity-time state-space manifold $\mathbb{M} = \mathbb{TC} \times I$ and the trajectory be an arbitrary path $\boldsymbol{\Gamma} \in \mathrm{C}^1(I; \mathbb{TC} \times I)$. Then the test subbundle TESTM of the tangent bundle $\mathbb{TM} = \mathbb{TTC} \times \mathbb{T}I$ is made of isometric velocities, i.e. pairs $\{\mathbf{Y}(\mathbf{v}), \Theta(t)\} \in \mathbb{T}_{\{\mathbf{v},t\}}(\mathbb{TC} \times I) = \mathbb{T}_{\mathbf{v}}\mathbb{TC} \times \mathbb{T}_t I$ such that the velocity $T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \cdot \mathbf{Y}(\mathbf{v}) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}$ is an infinitesimal isometry of the configuration $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \in \mathbb{C}$ at time $t \in I$.

In rigid-body dynamics all velocity fields are infinitesimal isometries. As a consequence, in the EULER-LAGRANGE condition, the test fields are completely arbitrary and can be dropped.

Another context in which the test fields are completely arbitrary is that of elastodynamics since there the rigidity constraint is eliminated by introducing the stress tensor field as a LAGRANGE multiplier (see section **??**).

In either case, under the further assumption that the force system acting on the body admits a potential we have that a trajectory $\boldsymbol{\Gamma} \subset \mathbb{M}$ is a path fulfilling the property $d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}} = 0$, for any tangent vector field $\mathbf{v}_{\boldsymbol{\Gamma}} \in \mathrm{C}^1(\boldsymbol{\Gamma}; \mathbb{T}\boldsymbol{\Gamma})$. We shall refer to these contexts as *perfect dynamical systems.*

Remark 2.1.2 *Trajectories of perfect dynamical systems are also called curl-lines of the differential one-form* $\boldsymbol{\omega}^1 \in \mathrm{C}^1(\mathbb{M}; \mathbb{T}^*\mathbb{M})$ *[6]. Indeed, in a 3D riemannian manifold* $\{\mathbb{M}, \mathbf{g}\}$, *with the metric-induced volume form* $\boldsymbol{\mu}_{\mathbf{g}}$, *setting* $\boldsymbol{\omega}^1 = \mathbf{g}\mathbf{w}$, *we have that* $\boldsymbol{\mu}_{\mathbf{g}} \cdot \mathrm{rot}\, \mathbf{w} = d(\mathbf{g}\mathbf{w}) = d\boldsymbol{\omega}^1$. *Hence the* EULER *condition* $d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}} = \boldsymbol{\mu}_{\mathbf{g}} \cdot \mathrm{rot}\, \mathbf{w} \cdot \mathbf{v}_{\boldsymbol{\Gamma}} = 0$ *means that the vector* $\mathbf{v}_{\boldsymbol{\Gamma}}$, *tangent to the trajectory, is parallel to* $\mathrm{rot}\, \mathbf{w}$.

Remark 2.1.3 *In optics the action principle is* FERMAT*'s time-stationarity principle which characterizes light rays. By virtue of this analogy, trajectories are also called rays and the following definitions are given.*

- *A ray-segment is a regular segment* $\boldsymbol{\Gamma} \in \mathbb{M}$ *of a trajectory. Then*

$$d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}} = 0 \implies \int_{\boldsymbol{\Gamma}} d\boldsymbol{\omega}^1 = 0 .$$

- *A ray-sheet is an hypersurface* $\boldsymbol{\Sigma} \in \mathbb{M}$ *generated by the trajectories crossing a given curve in* $\mathbb{M}$. *Then*

$$d\boldsymbol{\omega}^1 \cdot \mathbf{v}_{\boldsymbol{\Gamma}} = 0 \implies \int_{\boldsymbol{\Sigma}} d\boldsymbol{\omega}^1 = 0 .$$

- *A ray-tube in* $\mathbb{M}$ *is a tube whose generating lines are rays of the system.*

2.1.9 Abstract integral invariant

The EULER-LAGRANGE stationarity condition and STOKES formula provide the following invariance result.

Theorem 2.1.3 (Integral invariants) *The integral of the action one-form* $\boldsymbol{\omega}^1$ *around any loop* $\mathbf{c}$ *, surrounding a given ray-tube in* $\mathbb{M}$ *, is invariant.*

Proof. Given two loops $\mathbf{c}_1$ and $\mathbf{c}_2$ surrounding a ray-tube in $\mathbb{M}$, let $\boldsymbol{\Sigma}$ be the portion of the tube surface such that $\partial\boldsymbol{\Sigma} = \mathbf{c}_2 - \mathbf{c}_1$.

Then by STOKES formula:

$$\oint_{\mathbf{c}_2} \boldsymbol{\omega}^1 - \oint_{\mathbf{c}_1} \boldsymbol{\omega}^1 = \int_{\partial\boldsymbol{\Sigma}} \boldsymbol{\omega}^1 = \int_{\boldsymbol{\Sigma}} d\boldsymbol{\omega}^1 = 0\,,$$

where the last integral vanishes since $\boldsymbol{\Sigma}$ is a ray-sheet (see remark 2.1.3). ■

The next proposition shows that the invariance of the integral of the action one-form is indeed equivalent to the EULER-LAGRANGE stationarity condition.

Theorem 2.1.4 (Inverse of the integral invariants theorem) *Let the integral of the action one-form* $\boldsymbol{\omega}^1$ *around any loop* $\mathbf{c}$*, surrounding any given flow-tube in* $\mathbb{M}$ *generated by the flow* $\mathbf{Fl}^{\mathbf{X}}_{\lambda} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *of a vector field* $\mathbf{X} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ *be invariant. Then the flow-lines are trajectories, i.e.:*

$$d\boldsymbol{\omega}^1 \cdot \mathbf{X} = 0\,.$$

Proof. The invariance of the integral of the action one-form may be written as

$$\partial_{\lambda=0} \oint_{\boldsymbol{\psi}_\lambda(\mathbf{c})} \boldsymbol{\omega}^1 = \int_{\mathbf{c}} \mathcal{L}_{\mathbf{X}}\, \boldsymbol{\omega}^1 = 0\,.$$

By the homotopy formula

$$\mathcal{L}_{\mathbf{X}}\, \boldsymbol{\omega}^1 = d(\boldsymbol{\omega}^1 \cdot \mathbf{X}) + d\boldsymbol{\omega}^1 \cdot \mathbf{X}\,,$$

and by STOKES formula, being $\partial\mathbf{c} = 0$, we have that

$$\partial_{\lambda=0} \oint_{\boldsymbol{\psi}_\lambda(\mathbf{c})} \boldsymbol{\omega}^1 = \int_{\mathbf{c}} d\boldsymbol{\omega}^1 \cdot \mathbf{X} + \int_{\partial\mathbf{c}} \boldsymbol{\omega}^1 \cdot \mathbf{X} = \int_{\mathbf{c}} d\boldsymbol{\omega}^1 \cdot \mathbf{X} = 0\,.$$

By the arbitrarity of the intensity of the vector field $\mathbf{X}$ the result follows. Indeed, if at a point $\mathbf{x}$ of the flow-tube, it were $d\boldsymbol{\omega}^1 \cdot \mathbf{X} \cdot \mathbf{Y} \neq 0$, with $\mathbf{Y}$ tangent to the loop, we could take the field $\mathbf{X}$ vanishing outside a neighbourhood $U(\mathbf{x})$ of that point, so that, by continuity:

$$\int_{\mathbf{c}} d\boldsymbol{\omega}^1 \cdot \mathbf{X} = \int_{\mathbf{c} \cap U(\mathbf{x})} d\boldsymbol{\omega}^1 \cdot \mathbf{X} \neq 0 \,,$$

contrary to the assumption. Hence $d\boldsymbol{\omega}^1 \cdot \mathbf{X} \cdot \mathbf{Y} = 0$ and, by the arbitrarity of the loop $\mathbf{c}$, we have that:

$$d\boldsymbol{\omega}^1 \cdot \mathbf{X} \cdot \mathbf{Y} = 0 \,, \quad \forall\, \mathbf{Y} \in \mathbb{T}_{\mathbf{x}}\mathbb{M} \iff d\boldsymbol{\omega}^1 \cdot \mathbf{X} = 0 \,,$$

at any point of the flow-tube. ■

2.2 Classical Dynamics

Let us now turn to general dynamics. In the *lagrangian description*, the phase-space is the *velocity phase-space*, that is, the tangent bundle $\mathbb{TC}$ to the configuration manifold. The state variables are then velocity vector field based at a placement in the configuration manifold.

The projector $\boldsymbol{\tau}_{\mathbb{C}} \in \mathrm{C}^1(\mathbb{TC}\,; \mathbb{C})$ maps the velocity phase-space onto the configuration space by associating each velocity $\mathbf{v} \in \mathbb{TC}$ with its base placement $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \in \mathbb{C}$. The *Lagrangian* of the system is a time-dependent functional $L_t \in \mathrm{C}^1(\mathbb{TC}\,; \Re)$ on the velocity phase-space.

The usual expression of the Lagrangian is $L_t = K_t + P_t \circ \boldsymbol{\tau}_{\mathbb{C}}$, $K_t \in \mathrm{C}^1(\mathbb{TC}\,; \Re)$ is the positive definite quadratic kinetic energy and $P_t \in \mathrm{C}^1(\mathbb{C}\,; \Re)$ is the force potential.

The *fiber-derivative* $d_{\mathrm{F}} L_t \in \mathrm{C}^1(\mathbb{TC}\,; \mathbb{T}^*\mathbb{C})$ of the Lagrangian is defined by

$$d_{\mathrm{F}} L_t(\mathbf{v}_{\mathbf{x}}) \cdot \mathbf{w}_{\mathbf{x}} := \partial_{\lambda=0}\, L_t(\mathbf{v}_{\mathbf{x}} + \lambda\, \mathbf{w}_{\mathbf{x}}) \,,$$

where $\mathbf{v}_{\mathbf{x}}, \mathbf{w}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}$ are tangent vectors. In the tangent bundle $\mathbb{TC}$ the fiber-derivative plays the role of the partial derivative with respect to the vectorial part of tangent vectors, due to the linearity of the tangent fiber.

No analogue of the partial derivative of a Lagrangian with respect to the base point of the vectorial argument is available in a nonlinear configuration manifold, unless a connection is defined (see section 2.3).

When $L_t(\mathbf{v}) = K_t(\mathbf{v}) + P_t(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}))$, the fiber-derivative of the Lagrangian is equal to the fiber-derivative of the kinetic energy and has the mechanical meaning of a kinetic momentum. Let I be a time interval and $\gamma \in \mathrm{C}^1(I\,;\mathbb{C})$ a *time-parametrized path* in the configuration manifold with image $\boldsymbol{\gamma} := \gamma(I)$ and velocity field $\mathbf{v} \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\boldsymbol{\Gamma})$ with $\boldsymbol{\Gamma} := \mathbb{T}\boldsymbol{\gamma}$ defined by

$$\mathbf{v}(\gamma(t)) := \partial_{\tau=t}\,\gamma(\tau) = \dot{\gamma}(t)\,.$$

The classical variational statement of the law of dynamics concerns the action integral defined by the equivalent expressions:

$$\int_{\boldsymbol{\gamma}} (L \circ \mathbf{v})\,\gamma{\uparrow}dt = \int_I \gamma{\downarrow}(L \circ \mathbf{v})\,dt = \int_I (L \circ \mathbf{v} \circ \gamma)\,dt\,,$$

where we have applied the invariance formula of integrals under the pull-back by a morphism and the rule for the pull-back of the product between a scalar function and a one-form:

$$\gamma{\downarrow}((L_t \circ \mathbf{v})\,\gamma{\uparrow}dt) = \gamma{\downarrow}(L_t \circ \mathbf{v})\,dt = (L_t \circ \mathbf{v} \circ \gamma)\,dt\,.$$

The push forward $\gamma{\uparrow}dt \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ of the one form $dt \in \mathrm{C}^1(I\,;\mathbb{T}^*I)$ is defined by the commutative diagram

$$\begin{array}{ccc} \mathbb{T}^*I & \xleftarrow{T^*\gamma} & \mathbb{T}^*\mathbb{C} \\ {\scriptstyle dt}\uparrow & & \uparrow{\scriptstyle \gamma\uparrow dt} \\ I & \xrightarrow{\gamma} & \mathbb{C} \end{array} \quad \Longleftrightarrow \quad T^*\gamma \circ \gamma{\uparrow}dt \circ \gamma = dt\,.$$

The synchronous (first) variation of the action integral along a flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ in the configuration space is the derivative of the integral performed along the flow-dragged path. This compels to evaluate the integrand on the flow-dragged path. To this end the velocity field $\mathbf{v} \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{T}\mathbb{C})$ is extended, in a synchronous fashion, by dragging it along the flow, according to the relation:

$$\mathbf{v}(\boldsymbol{\varphi}_\lambda(\gamma(t))) := T\boldsymbol{\varphi}_\lambda(\gamma(t)) \cdot \mathbf{v}(\gamma(t)) = T\boldsymbol{\varphi}_\lambda(\gamma(t)) \cdot \dot{\gamma}(t)\,,$$

that is:

$$\mathbf{v} = \boldsymbol{\varphi}_\lambda{\uparrow}\mathbf{v}\,.$$

The classical statement of HAMILTON's principle, in the dynamics of continuous bodies, is the following.

Proposition 2.2.1 (Classical Hamilton's principle) *A dynamical trajectory of a continuous mechanical system in the configuration manifold is a time-parametrized path* $\gamma \in \mathrm{C}^1(I\,;\mathbb{C})$ *fulfilling the stationarity condition*

$$\partial_{\lambda=0} \int_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\gamma})} (L_t \circ \mathbf{v})\,(\boldsymbol{\varphi}_\lambda \circ \gamma)\!\uparrow\! dt = \partial_{\lambda=0} \int_I (\boldsymbol{\varphi}_\lambda \circ \gamma)\!\downarrow\!(L \circ \mathbf{v})\,dt$$
$$= \partial_{\lambda=0} \int_I (L \circ \mathbf{v} \circ \boldsymbol{\varphi}_\lambda \circ \gamma)\,dt = 0\,,$$

for any flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ *whose velocity field* $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{TC})$ *is an infinitesimal isometry and vanishes at the boundary (end points) of* $\boldsymbol{\gamma}$ *.*

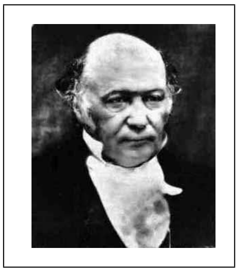

Figure 2.2: Sir William Rowan Hamilton (1805 - 1865)

Remark 2.2.1 *The kinetic energy in the Lagrangian functional* $L_t \in \mathrm{C}^1(\mathbb{TC}\,;\Re)$ *, is defined only on the trajectory of the body in the euclidean space, since the spatial mass-density is defined only there. On the other hand, to formulate* HAMILTON*'s principle, a definition of the Lagrangian on paths which are variations of the trajectory must be provided. In the literature on particle dynamics, this extension is tacitly performed by considering the point-mass of the particle to be constant along the virtual flow. Although such an assumption may appear as natural, when dealing with continuum dynamics the extension of the mass-form along virtual flows must be the object of an explicit statement (see Ansatz 3.6.1 on page 374).*

Remark 2.2.2 *In the literature* HAMILTON*'s principle is always stated in the context of perfect dynamics, is concerned with regular trajectories and the stationarity condition is imposed for any flow* $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ *, whose virtual velocity field* $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ *vanishes at the end points of the path [2], [6], [3].*

The basic step towards a general formulation of the law of dynamics consists in a suitable modification of the statement of Hamilton's principle, dropping out the assumption that the virtual velocity field vanishes at the end points of the path and allowing for singularities of the trajectory. The proper way to perform the modification follows from the discussion of the abstract action principle of section 2.1.2, when specialized to the velocity phase-space of lagrangian dynamics.

Remark 2.2.3 *The original definition of stationarity in the calculus of variations, and hence also of* HAMILTON*'s principle in dynamics, is unsatisfactory from the epistemological point of view. Indeed, it is a natural requirement that a property, characterizing a special class of paths, be formulated so that any piece of a special path is special too and the chain of two subsequent special paths is special too. The formulation of stationarity in terms of flows whose velocity vanishes at the end points of the path does not fulfill this natural requirement.*

2.2.1 The action one-form

We preliminarily recall some properties of flows in the tangent bundle of a manifold. Let $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ be a flow in the configuration manifold and $T\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$ the lifted flow induced, in the velocity phase-space, by the tangent functor, according to the relation $(T\boldsymbol{\varphi}_\lambda \cdot \mathbf{v})(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})) := T\boldsymbol{\varphi}_\lambda(\boldsymbol{\pi}(\mathbf{v})) \cdot \mathbf{v}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}))$ for all vector field, i.e. $\mathbf{v} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ with $T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{v} = \mathbf{id}_{\mathbb{C}}$.

We have the commutative diagram

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{T\boldsymbol{\varphi}_\lambda} & \mathbb{TC} \\ \boldsymbol{\pi}\downarrow & & \downarrow\boldsymbol{\pi} \\ \mathbb{C} & \xrightarrow{\boldsymbol{\varphi}_\lambda} & \mathbb{C} \end{array} \quad \Longleftrightarrow \quad \boldsymbol{\pi} \circ T\boldsymbol{\varphi}_\lambda = \boldsymbol{\varphi}_\lambda \circ \boldsymbol{\pi} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{C})\,.$$

The *canonical flip* $\mathbf{k}_{\mathbb{TTC}} \in \mathrm{C}^1(\mathbb{T}^2\mathbb{M}\,;\mathbb{T}^2\mathbb{M})$ defined by:

$$\mathbf{k}_{\mathbb{TTC}}\,(\partial_{\mu=0}\,\partial_{\lambda=0}\,\mathbf{c}(\lambda,\mu)) = \partial_{\lambda=0}\,\partial_{\mu=0}\,\mathbf{c}(\lambda,\mu)\,, \quad \forall\,\mathbf{c} \in \mathrm{C}^2(\Re \times \Re\,;\mathbb{C})\,,$$

is such that $\boldsymbol{\pi}_{\mathbb{TC}} \circ \mathbf{k}_{\mathbb{TTC}} = T\boldsymbol{\tau}_{\mathbb{C}}$, $T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{k}_{\mathbb{TTC}} = \boldsymbol{\pi}_{\mathbb{TC}}$ and $\mathbf{k}_{\mathbb{TTC}} \circ \mathbf{k}_{\mathbb{TTC}} = \mathbf{id}_{\mathbb{TTC}}$. The velocity of the lifted flow is then given by

$$\mathbf{v}_{T\boldsymbol{\varphi}} := \partial_{\lambda=0}\,T\boldsymbol{\varphi}_\lambda = \mathbf{k}_{\mathbb{C}} \circ T\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TTC})\,,$$

and from the relation $\boldsymbol{\pi}_{\mathbb{TC}} = T\boldsymbol{\pi} \circ \mathbf{k}_{\mathbb{C}}$ we have the commutative diagram

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{T\mathbf{v}_{\boldsymbol{\varphi}}} & \mathbb{TTC} \\ {\scriptstyle \mathbf{v}_{T\boldsymbol{\varphi}}}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\pi}_{\mathbb{TC}}} \\ \mathbb{TTC} & \xrightarrow{T\boldsymbol{\pi}} & \mathbb{TC} \end{array} \quad \Longleftrightarrow \quad \boldsymbol{\pi}_{\mathbb{TC}} \circ T\mathbf{v}_{\boldsymbol{\varphi}} = T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{v}_{T\boldsymbol{\varphi}} \,.$$

Moreover, the tangent functor fulfils the commutative diagram

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{T\mathbf{v}_{\boldsymbol{\varphi}}} & \mathbb{TTC} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{C}}}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\pi}_{\mathbb{TC}}} \\ \mathbb{C} & \xrightarrow{\mathbf{v}_{\boldsymbol{\varphi}}} & \mathbb{TC} \end{array} \quad \Longleftrightarrow \quad \boldsymbol{\pi}_{\mathbb{TC}} \circ T\mathbf{v}_{\boldsymbol{\varphi}} = \mathbf{v}_{\boldsymbol{\varphi}} \circ \boldsymbol{\tau}_{\mathbb{C}} \,,$$

and hence we get the commutative diagram

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{\mathbf{v}_{T\boldsymbol{\varphi}}} & \mathbb{TTC} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{C}}}\big\downarrow & & \big\downarrow{\scriptstyle T\boldsymbol{\tau}_{\mathbb{C}}} \\ \mathbb{C} & \xrightarrow{\mathbf{v}_{\boldsymbol{\varphi}}} & \mathbb{TC} \end{array} \quad \Longleftrightarrow \quad T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{v}_{T\boldsymbol{\varphi}} = \mathbf{v}_{\boldsymbol{\varphi}} \circ \boldsymbol{\tau}_{\mathbb{C}} \,.$$

HAMILTON's principle may be formulated in terms of the integral of an action one-form, by introducing a suitable space, the *velocity-time state-space*.

Let I be an open time interval and $J \subset \Re$ an open real interval.

- A *pseudo-time* is a stricly increasing scalar function $\theta \in \mathrm{C}^1(I\,;J)$ of the dynamical time $t \in I$ of classical mechanics: $t_2 > t_1 \implies \theta(t_2) > \theta(t_1)$.
- A path in the configuration manifold is a map $\gamma \in \mathrm{C}^1(I\,;\mathbb{C})$.

In the configuration-time manifold $\mathbb{C}\times I$ a path $\gamma_I := \gamma\times\theta^{-1} \in \mathrm{C}^1(J\,;\mathbb{C}\times I)$ is a product map with $\theta^{-1} \in \mathrm{C}^1(J\,;I)$ and $\gamma \in \mathrm{C}^1(I\,;\mathbb{C})$. Its image is denoted by $\boldsymbol{\gamma}_I := \gamma_I(J)$. In a *time-parametrized* path the pseudo-time is the identity map: $\theta = \mathbf{id}_I$.

- The *lifted path* in the velocity-time state-space is described by the tangent map $T\gamma_I = T\gamma \times T\theta^{-1} \in \mathrm{C}^1(\mathbb{T}J\,;\mathbb{TC} \times \mathbb{T}I)$.

We have the commutative diagram:

$$\begin{array}{ccc} \mathbb{T}I & \xrightarrow{T\gamma} & \mathbb{TC} \\ {\scriptstyle \boldsymbol{\pi}_I}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\tau}_{\mathbb{C}}} \\ I & \xrightarrow{\gamma} & \mathbb{C} \end{array} \quad \Longleftrightarrow \quad \boldsymbol{\tau}_{\mathbb{C}} \circ T\gamma = \gamma \circ \boldsymbol{\pi}_I \in \mathrm{C}^1(\mathbb{T}I\,;\mathbb{C}) \,.$$

A tangent space $\mathbb{T}_{\mathbf{e}_0}\mathbb{V}$ to a linear space $\mathbb{V}$ at the point $\mathbf{e}_0 \in \mathbb{V}$, is identified with the linear space $\mathbb{V}$ itself, by assuming, for any $\mathbf{e} \in \mathbb{V}$, the equivalences $\{\mathbf{e}_0\,,\mathbf{e}\} \simeq \{0\,,\mathbf{e}\} \simeq \mathbf{e}$.

By performing this identification for all tangent spaces $\mathbb{T}_{\mathbf{e}_0}\mathbb{V}$, the trivial tangent bundle $\mathbb{V} \times \mathbb{V}$ reduces to the linear space itself, i.e. $\mathbb{T}\mathbb{V} \simeq \mathbb{V}$. If the space $\mathbb{V}$ is the real line $\Re$ and $I \subseteq \Re$, we may set $\mathbb{T}I \simeq \Re$.

It is useful to introduce the cartesian projector $\mathrm{pr}_{\mathbb{TC}} \in \mathrm{C}^1(\mathbb{TC} \times I\,;\mathbb{TC})$, defined by

$$\mathrm{pr}_{\mathbb{TC}}(\mathbf{v},t) := \mathbf{v}\,, \quad \forall\, \mathbf{v} \in \mathbb{TC}\,, \quad t \in I\,.$$

The basic tool to define the *action one-form*, is Legendre transform.

- The *action functional* associated with the *Lagrangian* is defined by:

$$\mathrm{A}_t(\mathbf{v}) := \langle\, d_{\mathrm{F}} L_t(\mathbf{v}), \mathbf{v}\,\rangle\,, \quad \mathbf{v} \in \mathbb{TC}\,.$$

- The Hamiltonian $H_t \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\Re)$ is the functional Legendre-conjugate to the *Lagrangian* and the *energy* of the system $E_t \in \mathrm{C}^1(\mathbb{TC}\,;\Re)$ is defined by the relation: $E_t := H_t \circ d_{\mathrm{F}} L_t$, so that

$$L_t(\mathbf{v}) + E_t(\mathbf{v}) = \mathrm{A}_t(\mathbf{v}) = \langle\, d_{\mathrm{F}} L_t(\mathbf{v}), \mathbf{v}\,\rangle\,.$$

The Poincaré-Cartan *one-form* $\boldsymbol{\theta}_{L_t} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{T}^*\mathbb{TC})$ is defined by the identity:

$$(\boldsymbol{\theta}_{L_t} \cdot \mathbf{Y})(\mathbf{v}) := \langle\, d_{\mathrm{F}} L_t(\mathbf{v}), T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \cdot \mathbf{Y}(\mathbf{v})\,\rangle\,, \quad \forall\, \mathbf{Y}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{TC}\,.$$

Setting $\mathbf{X}(\mathbf{v}_t) = \dot{\mathbf{v}}_t = \partial_{\tau=t}\, \mathbf{v}_\tau$, from the relation

$$\mathbf{v}_t = \partial_{\tau=t}\, \boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau) = T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \cdot \partial_{\tau=t}\, \mathbf{v}_\tau = T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \cdot \dot{\mathbf{v}}_t\,,$$

we infer that $T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) = \mathbf{v}_t$, so that

$$(\boldsymbol{\theta}_{L_t} \cdot \mathbf{X})(\mathbf{v}_t) = \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t)\,\rangle = \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_t\,\rangle = \mathrm{A}_t(\mathbf{v}_t)\,.$$

Defining the pull-back

$$\boldsymbol{\theta}_L := \mathrm{pr}_{\mathbb{TC}}{\downarrow}\boldsymbol{\theta}_{L_t} \in \mathrm{C}^1(\mathbb{TC} \times I\,;\mathbb{T}^*(\mathbb{TC} \times I))\,,$$

and noting that $\mathrm{pr}_{\mathbb{TC}}{\uparrow}\{\mathbf{X}(\mathbf{v})\,,\Theta(t)\} = \mathbf{X}(\mathbf{v})$, we have that

$$(\boldsymbol{\theta}_L \cdot (\mathbf{X}\,,\Theta))(\mathbf{v},t) = (\boldsymbol{\theta}_{L_t} \cdot \mathbf{X})(\mathbf{v})\,,$$

for all $(\mathbf{X}\,,\Theta)(\mathbf{v},t) = (\mathbf{X}(\mathbf{v})\,,\Theta(t)) \in \mathbb{T}_{\mathbf{v}}\mathbb{TC} \times \mathbb{T}_t I$.

- The *lagrangian action one-form* $\boldsymbol{\omega}_L^1 \in \mathrm{C}^1(\mathbb{TC} \times I\,; \mathbb{T}^*(\mathbb{TC} \times I))$ is defined by
$$\boldsymbol{\omega}_L^1(\mathbf{v}, t) := (\boldsymbol{\theta}_L - E\,dt)(\mathbf{v}, t)\,,$$

where $E(\mathbf{v}, t) := E_t(\mathbf{v})$ and, with a little abuse of notation, $t(\mathbf{v}\,, t) = t$. Then, for a tangent vector $(\mathbf{Y}(\mathbf{v})\,, \Theta(t)) \in \mathbb{T}_{\mathbf{v}}\mathbb{TC} \times \mathbb{T}_t I$, we have:

$$\langle\, dt, (\mathbf{Y}(\mathbf{v})\,, \Theta(t))\,\rangle = \Theta(t) \quad \text{so that} \quad \langle\, dt, (\mathbf{Y}(\mathbf{v})\,, 1_t)\,\rangle = 1_t\,.$$

We recall hereafter some useful relations. Being

$$\langle\, E_t(\mathbf{v})dt, (\mathbf{X}(\mathbf{v})\,, \Theta(t))\,\rangle = E_t(\mathbf{v})\,\Theta(t)\,, \quad \mathbf{v} \in \mathbb{TC}\,, \quad t \in I\,,$$

we have that

$$\boldsymbol{\omega}_L^1(\mathbf{v}, t) \cdot (\dot{\mathbf{v}}_t\,, 1_t) = \mathrm{A}_t(\mathbf{v}) - E_t(\mathbf{v})\langle\, dt, (\dot{\mathbf{v}}_t\,, 1_t)\,\rangle = \mathrm{A}_t(\mathbf{v}) - E_t(\mathbf{v}) = \mathrm{L}_t(\mathbf{v})\,,$$

and also

$$\boldsymbol{\omega}_L^1(T\boldsymbol{\varphi}_\lambda(\mathbf{v}), t) \cdot (T\boldsymbol{\varphi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t\,, 1_t) = \mathrm{L}_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}))\,,$$

$$\begin{aligned}\boldsymbol{\omega}_L^1(\mathbf{v}, t) \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v})\,, 0) &= \langle\, d_{\mathrm{F}}\mathrm{L}_t(\mathbf{v}), T_{\mathbf{v}}\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v})\,\rangle \\ &= \langle\, d_{\mathrm{F}}\mathrm{L}_t(\mathbf{v}), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\pi}(\mathbf{v}))\,\rangle\,.\end{aligned}$$

Indeed, for any curve $\mathbf{v} \in \mathrm{C}^1(I\,; \mathbb{TC})$ in the velocity phase-space, setting

$$\dot{\mathbf{v}}_t := \partial_{\tau=t}\, \mathbf{v}_\tau = T_t\mathbf{v} \cdot 1_t\,,$$

and recalling the relation between the push and the tangent map:

$$(\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t) \circ \boldsymbol{\varphi}_\lambda = T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)\,,$$

we have that: $\partial_{\tau=t}\, T\boldsymbol{\varphi}_\lambda(\mathbf{v}_\tau) = T^2\boldsymbol{\varphi}_\lambda(\mathbf{v}_t) \cdot \dot{\mathbf{v}}_t = (T\boldsymbol{\varphi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t)_{T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)}$ and

$$\begin{aligned}T_{T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)}\boldsymbol{\tau}_{\mathbb{C}} \cdot \partial_{\tau=t}\, T\boldsymbol{\varphi}_\lambda(\mathbf{v}_\tau) &= \partial_{\tau=t}\, (\boldsymbol{\tau}_{\mathbb{C}} \circ T\boldsymbol{\varphi}_\lambda)(\mathbf{v}_\tau) \\ &= \partial_{\tau=t}\, (\boldsymbol{\varphi}_\lambda \circ \boldsymbol{\tau}_{\mathbb{C}})(\mathbf{v}_\tau) \\ &= T_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\boldsymbol{\varphi}_\lambda \cdot \partial_{\tau=t}\, \boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau) \\ &= T_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\boldsymbol{\varphi}_\lambda \cdot \mathbf{v}_t\,.\end{aligned}$$

2.2.2 Geometric Hamilton principle

The test flows for HAMILTON's principle in the configuration manifold are variations of the trajectory induced by flows $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C};\mathbb{C})$ with velocity field $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C};\mathbb{TC})$. When synchronous variations are considered, there is an induced flow $\boldsymbol{\xi}_\lambda := \mathrm{pr}_{\mathbb{C}}{\downarrow}\boldsymbol{\varphi}_\lambda \times \mathrm{pr}_I{\downarrow}\mathbf{id}_I \in \mathrm{C}^1(\mathbb{C}\times I;\mathbb{C}\times I)$ in the configuration-time state-space, so that

$$\boldsymbol{\xi}_\lambda(\mathbf{x},t) = (\mathrm{pr}_{\mathbb{C}}{\downarrow}\boldsymbol{\varphi}_\lambda \times \mathrm{pr}_I{\downarrow}\mathbf{id}_I)(\mathbf{x},t) = \{\boldsymbol{\varphi}_\lambda(\mathrm{pr}_{\mathbb{C}}(\mathbf{x},t))\,, \mathrm{pr}_I(\mathbf{x},t)\} = \{\boldsymbol{\varphi}_\lambda(\mathbf{x})\,, t\}\,.$$

By applying the tangent functor, the flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C};\mathbb{C})$ induces, in the *velocity* phase-space, a *lifted phase-flow* $T\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{TC};\mathbb{TC})$ with *phase-velocity* field

$$\mathbf{v}_{T\boldsymbol{\varphi}} = \partial_{\lambda=0}\,T\boldsymbol{\varphi}_\lambda = \mathbf{k}_{\mathbb{TTC}} \circ T\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^1(\mathbb{TC};\mathbb{TTC})\,,$$

where $\mathbf{k}_{\mathbb{TTC}} \in \mathrm{C}^1(\mathbb{TTC};\mathbb{TTC})$ is the canonical flip.

In the classical HAMILTON's principle synchronous variations are considered and the related action principle may be stated as follows.

Proposition 2.2.2 (Synchronous action principle) *The trajectory of a continuous dynamical system in the configuration manifold is a time-parametrized piecewise regular path* $\gamma \in \mathrm{C}^1(\mathrm{PAT}(I);\mathbb{C})$ *with velocity* $\mathbf{v} = \dot{\gamma}$ *fulfilling the stationarity condition*

$$\partial_{\lambda=0}\int_I L(T\boldsymbol{\varphi}_\lambda(\mathbf{v}))\,dt = \oint_{\partial I} \langle d_{\mathrm{F}}L(\mathbf{v}), \delta\mathbf{v}\rangle\,dt\,,$$

for any virtual flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C};\mathbb{C})$ *with* $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\boldsymbol{\gamma};\mathbb{TC})$ *such that* $\delta\mathbf{v} = \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}))$ *is an infinitesimal isometry at* $\boldsymbol{\gamma}$*. The trajectory in the velocity-time state-space fulfils the action principle expressed by the following stationarity condition for the one-form* $\boldsymbol{\omega}^1_L \in \mathrm{C}^1(\mathbb{TC}\times I;\mathbb{T}^*(\mathbb{TC}\times I))$*:*

$$\partial_{\lambda=0}\int_{T\boldsymbol{\xi}_\lambda(\boldsymbol{\Gamma}_I)} \boldsymbol{\omega}^1_L = \oint_{\partial\boldsymbol{\Gamma}_I} \boldsymbol{\omega}^1_L \cdot \{\mathbf{v}_{T\boldsymbol{\varphi}}\,, 0\}\,,$$

for any virtual flow $\boldsymbol{\xi}_\lambda := \mathrm{pr}_{\mathbb{C}}{\downarrow}\boldsymbol{\varphi}_\lambda \times \mathrm{pr}_I{\downarrow}\mathbf{id}_I \in \mathrm{C}^1(\mathbb{C}\times I;\mathbb{C}\times I)$*.*

Proof. From the commutative diagram

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{T\boldsymbol{\varphi}_\lambda} & \mathbb{TC} \\ \Big\downarrow{\boldsymbol{\tau}_{\mathbb{C}}} & & \Big\downarrow{\boldsymbol{\tau}_{\mathbb{C}}} \\ \mathbb{C} & \xrightarrow{\boldsymbol{\varphi}_\lambda} & \mathbb{C} \end{array} \iff \quad \boldsymbol{\tau}_{\mathbb{C}} \circ T\boldsymbol{\varphi}_\lambda = \boldsymbol{\varphi}_\lambda \circ \boldsymbol{\tau}_{\mathbb{C}} \in \mathrm{C}^1(\mathbb{TC};\mathbb{C})\,,$$

taking the derivative $\partial_{\lambda=0}$, we infer the relation:

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{\mathbf{v}_{T\boldsymbol{\varphi}}} & \mathbb{TTC} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{C}}}\big\downarrow & & \big\downarrow{\scriptstyle T\boldsymbol{\tau}_{\mathbb{C}}} \\ \mathbb{C} & \xrightarrow{\mathbf{v}_{\boldsymbol{\varphi}}} & \mathbb{TC} \end{array} \iff \quad T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{v}_{T\boldsymbol{\varphi}} = \mathbf{v}_{\boldsymbol{\varphi}} \circ \boldsymbol{\tau}_{\mathbb{C}} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})\,.$$

Then, being $\partial\mathbb{T}\boldsymbol{\gamma} = \partial T\gamma(I) = T\gamma(\partial I)$, we have that:

$$\oint_{\partial\boldsymbol{\Gamma}} \boldsymbol{\omega}^1_L \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}\,, 0) = \oint_{\partial I} \langle\, d_{\mathrm{F}} L(\mathbf{v}), \delta\mathbf{v}\,\rangle\, dt\,,$$

and, being $T\boldsymbol{\xi}_\lambda(\boldsymbol{\Gamma}) = (T\boldsymbol{\xi}_\lambda \circ T\gamma)(I) = T(\boldsymbol{\xi}_\lambda \circ \gamma)(I)$, we have that:

$$\int_{T\boldsymbol{\xi}_\lambda(\mathbb{T}\boldsymbol{\gamma})} \boldsymbol{\omega}^1_L = \int_I \boldsymbol{\omega}^1_L(T\boldsymbol{\varphi}_\lambda(\mathbf{v}), t) \cdot (T\boldsymbol{\varphi}_\lambda \!\uparrow\! \dot{\mathbf{v}}\,, 1)\, dt = \int_I L(T\boldsymbol{\varphi}_\lambda(\mathbf{v}))\, dt\,.$$

This proves the equivalence of the two formulations. ■

If the initial and final configurations are held fixed by the virtual flow, the boundary term vanishes being $\delta\mathbf{v} = 0$ at the end points of $\boldsymbol{\gamma}$. This assumption is usually made in literature to formulate HAMILTON's principle [176], [2], [6]. More in general, the vanishing of the boundary term is equivalent to assume that the virtual velocity fulfils a equiprojectivity condition at the end points of $\boldsymbol{\gamma}$, that is $\langle\, d_{\mathrm{F}} L(\mathbf{v}), \delta\mathbf{v}\,\rangle_b = \langle\, d_{\mathrm{F}} L(\mathbf{v}), \delta\mathbf{v}\,\rangle_a$ where $I = [a, b]$.

2.2.3 Asynchronous action principle

Asynchronous variations of the trajectory are expressed by considering, in addition to the flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ in the configuration space, a flow $\theta_\lambda \in \mathrm{C}^1(I\,;\Re)$ in the time domain, with $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ and $\Theta = \partial_{\lambda=0}\, \theta_\lambda \in \mathrm{C}^1(I\,;\mathbb{T}\Re)$.

So we have a product flow $\boldsymbol{\varphi}_\lambda \times \theta_\lambda \in \mathrm{C}^1(\mathbb{C}\times\Re\,;\mathbb{C}\times\Re)$ in the configuration-time state-space.

In the velocity-time state-space the lifted flow is $T\boldsymbol{\varphi}_\lambda \times T\theta_\lambda \in \mathrm{C}^1(\mathbb{TC}\times\mathbb{T}\Re\,;\mathbb{TC}\times\mathbb{T}\Re)$ with velocity $(\mathbf{v}_{T\boldsymbol{\varphi}}\,, v_\theta) \in \mathrm{C}^1(\mathbb{TC}\times\mathbb{T}\Re\,;\mathbb{TTC}\times\mathbb{TT}\Re)$.

Each path of the one-parameter family $\boldsymbol{\varphi}_\lambda \circ \gamma \in \mathrm{C}^1(I\,;\mathbb{C})$ generated by the action of the flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$, with $\boldsymbol{\varphi}_0 = \mathbf{id}_{\mathbb{C}}$, on the trajectory $\gamma \in \mathrm{C}^1(I\,;\mathbb{C})$ is parametrized by a *pseudo time* $\theta_\lambda \in \mathrm{C}^1(I\,;\Re)$, with $\theta_0 = \mathbf{id}_I$.

Accordingly, the action of the flow $\boldsymbol{\varphi}_\lambda \times \theta_\lambda \in \mathrm{C}^1(\mathbb{C} \times I\,;\mathbb{C} \times \Re)$ transforms the trajectory $\gamma_I = \gamma \times \mathbf{id}_I \in \mathrm{C}^1(I\,;\mathbb{C} \times I)$, into the trajectory $\boldsymbol{\xi}_\lambda \circ \gamma_I = (\boldsymbol{\varphi}_\lambda \times \theta_\lambda) \circ \gamma_I \in \mathrm{C}^1(\Re\,;\mathbb{C} \times \Re)$ according to the law:

$$(\gamma(t)\,, t) \to ((\boldsymbol{\varphi}_\lambda \circ \gamma)(t)\,, \theta_\lambda(t)) = ((\boldsymbol{\varphi}_\lambda \circ \gamma \circ \theta_\lambda^{-1})(\theta_\lambda(t))\,, \theta_\lambda(t))\,, \; \forall t \in I\,.$$

Setting:

$$\gamma_\lambda := \boldsymbol{\varphi}_\lambda \circ \gamma \circ \theta_\lambda^{-1}\,,$$

the virtual velocity along the flow is given by:

$$\partial_{\lambda=0}\, \gamma_\lambda = \partial_{\lambda=0}\, (\boldsymbol{\varphi}_\lambda \circ \gamma \circ \theta_\lambda^{-1}) = \mathbf{v}_{\boldsymbol{\varphi}} \circ \gamma - T\gamma \cdot \Theta \in \mathrm{C}^1(\Re\,;\mathbb{TC})\,,$$

where $\Theta = \partial_{\lambda=0}\, \theta_\lambda = -\partial_{\lambda=0}\, \theta_\lambda^{-1}$ is the pseudo-time dilation rate in the asynchronous variation. For subsequent developments it is crucial to observe that:

Lemma 2.2.1 *Given a flow* $\theta_\lambda \in \mathrm{C}^1(I\,;\Re)$ *with velocity* $\Theta = \partial_{\lambda=0}\, \theta_\lambda \in \mathrm{C}^1(I\,;\mathbb{T}I)$ *and the maps* $f \in \mathrm{C}^1(\Re\,;\Re)$ *and* $E \in \mathrm{C}^1(\mathbb{TC} \times \Re\,;\Re)$, *the following relations hold*

$$\begin{aligned}
&\partial_{\lambda=0}\, (\theta_\lambda \!\uparrow\! 1_t)_{\theta_\lambda(t)} = \partial_{\tau=t}\, \Theta(\tau)\,,\\
&\partial_{\lambda=0}\, f(\theta_\lambda(t)) = \partial_{\tau=t}\, f(\tau)\, \Theta(t)\,,\\
&\partial_{\lambda=0}\, E_{\theta_\lambda(t)}(\mathbf{v}_t)\, (\theta_\lambda \!\uparrow\! 1_t)_{\theta_\lambda(t)} = \partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\, \Theta(\tau)\,.
\end{aligned}$$

Proof. By a direct computation we get

$$\begin{aligned}
\partial_{\lambda=0}\, (\theta_\lambda \!\uparrow\! 1_t)_{\theta_\lambda(t)} &= \partial_{\lambda=0}\, T\theta_\lambda(t) \cdot 1_t = T\partial_{\lambda=0}\, \theta_\lambda(t) \cdot 1_t\\
&= T\Theta(t) \cdot 1_t = \dot{\Theta}(t) = \partial_{\tau=t}\, \Theta(\tau)\,,\\
\partial_{\lambda=0}\, f(\theta_\lambda(t)) &= \partial_{\tau=t}\, f(\tau)\, \partial_{\lambda=0}\, \theta_\lambda(t) = \partial_{\tau=t}\, f(\tau)\, \Theta(t)\,,
\end{aligned}$$

so that

$$\begin{aligned}
\partial_{\lambda=0}\, E_{\theta_\lambda(t)}(\mathbf{v}_t)\, (\theta_\lambda \!\uparrow\! 1_t)_{\theta_\lambda(t)} &= \partial_{\lambda=0}\, E_{\theta_\lambda(t)}(\mathbf{v}_t) + E_t(\mathbf{v}_t)\, \partial_{\lambda=0}\, (\theta_\lambda \!\uparrow\! 1_t)_{\theta_\lambda(t)}\\
&= \partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\, \Theta(t) + E_t(\mathbf{v}_t)\, \partial_{\tau=t}\, \Theta(\tau)\\
&= \partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\, \Theta(\tau)\,,
\end{aligned}$$

and the result is proven. ■

The *asynchronous action principle* (**A.A.P.**) for the dynamical trajectory is expressed by the following statement.

Proposition 2.2.3 (A.A.P.) *The trajectory in the velocity-time state-space* $\mathbb{TC} \times I$ *is a lifted path* $\Gamma = T\gamma \in \mathrm{C}^1(I\,;\mathbb{TC} \times I)$ *such that the differential one-form* $\boldsymbol{\omega}^1_L \in \mathrm{C}^1(\mathbb{TC} \times I\,;\mathbb{T}^*\mathbb{TC} \times \mathbb{T}^*I)$ *fulfils the stationarity condition:*

$$\partial_{\lambda=0} \int_{T\boldsymbol{\xi}_\lambda(\boldsymbol{\Gamma}_I)} \boldsymbol{\omega}^1_L = \oint_{\partial\boldsymbol{\Gamma}_I} \boldsymbol{\omega}^1_L \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}\,, \Theta)\,,$$

for any flow $\boldsymbol{\xi}_\lambda = \mathrm{pr}_{\mathbb{C}}\!\downarrow\!\boldsymbol{\varphi}_\lambda \times \mathrm{pr}_I\!\downarrow\!\theta_\lambda \in \mathrm{C}^1(\mathbb{C}\times I\,;\mathbb{C}\times\Re)$ *, with velocity fields* $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ *and* $\Theta = \partial_{\lambda=0}\,\theta_\lambda \in \mathrm{C}^1(I\,;\mathbb{T}\Re)$ *with* $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{TC})$ *infinitesimal isometry at* $\boldsymbol{\gamma}$ *.*

Proof. To provide an explicit expression to the **A.A.P.** variational condition, we recall the definition of the one-form $\boldsymbol{\omega}^1_L(\mathbf{v},t) := \boldsymbol{\theta}_L(\mathbf{v},t) - E(\mathbf{v},t)\,\mathrm{pr}_{\mathbb{T}I}\!\downarrow\!dt$. Defining the *energy one-form* $\boldsymbol{\eta} \in \mathrm{C}^1(\mathbb{TC} \times \mathbb{T}I\,;\mathbb{T}^*\mathbb{TC} \times \mathbb{T}^*I)$ by $\boldsymbol{\eta}(\mathbf{v}_t,t) := E_t(\mathbf{v}_t)\,\mathrm{pr}_{\mathbb{T}I}\!\downarrow\!dt$, we may write $\boldsymbol{\omega}^1_L := \boldsymbol{\theta}_L - \boldsymbol{\eta}$. Then, being

$$T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = (T\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{v}_{T\boldsymbol{\varphi}})(\mathbf{v}_t) = \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) = \delta\mathbf{v}_t\,,$$

and $\langle T_{\theta_\lambda(t)}\mathbf{id}_{\Re}, (\theta_\lambda\!\uparrow\!1_t)_{\theta_\lambda(t)}\rangle = (\theta_\lambda\!\uparrow\!1_t)_{\theta_\lambda(t)}$, we have that

$$\oint_{\partial\boldsymbol{\Gamma}_I} \boldsymbol{\omega}^1_L \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}\,, \Theta) = \oint_{\partial I} \langle d_{\mathrm{F}} L_t((\mathbf{v}_t)), \delta\mathbf{v}_t \rangle\,dt - \oint_{\partial I} E_t(\mathbf{v}_t)\,\Theta(t)\,dt\,.$$

On the other hand, being $\boldsymbol{\theta}_L := \mathrm{pr}_{\mathbb{TC}}\!\downarrow\!\boldsymbol{\theta}_{L_t}$ and $\mathrm{pr}_{\mathbb{TC}}\circ(T\boldsymbol{\varphi}_\lambda\times\theta_\lambda)\circ\gamma_I = T(\boldsymbol{\varphi}_\lambda\circ\gamma)$, we have that

$$\begin{aligned}\int_{T\boldsymbol{\xi}_\lambda(\boldsymbol{\Gamma}_I)} \boldsymbol{\omega}^1_L &= \int_{T\boldsymbol{\xi}_\lambda(\boldsymbol{\Gamma}_I)} \boldsymbol{\theta}_L - \boldsymbol{\eta} \\ &= \int_{T\boldsymbol{\varphi}_\lambda(\boldsymbol{\Gamma})} \boldsymbol{\theta}_{L_t} - \int_{T\boldsymbol{\xi}_\lambda(\boldsymbol{\Gamma}_I)} \boldsymbol{\eta}\,,\end{aligned}$$

so that

$$\begin{aligned}\int_{T\boldsymbol{\xi}_\lambda(\boldsymbol{\Gamma}_I)} \boldsymbol{\omega}^1_L = &\int_I \mathrm{A}_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\,dt \\ &- \int_I E_{\theta_\lambda(t)}(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\,(\theta_\lambda\!\uparrow\!1_t)_{\theta_\lambda(t)}\,dt\,.\end{aligned}$$

By LEIBNIZ rule, the derivative of the last integral may be split into

$$\partial_{\lambda=0} \int_I E_{\theta_\lambda(t)}(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\,(\theta_\lambda\!\uparrow\!1_t)_{\theta_\lambda(t)}\,dt$$

$$= \partial_{\lambda=0} \int_I E_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\,dt + \partial_{\lambda=0} \int_I E_{\theta_\lambda(t)}(\mathbf{v}_t)\,(\theta_\lambda\!\uparrow\!1_t)_{\theta_\lambda(t)}\,dt\,.$$

By Lemma 2.2.1 have that $\partial_{\lambda=0}\, E_{\theta_\lambda(t)}(\mathbf{v}_t)\,(\theta_\lambda\!\uparrow\!1_t)_{\theta_\lambda(t)} = \partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\,\Theta(\tau)\,.$ Then, being

$$\oint_{\partial I} E_t(\mathbf{v}_t)\,\Theta(t)\,dt = \int_I \partial_{\tau=t}\, E_\tau(\mathbf{v}_\tau)\,\Theta(\tau)\,dt$$

$$= \int_I \partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\,\Theta(\tau)\,dt + \int_I \partial_{\tau=t}\, E_t(\mathbf{v}_\tau)\,\Theta(t)\,dt\,,$$

the **A.A.P.** may be written as

$$\partial_{\lambda=0} \int_I \mathrm{A}_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\,dt - \oint_{\partial I} \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \delta\mathbf{v}_t \rangle\,dt$$

$$= \partial_{\lambda=0} \int_I E_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\,dt - \partial_{\tau=t} \int_I E_t(\mathbf{v}_\tau)\,\Theta(t)\,dt\,.$$

Being $L_t + E_t = \mathrm{A}_t$, the independency of time and velocity variations and the arbitrarity of time variations, the **A.A.P.** may be split into:

$$\begin{cases} \partial_{\lambda=0} \displaystyle\int_I L_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\,dt = \oint_{\partial I} \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \delta\mathbf{v}_t \rangle\,dt\,, \\ \partial_{\tau=t}\, E_t(\mathbf{v}_\tau) = 0\,. \end{cases}$$

The former is the new form of HAMILTON's principle introduced earlier in proposition 2.2.2, while the latter is the statement of conservation of energy for time dependent lagrangians: the energy functional is dragged by the motion.

This means that the convective LIE derivative of the energy functional along the trajectory, which is also the directional derivative along a tangent to the trajectory, vanishes at any time. The total time rate is then equal to the partial time derivative evaluated at fixed velocity. ■

We shall see that the conservation of energy is a consequence of HAMILTON's principle and hence the enlargment of the test flows to include also asynchronous flows in the velocity-time state-space is permitted since it does not impose further conditions to the motion.

HAMILTON's principle, which deals with the special case of a time-flow $\theta_\lambda \in \mathrm{C}^1(I\,;\Re)$ equal to the identity, will be called the *synchronous action principle* (**S.A.P.**).

2.2.4 Free asynchronous action principle

A more general variational condition, which will be called the *free asynchronous action principle* (**F.A.A.P.**), may be formulated by considering a path $\mathbf{\Gamma}_I \in \mathrm{C}^1(I\,;\mathbb{TC}\times I)$ in the velocity-time state-space, with cartesian projection $\mathbf{\Gamma} := \mathrm{pr}_{\mathbb{TC}} \circ \mathbf{\Gamma}_I \in \mathrm{C}^1(I\,;\mathbb{TC})$ on the velocity phase-space, defined by $(\mathbf{\Gamma}(t)\,,t) = \mathbf{\Gamma}_I(t)$. Virtual flows $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$, in the velocity phase-space are assumed to be *automorphic flows*. This means that for each $\lambda \in I$ the map $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$ is an automorphism, i.e. fiber-preserving and invertible, and thus projects to an invertible map $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ in the configuration space. By Lemma 1.2.11 on page 57, the virtual velocities $\mathbf{Y} = \partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TTC})$ are vector fields, i.e. $\boldsymbol{\tau}_{\mathbb{TC}} \circ \mathbf{Y} = \mathbf{id}_{\mathbb{TC}}$, which are decomposable as sum of two contributions:

$$\mathbf{Y} = \mathbf{V} + \mathbf{v}_{T\boldsymbol{\varphi}}\,,$$

where $\mathbf{V} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TTC})$ is a vertical vector field, i.e. $T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{V} = 0$.

Proposition 2.2.4 (Free asynchronous action principle) *A trajectory in the velocity-time state-space is a path* $\mathbf{\Gamma} \in \mathrm{C}^1(I\,;\mathbb{TC}\times I)$ *such that the one-form* $\boldsymbol{\omega}^1_L \in \mathrm{C}^1(\mathbb{TC}\times I\,;\mathbb{T}^*(\mathbb{TC}\times I))$ *fulfils the stationarity condition,*

$$\partial_{\lambda=0} \int_{(\mathbf{Fl}^{\mathbf{Y}}_\lambda \times \mathbf{Fl}^{\Theta}_\lambda)(\mathbf{\Gamma}_I)} \boldsymbol{\omega}^1_L = \oint_{\partial \mathbf{\Gamma}_I} \boldsymbol{\omega}^1_L \cdot (\mathbf{Y}\,,\Theta)\,,$$

for all virtual flows $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$ *and* $\mathbf{Fl}^{\Theta}_\lambda \in \mathrm{C}^1(I\,;I)$ *such that the flow* $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$ *projects to a flow* $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ *in the configuration space whose velocity* $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^1(\gamma(I)\,;\mathbb{TC})$ *is an infinitesimal isometry of the projected trajectory* $\gamma = \boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{\Gamma}$.

Proof. By theorem 2.1.1, the EULER differential condition of stationarity on the velocity $(\mathbf{X}(\mathbf{v}_t)\,,1_t) \in \mathbb{T}_{(\mathbf{v}_t\,,t)}\mathbf{\Gamma}_I$ at regular points along the trajectory is given by:

$$d\boldsymbol{\omega}^1_L(\mathbf{v}_t,t)\cdot(\mathbf{X}(\mathbf{v}_t)\,,1_t)\cdot(\mathbf{Y}(\mathbf{v}_t)\,,\Theta(t)) = 0\,,$$

and the jump condition at singular points of the trajectory is given by:

$$[[\boldsymbol{\omega}_L^1 \cdot (\mathbf{X}\,,1)]]_{(\mathbf{v}_t\,,t)} \cdot (\mathbf{Y}(\mathbf{v}_t)\,,\Theta(t)) = 0\,.$$

Recalling that $\boldsymbol{\omega}_L^1 = \boldsymbol{\theta}_L - \boldsymbol{\eta}$, we may write the differential condition as:

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) = d\boldsymbol{\eta}(\mathbf{v}_t, t) \cdot (\mathbf{X}(\mathbf{v}_t)\,, 1_t) \cdot (\mathbf{Y}(\mathbf{v}_t)\,, \Theta(t))\,.$$

The computation of the exterior derivative of the energy one-form by Palais formula requires the extension of the tangent vector $(\mathbf{X}(\mathbf{v}_t)\,, 1_t) \in \mathbb{T}_{(\mathbf{v}_t\,,t)}\boldsymbol{\Gamma}_I$ to a vector field $\dot{\mathcal{F}} \in \mathrm{C}^1(\mathbb{TC} \times I\,; \mathbb{TTC} \times \mathbb{T}I)$ by pushing it along the phase-flow $\mathbf{Fl}_\lambda^{\mathbf{Y}} \times \theta_\lambda \in \mathrm{C}^1(\mathbb{TC} \times \Re\,; \mathbb{TC} \times \Re)$, according to the relation:

$$\dot{\mathcal{F}}(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t), \theta_\lambda(t)) := (\mathbf{Fl}_\lambda^{\mathbf{Y}} \Uparrow \mathbf{X}(\mathbf{v}_t)\,, \theta_\lambda \Uparrow 1_t)_{(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t)\,,\theta_\lambda(t))}\,.$$

Palais formula tells us that

$$\begin{aligned} d\boldsymbol{\eta}(\mathbf{v}_t, t) \cdot (\mathbf{X}(\mathbf{v}_t)\,, 1_t) \cdot (\mathbf{Y}(\mathbf{v}_t)\,, \Theta(t)) &= d_{(\mathbf{X}(\mathbf{v}_t)\,,1_t)} \langle \boldsymbol{\eta}, (\mathbf{Y}\,, \Theta) \rangle \\ &\quad - d_{(\mathbf{Y}(\mathbf{v}_t)\,,\Theta(t))} \langle \boldsymbol{\eta}, \dot{\mathcal{F}} \rangle + \langle \boldsymbol{\eta}, \mathcal{L}_{(\mathbf{Y}\,,\Theta)} \dot{\mathcal{F}} \rangle (\mathbf{v}_t, t)\,. \end{aligned}$$

Since, by the chosen extension, the Lie derivative $\mathcal{L}_{(\mathbf{Y}\,,\Theta)}\dot{\mathcal{F}}$ vanishes, we may evaluate as follows:

$$\begin{aligned} d_{(\mathbf{X}(\mathbf{v}_t)\,,1_t)} \langle \boldsymbol{\eta}, (\mathbf{Y}\,, \Theta) \rangle &= \partial_{\tau=t} \langle \boldsymbol{\eta}(\mathbf{v}_\tau, \tau), (\mathbf{Y}(\mathbf{v}_\tau)\,, \Theta(\tau)) \rangle \\ &= \partial_{\tau=t} \langle E_\tau(\mathbf{v}_\tau)\, \mathbf{id}_{\mathbb{T}\Re}(\tau), \Theta(\tau) \rangle \\ &= \partial_{\tau=t}\, E_\tau(\mathbf{v}_\tau)\, \Theta(\tau) \\ &= \partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\, \Theta(\tau) + \partial_{\tau=t}\, E_t(\mathbf{v}_\tau)\, \Theta(t)\,, \end{aligned}$$

and, by Lemma 2.2.1:

$$\begin{aligned} d_{(\mathbf{Y}(\mathbf{v}_t)\,,\Theta(t))} \langle \boldsymbol{\eta}, \dot{\mathcal{F}} \rangle &= \partial_{\lambda=0}\, E_{\theta_\lambda(t)}(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t))\, \langle dt, \theta_\lambda \Uparrow 1_t \rangle \\ &= \partial_{\lambda=0}\, E_{\theta_\lambda(t)}(\mathbf{v}_t)\, \theta_\lambda \Uparrow 1_t + \partial_{\lambda=0}\, E_t(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t)) \\ &= \partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\, \Theta(\tau) + \partial_{\lambda=0}\, E_t(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t))\,. \end{aligned}$$

Summing up, the terms $\partial_{\tau=t}\, E_\tau(\mathbf{v}_t)\, \Theta(\tau)$ cancel one another, in agreement with the tensoriality of the exterior derivative, and, being

$$\partial_{\lambda=0}\, E_t(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t)) = dE_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\,,$$

we get:

$$d\boldsymbol{\eta}(\mathbf{v}_t, t) \cdot (\mathbf{X}(\mathbf{v}_t)\,, 1_t) \cdot (\mathbf{Y}(\mathbf{v}_t)\,, \Theta(t)) = \partial_{\tau=t}\, E_t(\mathbf{v}_\tau)\, \Theta(t) - dE_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\,.$$

The differential condition takes then the canonical expression:

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) = \partial_{\tau=t}\, E_t(\mathbf{v}_\tau)\, \Theta(t) - dE_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\,,$$

which, by the arbitrarity of $\Theta(t) \in \mathbb{T}_t I$, is equivalent to

$$\begin{cases} d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) = -dE_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\,, \\ \partial_{\tau=t}\, E_t(\mathbf{v}_\tau) = dE_t(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) = 0\,. \end{cases}$$

By the skew symmetry of $d\boldsymbol{\theta}_L$ we have that $d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) = 0$. The latter condition may then be dropped, being implied by the former one which is HAMILTON's equation in lagrangian form. ■

The conclusion of the previous proposition is in accordance with the analysis performed, in the context of perfect dynamical sistems, by G.A. DESCHAMPS ([36], section 7.7) and by ABRAHAM & MARSDEN ([2], Theorem 5.1.13), on the basis of a formal treatment which follows E. CARTAN original one [24].

In the next Lemma 2.2.2 we show that the fulfilment of the canonical equation implies that an integral curve of the field $\mathbf{X}$ is indeed a velocity curve for a trajectory in the configuration manifold. To this end let us resume for convenience the contents of Lemma 1.2.11. A virtual flow $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,; \mathbb{T}\mathbb{C})$ in the velocity phase-space is assumed to be fiber preserving and hence projects to a virtual flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,; \mathbb{C})$ on the configuration manifold, according to the commutative diagrams:

$$\begin{array}{ccccc} \mathbb{T}\mathbb{C} & \xrightarrow{\mathbf{Fl}^{\mathbf{Y}}_\lambda} & \mathbb{T}\mathbb{C} & \xrightarrow{\mathbf{Y}} & \mathbb{T}\mathbb{T}\mathbb{C} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{C}}}\downarrow & & {\scriptstyle \boldsymbol{\tau}_{\mathbb{C}}}\downarrow & & \downarrow{\scriptstyle T\boldsymbol{\tau}_{\mathbb{C}}} \\ \mathbb{C} & \xrightarrow{\boldsymbol{\varphi}_\lambda} & \mathbb{C} & \xrightarrow{\mathbf{v}_{\boldsymbol{\varphi}}} & \mathbb{T}\mathbb{C} \end{array} \iff \begin{cases} \boldsymbol{\varphi}_\lambda \circ \boldsymbol{\tau}_{\mathbb{C}} = \boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,; \mathbb{C})\,, \\ \mathbf{v}_{\boldsymbol{\varphi}} \circ \boldsymbol{\tau}_{\mathbb{C}} = T\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{Y} \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,; \mathbb{T}\mathbb{C})\,. \end{cases}$$

Since the map $T\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,; \mathbb{T}\mathbb{C})$ is an automorphism which also projects to $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,; \mathbb{C})$, the *correction flow* $\mathbf{Fl}^{\mathbf{V}}_\lambda \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,; \mathbb{T}\mathbb{C})$ defined by:

$$\mathbf{Fl}^{\mathbf{V}}_\lambda := \mathbf{Fl}^{\mathbf{Y}}_\lambda \circ (T\boldsymbol{\varphi}_\lambda)^{-1} \iff \mathbf{Fl}^{\mathbf{Y}}_\lambda = \mathbf{Fl}^{\mathbf{V}}_\lambda \circ T\boldsymbol{\varphi}_\lambda\,,$$

projects to the identity: $\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{Fl}^{\mathbf{V}}_\lambda = \mathbf{id}_{\mathbb{C}} \circ \boldsymbol{\tau}_{\mathbb{C}}$. The velocity of the flow $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,; \mathbb{T}\mathbb{C})$ is thus split into:

$$\begin{aligned} \mathbf{Y}(\mathbf{v}_t) &= \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{Y}}_\lambda(\mathbf{v}_t) = \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{V}}_\lambda(\mathbf{v}_t) + \partial_{\lambda=0}\, T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t) \\ &= \mathbf{V}(\mathbf{v}_t) + \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t)\,, \end{aligned}$$

with $T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)\cdot\mathbf{V}(\mathbf{v}_t)=0$, which means that $\mathbf{V}(\mathbf{v}_t)\in\mathbb{T}_{\mathbf{v}_t}\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\mathbb{C}\equiv\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\mathbb{C}$ is a vertical vector.

HAMILTON's canonical equation is accordingly split into:

$$\begin{cases} d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t)\cdot\mathbf{X}(\mathbf{v}_t)\cdot\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t)=-\langle dE_t(\mathbf{v}_t),\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t)\rangle\,,\\ d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t)\cdot\mathbf{X}(\mathbf{v}_t)\cdot\mathbf{V}(\mathbf{v}_t)=-\langle dE_t(\mathbf{v}_t),\mathbf{V}(\mathbf{v}_t)\rangle\,. \end{cases}$$

The following result, first given in [4], reveals the role of vertical virtual velocities. The proof we give is original.

Lemma 2.2.2 *If the linear map* $d^2_{\mathrm{F}}L_t(\mathbf{v})\in BL(\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}\,;\mathbb{T}^*_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C})$ *is invertible, the fulfillment of the variational condition*

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v})\cdot\mathbf{X}(\mathbf{v})\cdot\mathbf{V}(\mathbf{v})=-\langle dE_t(\mathbf{v}),\mathbf{V}(\mathbf{v})\rangle\,,$$

for any vertical vector $\mathbf{V}(\mathbf{v})\in\mathbb{T}_{\mathbf{v}}\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}\simeq\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\mathbb{C}$*, is equivalent to require that* $T_{\mathbf{v}}\boldsymbol{\tau}_{\mathbb{C}}\circ\mathbf{X}(\mathbf{v})=\mathbf{v}$ *i.e. that* $\mathbf{X}(\mathbf{v})$ *is second order along the lifted trajectory.*

Proof. Let us denote by $\mathcal{F}_{\mathbf{X}}(\mathbf{Fl}^{\mathbf{V}}_{\lambda}(\mathbf{v})):=(\mathbf{Fl}^{\mathbf{V}}_{\lambda}\!\uparrow\!\mathbf{X}\circ\mathbf{Fl}^{\mathbf{V}}_{\lambda})(\mathbf{v})$ the extension of a vector field $\mathbf{X}(\mathbf{v})\in\mathbb{T}_{\mathbf{v}}\mathbb{TC}$ performed by pushing it along the flow $\mathbf{Fl}^{\mathbf{V}}_{\lambda}\in\mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$. Then PALAIS' formula gives:

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v})\cdot\mathbf{X}(\mathbf{v})\cdot\mathbf{V}(\mathbf{v})=d_{\mathbf{X}(\mathbf{v})}(\boldsymbol{\theta}_{L_t}\cdot\mathbf{V})(\mathbf{v})-d_{\mathbf{V}(\mathbf{v})}(\boldsymbol{\theta}_{L_t}\cdot\mathcal{F}_{\mathbf{X}})(\mathbf{v})\,.$$

The first term on the r.h.s. vanishes since $\boldsymbol{\theta}_{L_t}$ is horizontal:

$$(\boldsymbol{\theta}_{L_t}\cdot\mathbf{V})(\mathbf{v})=\langle d_{\mathrm{F}}L_t(\mathbf{v}),T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})\cdot\mathbf{V}(\mathbf{v})\rangle=0\,.$$

Observing that $\boldsymbol{\tau}_{\mathbb{C}}\circ\mathbf{Fl}^{\mathbf{V}}_{\lambda}=\boldsymbol{\tau}_{\mathbb{C}}$, and recalling the definition of the canonical soldering form $\mathbf{J}:=\mathbf{Vl}_{\mathbb{TC}}\circ(\boldsymbol{\pi}_{\mathbb{TC}}\,,T\boldsymbol{\tau}_{\mathbb{C}})$, the second term evaluates to:

$$\begin{aligned} d_{\mathbf{V}(\mathbf{v})}(\boldsymbol{\theta}_{L_t}\cdot\mathcal{F}_{\mathbf{X}})(\mathbf{v}) &= \partial_{\lambda=0}\,\langle d_{\mathrm{F}}L_t(\mathbf{Fl}^{\mathbf{V}}_{\lambda}(\mathbf{v})),T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{Fl}^{\mathbf{V}}_{\lambda}(\mathbf{v}))\cdot T\mathbf{Fl}^{\mathbf{V}}_{\lambda}(\mathbf{v})\cdot\mathbf{X}(\mathbf{v})\rangle\\ &= \partial_{\lambda=0}\,\langle d_{\mathrm{F}}L_t(\mathbf{Fl}^{\mathbf{V}}_{\lambda}(\mathbf{v})),T(\boldsymbol{\tau}_{\mathbb{C}}\circ\mathbf{Fl}^{\mathbf{V}}_{\lambda})(\mathbf{v})\cdot\mathbf{X}(\mathbf{v})\rangle\\ &= \partial_{\lambda=0}\,\langle d_{\mathrm{F}}L_t(\mathbf{Fl}^{\mathbf{V}}_{\lambda}(\mathbf{v})),T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})\cdot\mathbf{X}(\mathbf{v})\rangle\\ &= \langle d^2_{\mathrm{F}}L_t(\mathbf{v})\cdot\mathbf{V}(\mathbf{v}),(\mathbf{Vl}_{\mathbb{TC}}\circ(\boldsymbol{\pi}_{\mathbb{TC}}\,,T\boldsymbol{\tau}_{\mathbb{C}}))(\mathbf{v})\cdot\mathbf{X}(\mathbf{v})\rangle\\ &= \langle d^2_{\mathrm{F}}L_t(\mathbf{v})\cdot\mathbf{J}(\mathbf{v})\cdot\mathbf{X}(\mathbf{v}),\mathbf{V}(\mathbf{v})\rangle\,, \end{aligned}$$

where the last equality holds by the symmetry of $d^2_{\mathrm{F}}L_t(\mathbf{v})\in BL(\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}^2\,;\Re)=BL(\mathbb{T}_{\mathbf{v}}\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}^2\,;\Re)$. Setting $\mathbf{V}(\mathbf{v})=\mathbf{J}(\mathbf{v})\cdot\mathbf{Z}(\mathbf{v})$, we get the equality

$$\langle d\boldsymbol{\theta}_{L_t}(\mathbf{v})\cdot\mathbf{X}(\mathbf{v}),\mathbf{J}(\mathbf{v})\cdot\mathbf{Z}(\mathbf{v})\rangle=-\langle d^2_{\mathrm{F}}L_t(\mathbf{v})\cdot\mathbf{J}(\mathbf{v})\cdot\mathbf{X}(\mathbf{v}),\mathbf{J}(\mathbf{v})\cdot\mathbf{Z}(\mathbf{v})\rangle\,,$$

for any $\mathbf{Z}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{TC}$, that is:

$$\mathbf{J}^*(\mathbf{v}) \cdot d\boldsymbol{\theta}_{L_t}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) = -\mathbf{J}^*(\mathbf{v}) \cdot d^2_{\mathrm{F}} L_t(\mathbf{v}) \cdot \mathbf{J}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) \,.$$

On the other hand, being $E_t = H_t \circ d_{\mathrm{F}} L_t$ and $d_{\mathrm{F}} H_t \circ d_{\mathrm{F}} L_t = \mathbf{id}_{\mathbb{TC}}$ and noting that $\langle dE_t(\mathbf{v}), \mathbf{V}(\mathbf{v}) \rangle = \langle d_{\mathrm{F}} E_t(\mathbf{v}), \mathbf{V}(\mathbf{v}) \rangle$, we have:

$$d_{\mathrm{F}} E_t(\mathbf{v}) = d_{\mathrm{F}} H_t(d_{\mathrm{F}} L_t(\mathbf{v})) \cdot d^2_{\mathrm{F}} L_t(\mathbf{v}) = d^2_{\mathrm{F}} L_t(\mathbf{v}) \cdot \mathbf{v} = d^2_{\mathrm{F}} L_t(\mathbf{v}) \cdot \mathbf{C}(\mathbf{v}) \,,$$

where $\mathbf{C} \in \mathrm{C}^1(\mathbb{TC}\,; \mathbb{VTC})$ is the LIOUVILLE vector field defined by $\mathbf{C}(\mathbf{v}) := \mathbf{vl}_{\mathbb{TC}}(\mathbf{v}) \cdot \mathbf{v}$. Then, the assumption in the statement writes:

$$\mathbf{J}^*(\mathbf{v}) \cdot d\boldsymbol{\theta}_{L_t}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) = -\mathbf{J}^*(\mathbf{v}) \cdot d_{\mathrm{F}} E_t(\mathbf{v}) = -\mathbf{J}^*(\mathbf{v}) \cdot d^2_{\mathrm{F}} L_t(\mathbf{v}) \cdot \mathbf{C}(\mathbf{v}) \,.$$

If the linear map $d^2_{\mathrm{F}} L_t(\mathbf{v}) \in BL\,(\mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}\,; \mathbb{T}^*_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C})$ is invertible, we infer that $\mathbf{J}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) = \mathbf{C}(\mathbf{v})$ which is equivalent to $T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) = \mathbf{v}$ by the injectivity of the vertical lift. ■

Recalling that

$$\mathrm{A}_t(\mathbf{v}_t) = \langle \boldsymbol{\theta}_{L_t}(\mathbf{v}_t), \mathbf{X}(\mathbf{v}_t) \rangle \,,$$

and the homotopy formula

$$\mathcal{L}_{\mathbf{X}} \boldsymbol{\theta}_{L_t} = d(\boldsymbol{\theta}_{L_t} \cdot \mathbf{X}) + d\boldsymbol{\theta}_{L_t} \cdot \mathbf{X} \,,$$

the equation of motion

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) = -dE_t(\mathbf{v}_t) \,,$$

may be written as

$$\mathcal{L}_{\mathbf{X}} \boldsymbol{\theta}_{L_t}(\mathbf{v}_t) = d\mathrm{A}_t(\mathbf{v}_t) - dE_t(\mathbf{v}_t) = dL_t(\mathbf{v}_t) \,.$$

2.2.5 Lagrange's equation

Setting $\mathbf{X}(\mathbf{v}_t) = \partial_{\tau=t}\, \mathbf{v}_\tau = \dot{\mathbf{v}}_t$, HAMILTON's canonical equation is equivalent to the variational condition

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \dot{\mathbf{v}}_t \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = -\langle dE_t(\mathbf{v}_t), \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) \rangle \,.$$

An explicit form is provided by the next result.

Theorem 2.2.1 (Law of dynamics) Hamilton*'s canonical equation for the trajectory is equivalent to the differential condition:*

$$\partial_{\tau=t}\,\langle\, d_{\mathrm{F}}L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau\,\rangle = \partial_{\lambda=0}\,L_t(T\boldsymbol{\varphi}_\lambda \cdot \mathbf{v}_t)\,,$$

and the jump conditions

$$\langle\,[[d_{\mathrm{F}}L_t(\mathbf{v}_t)]], \delta\mathbf{v}_t\,\rangle = 0\,,$$

for all flows $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ *such that* $\delta\mathbf{v}$ *is an infinitesimal isometry.*

Proof. Palais formula yields the expression:

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t)\cdot\dot{\mathbf{v}}_t\cdot\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = d_{\dot{\mathbf{v}}_t}(\boldsymbol{\theta}_{L_t}\cdot\mathbf{v}_{T\boldsymbol{\varphi}}) - d_{\mathbf{v}_{T\boldsymbol{\varphi}}}(\boldsymbol{\theta}_{L_t}\cdot\dot{\mathcal{F}}) + (\boldsymbol{\theta}_{L_t}\cdot\mathcal{L}_{\mathbf{v}_{T\boldsymbol{\varphi}}}\dot{\mathcal{F}})(\mathbf{v}_t)\,,$$

where $\dot{\mathcal{F}} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TTC})$ is the extension of the vector $\dot{\mathbf{v}}_t \in \mathbb{T}_{\mathbf{v}_t}\boldsymbol{\Gamma}$ performed by pushing it along the phase-flow $T\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$, that is:

$$\dot{\mathcal{F}}(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)) := (T\boldsymbol{\varphi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t)_{T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)}\,.$$

Then the Lie derivative $\mathcal{L}_{(\mathbf{Y},\Theta)}\dot{\mathcal{F}}$ vanishes. Evaluating the first term we get:

$$\begin{aligned} d_{\dot{\mathbf{v}}_t}(\boldsymbol{\theta}_{L_t}\cdot\mathbf{v}_{T\boldsymbol{\varphi}}) &= \partial_{\tau=t}\,\langle\, d_{\mathrm{F}}L_\tau(\mathbf{v}_\tau), T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau)\cdot\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_\tau)\,\rangle \\ &= \partial_{\tau=t}\,\langle\, d_{\mathrm{F}}L_\tau(\mathbf{v}_\tau), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau))\,\rangle = \partial_{\tau=t}\,\langle\, d_{\mathrm{F}}L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau\,\rangle\,. \end{aligned}$$

Moreover, being

$$\begin{aligned} T\boldsymbol{\tau}_{\mathbb{C}}(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t)\cdot T\boldsymbol{\varphi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t &= T\boldsymbol{\tau}_{\mathbb{C}}(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t)\cdot\partial_{\tau=t}\,T\boldsymbol{\varphi}_\lambda(\mathbf{v}_\tau) \\ &= \partial_{\tau=t}\,(\boldsymbol{\tau}_{\mathbb{C}}\circ T\boldsymbol{\varphi}_\lambda)(\mathbf{v}_\tau) = \partial_{\tau=t}\,(\boldsymbol{\varphi}_\lambda\circ\boldsymbol{\tau}_{\mathbb{C}})(\mathbf{v}_\tau) \\ &= T\boldsymbol{\varphi}_\lambda\cdot T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)\cdot\dot{\mathbf{v}}_t = T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t\,. \end{aligned}$$

and

$$\begin{aligned} (\boldsymbol{\theta}_L\cdot\dot{\mathcal{F}})(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t) &= \langle\, d_{\mathrm{F}}L_t(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t), T\boldsymbol{\tau}_{\mathbb{C}}(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t)\cdot T\boldsymbol{\varphi}_\lambda\!\uparrow\!\dot{\mathbf{v}}_t\,\rangle \\ &= \langle\, d_{\mathrm{F}}L_t(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t), T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t\,\rangle\,, \end{aligned}$$

evaluating the second term we get

$$\begin{aligned} d_{\mathbf{v}_{T\boldsymbol{\varphi}}}(\boldsymbol{\theta}_{L_t}\cdot\dot{\mathcal{F}}) &= \partial_{\lambda=0}\,(\boldsymbol{\theta}_{L_t}\cdot\dot{\mathcal{F}})(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t) \\ &= \partial_{\lambda=0}\,\mathrm{A}_t(T\boldsymbol{\varphi}_\lambda\cdot\mathbf{v}_t) = d\mathrm{A}_t(\mathbf{v}_t)\cdot\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t)\,. \end{aligned}$$

Summing up:

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \dot{\mathbf{v}}_t \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = \partial_{\tau=t} \langle d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau \rangle \\ - d\mathrm{A}_t(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) \,.$$

Being $\mathrm{A}_t = L_t + E_t$, the explicit form of HAMILTON's canonical equation is given by

$$\partial_{\tau=t} \langle d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau \rangle = dL_t(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) \,,$$

with $dL_t(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = \partial_{\lambda=0} \, L_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))$. ■

The law of dynamics states that the time-rate of increase of the virtual power of the kinetic momentum along the trajectory is equal to the rate of variation of the Lagrangian along any flow whose velocity at the actual configuration is an admissible infinitesimal isometry.

In the author's knowledge, the general law of dynamics in a non-linear configuration manifold contributed above, is not quoted in the literature. This law provides the most general formulation of the governing rules of dynamics in terms of the Lagrangian of the system.

Remark 2.2.4 *To evaluate the expression of the law of dynamics in the form derived above, it is compelling to assign the flows* $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ *at least in a neighborhood of* $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \in \boldsymbol{\gamma}$ *and not just the initial velocity* $\mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))$ *at the actual configuration* $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \in \boldsymbol{\gamma}$. *By tensoriality, the flows* $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ *leading to the same value of* $\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) \in \mathbb{T}_{\mathbf{v}_t}\mathbb{TC}$ *are equivalent. Anyway, by introducing a connection, we shall see that this expression of the law of dynamics is equivalent to one in which virtual flows enters in the analysis only through their virtual velocity, thus revealing that dynamical equilibrium depends only on the kinematical constraints pertaining to the body-placement under consideration.*

Remark 2.2.5 *In the variational expression of the law of dynamics, the test fields* $\mathbf{v}_{\boldsymbol{\varphi}} \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{TC})$ *are infinitesimal isometries at the trajectory* $\gamma \in \mathrm{C}^1(I\,;\mathbb{C})$. *This rigidity constraint has a basic physical meaning since it reveals that the dynamical equilibrium at a given configuration is independent of the material properties of the body. The evaluation of the equilibrium configuration requires in general to take into account the constitutive properties of the material and hence to get rid of the rigidity constraint. This task can be accomplished in complete generality by the method of* LAGRANGE *multipliers. In continuum mechanics, the* LAGRANGE *multipliers in duality with the rigidity constraints are called the stress fields in the body [165], (see section 3.6 on page 367 and 3.5.3 on page 359).*

Remark 2.2.6 *The law of dynamics can be directly deduced from* HAMILTON*'s principle in the extended form provided by proposition 2.2.2. Indeed, by applying the fundamental theorem of calculus, the principle may be rewritten as:*

$$\int_I \partial_{\lambda=0}\, L_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\, dt = \int_I \partial_{\tau=t}\, \langle\, d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau\, \rangle\, dt\,.$$

By the arbitrarity of the flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ *and the piecewise continuity of the integrands, we get the result.*

Remark 2.2.7 *The general expression of the law of dynamics implies, as a trivial corollary, a statement which extends to continuum dynamics* E. NOETHER*'s theorem as formulated in [129], [6], [106], [3]. Indeed from the law of dynamics we directly infer that*

$$\partial_{\lambda=0}\, L_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)) = 0 \iff \partial_{\tau=t}\, \langle\, d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau\, \rangle = 0\,,$$

while the extension of NOETHER*'s theorem consists in the weaker statement:*

$$L_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)) = L_t(\mathbf{v}_t) \implies \partial_{\tau=t}\, \langle\, d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau\, \rangle = 0\,.$$

2.2.6 The Legendrian functor

The LEGENDRE transform associated with a regular Lagrangian $L_t \in \mathrm{C}^1(\mathbb{TC}\,;\Re)$ induces the covariant *Legendrian functor* LEG between the categories of tangent and cotangent bundles over the base manifold $\mathbb{C}$.

The Legendrian functor, transforms a morphism $\mathbf{f} \in \mathrm{C}^k(\mathbb{TC}\,;\mathbb{TC})$ into a morphism $\mathrm{LEG}(\mathbf{f}) \in \mathrm{C}^k(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ defined by the commutative diagram:

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{\ \mathbf{f}\ } & \mathbb{TC} \\ {\scriptstyle d_{\mathrm{F}} L_t}\big\downarrow & & \big\downarrow{\scriptstyle d_{\mathrm{F}} L_t} \\ \mathbb{T}^*\mathbb{C} & \xrightarrow{\ \mathrm{LEG}(\mathbf{f})\ } & \mathbb{T}^*\mathbb{C} \end{array} \iff \quad \mathrm{LEG}(\mathbf{f}) \circ d_{\mathrm{F}} L_t := d_{\mathrm{F}} L_t \circ \mathbf{f}\,.$$

This means that the morphisms $\mathbf{f} \in \mathrm{C}^k(\mathbb{TC}\,;\mathbb{TC})$ and $\mathrm{LEG}(\mathbf{f}) \in \mathrm{C}^k(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ are $d_{\mathrm{F}} L_t$-related. If the Lagrangian is regular, we have that $d_{\mathrm{F}} H_t = (d_{\mathrm{F}} L_t)^{-1}$ and $T d_{\mathrm{F}} H_t = (T d_{\mathrm{F}} L_t)^{-1}$. The Legendrian functor is then invertible, according to the commutative diagram:

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{\ \mathrm{LEG}^{-1}(\mathbf{g})\ } & \mathbb{TC} \\ {\scriptstyle d_{\mathrm{F}} H_t}\big\uparrow & & \big\uparrow{\scriptstyle d_{\mathrm{F}} H_t} \\ \mathbb{T}^*\mathbb{C} & \xrightarrow{\ \mathbf{g}\ } & \mathbb{T}^*\mathbb{C} \end{array} \iff \quad \mathrm{LEG}^{-1}(\mathbf{g}) \circ d_{\mathrm{F}} H_t := d_{\mathrm{F}} H_t \circ \mathbf{g}\,.$$

Let $\varphi_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ be a flow with velocity field $\mathbf{v}_{\varphi} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ and $T\varphi_\lambda \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TC})$ the lifted flow with velocity field $\mathbf{v}_{T\varphi} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{TTC})$.

The flow $\mathrm{LEG}(T\varphi_\lambda) := d_{\mathrm{F}}L_t \circ T\varphi_\lambda \circ d_{\mathrm{F}}H_t \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$, is defined according to the commutative diagrams:

$$\begin{array}{ccccc} \mathbb{C} & \xleftarrow{\;\tau_{\mathbb{C}}\;} & \mathbb{TC} & \xleftarrow{\;d_{\mathrm{F}}H_t\;} & \mathbb{T}^*\mathbb{C} \\ \varphi_\lambda \uparrow & & T\varphi_\lambda \uparrow & & \mathrm{LEG}(T\varphi_\lambda) \uparrow \\ \mathbb{C} & \xleftarrow{\;\tau_{\mathbb{C}}\;} & \mathbb{TC} & \xleftarrow{\;d_{\mathrm{F}}H_t\;} & \mathbb{T}^*\mathbb{C} \end{array} \iff \begin{cases} \tau_{\mathbb{C}} \circ T\varphi_\lambda = \varphi_\lambda \circ \tau_{\mathbb{C}}\,, \\ d_{\mathrm{F}}H_t \circ \mathrm{LEG}(T\varphi_\lambda) = T\varphi_\lambda \circ d_{\mathrm{F}}H_t\,, \\ \tau^*_{\mathbb{C}} \circ \mathrm{LEG}(T\varphi_\lambda) = \varphi_\lambda \circ \tau^*_{\mathbb{C}}\,, \end{cases}$$

and its velocity $\mathbf{v}_{\mathrm{LEG}(T\varphi)} := Td_{\mathrm{F}}L_t \circ \mathbf{v}_{T\varphi} \circ d_{\mathrm{F}}H_t = d_{\mathrm{F}}L_t \!\uparrow\! \mathbf{v}_{T\varphi} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{TT}^*\mathbb{C})$ by the commutative diagrams:

$$\begin{array}{ccccc} \mathbb{TC} & \xleftarrow{\;T\tau_{\mathbb{C}}\;} & \mathbb{TTC} & \xleftarrow{\;Td_{\mathrm{F}}H_t\;} & \mathbb{TT}^*\mathbb{C} \\ \mathbf{v}_{\varphi} \uparrow & & \mathbf{v}_{T\varphi} \uparrow & & \mathbf{v}_{\mathrm{LEG}(T\varphi)} \uparrow \\ \mathbb{C} & \xleftarrow{\;\tau_{\mathbb{C}}\;} & \mathbb{TC} & \xleftarrow{\;d_{\mathrm{F}}H_t\;} & \mathbb{T}^*\mathbb{C} \end{array} \iff \begin{cases} T\tau_{\mathbb{C}} \circ \mathbf{v}_{T\varphi} = \mathbf{v}_{\varphi} \circ \tau_{\mathbb{C}}\,, \\ Td_{\mathrm{F}}H_t \circ \mathbf{v}_{\mathrm{LEG}(T\varphi)} = \mathbf{v}_{T\varphi} \circ d_{\mathrm{F}}H_t\,, \\ T\tau^*_{\mathbb{C}} \circ \mathbf{v}_{\mathrm{LEG}(T\varphi)} = \mathbf{v}_{\varphi} \circ \tau^*_{\mathbb{C}}\,. \end{cases}$$

2.2.7 Hamiltonian description

A general form of the action principle for a trajectory $\Gamma^*_I \in \mathrm{C}^1(I\,;\mathbb{T}^*\mathbb{C} \times I)$ in the covelocity-time state-space, is expressed by the variational condition:

$$\partial_{\lambda=0} \int_{((\mathbf{Fl}^{\mathbf{Y}}_\lambda \times \mathbf{Fl}^{\Theta}_\lambda)\circ \boldsymbol{\Gamma}^*)(I)} \boldsymbol{\omega}^1 = \oint_{\partial \boldsymbol{\Gamma}^*(I)} \boldsymbol{\omega}^1 \cdot (\mathbf{Y}\,,\Theta)\,,$$

for any time-flow $\mathbf{Fl}^{\Theta}_\lambda \in \mathrm{C}^1(I\,;I)$ with velocity vector field $\Theta \in \mathrm{C}^1(I\,;\mathbb{T}I)$ and any automorphic flow $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$, with projected flow $\varphi_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ defined by the commutative diagram:

$$\begin{array}{ccc} \mathbb{T}^*\mathbb{C} & \xrightarrow{\;\mathbf{Fl}^{\mathbf{Y}}_\lambda\;} & \mathbb{T}^*\mathbb{C} \\ \tau^*_{\mathbb{C}} \downarrow & & \tau^*_{\mathbb{C}} \downarrow \\ \mathbb{C} & \xrightarrow{\;\varphi_\lambda\;} & \mathbb{C} \end{array} \iff \tau^*_{\mathbb{C}} \circ \mathbf{Fl}^{\mathbf{Y}}_\lambda = \varphi_\lambda \circ \tau^*_{\mathbb{C}}\,.$$

We set $\mathbf{v}^*_t = \mathrm{pr}_{\mathbb{T}^*\mathbb{C}} \circ \Gamma^*_I(t)$, so that $\Gamma^*_I(t) = (\mathbf{v}^*_t\,, t)$. Localizing the action principle, the differential condition reads:

$$d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t) = -\langle dH_t(\mathbf{v}^*_t), \mathbf{Y}(\mathbf{v}^*_t) \rangle\,, \quad \forall\, \mathbf{Y}(\mathbf{v}^*_t) \in \mathbb{T}_{\mathbf{v}^*_t}\mathbb{T}^*\mathbb{C}\,.$$

The flows $\mathbf{Fl}^{\mathbf{Y}}_\lambda \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ and $\textsc{Leg}(T\boldsymbol{\varphi}_\lambda) := d_{\mathrm{F}}L_t \circ T\boldsymbol{\varphi}_\lambda \circ d_{\mathrm{F}}H_t \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ both project to the same base-flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$. If the map $T\boldsymbol{\varphi}_\lambda$ is invertible, then $\textsc{Leg}(T\boldsymbol{\varphi}_\lambda)$ is invertible too and we may define the map $\mathbf{Fl}^{\mathbf{V}}_\lambda := \mathbf{Fl}^{\mathbf{Y}}_\lambda \circ (\textsc{Leg}(T\boldsymbol{\varphi}_\lambda))^{-1} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ and write:

$$\mathbf{Fl}^{\mathbf{Y}}_\lambda = \mathbf{Fl}^{\mathbf{V}}_\lambda \circ \textsc{Leg}(T\boldsymbol{\varphi}_\lambda)\,,$$

with the flow $\mathbf{Fl}^{\mathbf{V}}_\lambda \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ such that $\boldsymbol{\tau}^*_{\mathbb{C}} \circ \mathbf{Fl}^{\mathbf{V}}_\lambda = \boldsymbol{\tau}^*_{\mathbb{C}}$.

Then, by Leibniz rule, we have the virtual velocity split:

$$\begin{aligned}\mathbf{Y}(\mathbf{v}^*_t) = \partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{Y}}_\lambda(\mathbf{v}^*_t) &= \partial_{\lambda=0}\,\mathbf{Fl}^{\mathbf{V}}_\lambda(\mathbf{v}^*_t) + \partial_{\lambda=0}\,\textsc{Leg}(T\boldsymbol{\varphi}_\lambda)(\mathbf{v}^*_t)\\ &= \mathbf{V}(\mathbf{v}^*_t) + \mathbf{v}_{\textsc{Leg}(T\boldsymbol{\varphi})}(\mathbf{v}^*_t)\,,\end{aligned}$$

and the verticality property:

$$\partial_{\lambda=0}\,(\boldsymbol{\tau}^*_{\mathbb{C}} \circ \mathbf{Fl}^{\mathbf{V}}_\lambda) = T\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)\cdot\mathbf{V}(\mathbf{v}^*_t) = 0\,,$$

so that $\mathbf{V}(\mathbf{v}^*_t) \in \mathbb{T}_{\mathbf{v}^*_t}\mathbb{T}^*_{\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)}\mathbb{C} \simeq \mathbb{T}^*_{\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)}\mathbb{C}$. The differential condition may thus be split into:

$$\begin{cases} d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*_t,t)\cdot\mathbf{X}(\mathbf{v}^*_t)\cdot\mathbf{v}_{\textsc{Leg}(T\boldsymbol{\varphi})}(\mathbf{v}^*_t) = -\langle\, dH_t(\mathbf{v}^*_t), \mathbf{v}_{\textsc{Leg}(T\boldsymbol{\varphi})}(\mathbf{v}^*_t)\,\rangle\,,\\ d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*_t,t)\cdot\mathbf{X}(\mathbf{v}^*_t)\cdot\mathbf{V}(\mathbf{v}^*_t) = -\langle\, d_{\mathrm{F}}H_t(\mathbf{v}^*_t), \mathbf{V}(\mathbf{v}^*_t)\,\rangle\,,\end{cases}$$

The second equations is fulfilled if and only if the velocity of the base trajectory associated with $\boldsymbol{\Gamma}^*$ is Legendre conjugate to the velocity $\mathbf{X}(\mathbf{v}^*_t)$ of $\boldsymbol{\Gamma}^*$, as is clarified by the next result.

Lemma 2.2.3 *The fulfillment of the differential condition*

$$d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*_t)\cdot\mathbf{X}(\mathbf{v}^*_t)\cdot\mathbf{V}(\mathbf{v}^*_t) = -\langle\, d_{\mathrm{F}}H_t(\mathbf{v}^*_t), \mathbf{V}(\mathbf{v}^*_t)\,\rangle\,,$$

for any vertical vector $\mathbf{V}(\mathbf{v}^*_t) \in \mathbb{T}_{\mathbf{v}^*_t}\mathbb{T}^*_{\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)}\mathbb{C} = \mathbb{T}^*_{\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)}\mathbb{C}$*, is equivalent to require that:*

$$T\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)\cdot\mathbf{X}(\mathbf{v}^*_t) = d_{\mathrm{F}}H_t(\mathbf{v}^*_t)\,.$$

Proof. By PALAIS formula with $\mathbf{X}(\mathbf{v}_t^*) \in \mathbb{T}_{\mathbf{v}_t^*}\mathbb{T}^*\mathbb{C}$ extended to a vector field $\dot{\mathcal{F}}^*(\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*)) := (\mathbf{Fl}_\lambda^{\mathbf{V}}\!\uparrow\!\mathbf{X})(\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*))$ pushed along the flow $\mathbf{Fl}_\lambda^{\mathbf{V}} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$:

$$d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)\cdot\mathbf{V}(\mathbf{v}_t^*) = d_{\mathbf{X}(\mathbf{v}_t^*)}(\boldsymbol{\theta}\mathbb{C}\cdot\mathbf{V})(\mathbf{v}_t^*) - d_{\mathbf{V}(\mathbf{v}_t^*)}(\boldsymbol{\theta}\mathbb{C}\cdot\dot{\mathcal{F}}^*)(\mathbf{v}_t^*)\,,$$

with $d_{\mathbf{X}(\mathbf{v}_t^*)}(\boldsymbol{\theta}\mathbb{C}\cdot\mathbf{V})(\mathbf{v}_t^*) = \partial_{\tau=t}\,\langle\mathbf{v}_\tau^*, T\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_\tau^*)\cdot\mathbf{V}(\mathbf{v}_\tau^*)\rangle = 0$ and

$$\begin{aligned}
d_{\mathbf{V}(\mathbf{v}_t^*)}(\boldsymbol{\theta}\mathbb{C}\cdot\dot{\mathcal{F}}^*)(\mathbf{v}_t^*) &= \partial_{\lambda=0}\,\langle\boldsymbol{\theta}\mathbb{C}(\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*)), (\mathbf{Fl}_\lambda^{\mathbf{V}}\!\uparrow\!\mathbf{X})(\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*))\rangle \\
&= \partial_{\lambda=0}\,\langle\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*), T\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*))\cdot T\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)\rangle \\
&= \partial_{\lambda=0}\,\langle\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*), T(\boldsymbol{\tau}_{\mathbb{C}}^*\circ\mathbf{Fl}_\lambda^{\mathbf{V}})(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)\rangle \\
&= \partial_{\lambda=0}\,\langle\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*), T\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)\rangle \\
&= \langle\partial_{\lambda=0}\,\mathbf{Fl}_\lambda^{\mathbf{V}}(\mathbf{v}_t^*), T\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)\rangle \\
&= \langle\mathbf{V}(\mathbf{v}_t^*), T\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)\rangle\,.
\end{aligned}$$

By the arbitrarity of $\mathbf{V}(\mathbf{v}_t^*) \in \mathbb{T}_{\mathbf{v}_t^*}\mathbb{T}^*_{\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_t^*)}\mathbb{C} = \mathbb{T}^*_{\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_t^*)}\mathbb{C}$, the differential condition may be written as $d_{\mathrm{F}}H_t(\mathbf{v}_t^*) = T\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)$. ■

Remark 2.2.8 *In the lecture notes by* F. GANTMACHER *[59] and in the nice book by* V.I. ARNOLD *[6], the action principle of dynamics is formulated in the covelocity-time state-space, that is, in the product space* $\mathbb{T}^*\mathbb{C}\times I$, *with the covelocity-phase-space* $\mathbb{T}^*\mathbb{C}$ *one-to-one related to the velocity state-space* $\mathbb{T}\mathbb{C}$ *by means of the* LEGENDRE *transform. The action principle is stated, in the special context of rigid-body dynamics and in finite dimensional configuration manifolds, as an extremality property of the integral of the one form* $\boldsymbol{\theta} - H\,dt$ *along the trajectory* $\boldsymbol{\Gamma}^*$ *in the covelocity-time state-space.*

The extremality property stated in [59, chap.3, sec.17], and in [6, chap.IX, sec. **C** *] considers arbitrary flows with the initial and the final configurations of the trajectory held fixed and it is claimed that the class of trajectory-variations in the covelocity-time state-space is greatly enlarged with respect to the ones considered in the usual statement of* HAMILTON*'s principle. In [6] this result is attributed to the extremality property of the* LEGENDRE *transformation under a convexity assumption on the Lagrangian. The analysis developed in the previous sections clarifies the situation. The class of trajectory-variations may be enlarged to include arbitrary flows in the covelocity-phase-space, which project to well-defined flows in the configuration manifold. This enlargement is exactly what is needed to get, as a natural condition of the variational action principle,* *the* LEGENDRE *transform between the momentum along the trajectory in the covelocity-phase-space and the velocity of the projected trajectory in the configuration manifold. The enlargement to asynchronous flows in the covelocity-time state-space is instead performed for free, due to the energy conservation law.*

2.2.8 Non-potential forces

When non-potential forces are considered acting on the mechanical system, the action principle and the relevant EULER conditions, must be suitably modified. The appropriate version of the action principle may be derived, from the abstract version stated in proposition 2.1.5, by defining the *force forms* as follows.

Non-potential forces acting on the mechanical system, are represented by a time-dependent field of one-forms $\mathbf{F}_t \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{T}^*\mathbb{C})$ on the trajectory in the configuration manifold, so that $\mathbf{F}_t(\mathbf{x}) \in \mathbb{T}^*_{\mathbf{x}}\mathbb{C}$ with $\mathbf{x} \in \boldsymbol{\gamma}$. To formulate the law of dynamics on the velocity-time state space we need first to express forces as one-forms on the velocity bundle. Physical consistency requires that force forms be represented by horizontal forms on the velocity bundle since the virtual work must vanish for a vanishing velocity of the base point in the configuration manifold. The correspondence between force one-forms $\mathbf{F}_t \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{T}^*\mathbb{C})$ acting along the trajectory in the configuration manifold, and horizontal one-forms $\mathbf{f}_t \in \mathrm{C}^1(\boldsymbol{\Gamma}\,;\mathbb{T}^*\mathbb{T}\mathbb{C})$ acting along the lifted trajectory in the velocity bundle is the bijection defined by: $\mathbf{f}_t := T^*\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{F}_t$, where $T^*\boldsymbol{\tau}_{\mathbb{C}} = (T\boldsymbol{\tau}_{\mathbb{C}})^*$, so that:

$$\begin{aligned}\mathbf{f}_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) &:= \langle T^*\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{F}_t(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)), \mathbf{Y}(\mathbf{v}_t)\rangle\,,\\ &= \langle \mathbf{F}_t(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)), T\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{Y}(\mathbf{v}_t)\rangle\,, \quad \forall\, \mathbf{Y}(\mathbf{v}_t) \in \mathbb{T}_{\mathbf{v}_t}\mathbb{T}^*\mathbb{C}\,,\end{aligned}$$

In the velocity-time state-space, at regular points of the trajectory, forces are represented by *force two-forms* defined by: $\boldsymbol{\alpha}^2_{\text{REG}}(\mathbf{v}_t, t) := -\mathbf{f}_t(\mathbf{v}_t) \wedge dt$. From the definition it follows that:

$$\begin{aligned}[\boldsymbol{\alpha}^2_{\text{REG}} \cdot (\mathbf{Y}\,, \Theta_t) \cdot (\mathbf{X}\,, 1)](\mathbf{v}_t, t) &= -(\mathbf{f}_t(\mathbf{v}_t) \wedge dt) \cdot (\mathbf{Y}(\mathbf{v}_t)\,, \Theta_t) \cdot (\mathbf{X}(\mathbf{v}_t)\,, 1_t)\\ &= (\mathbf{f}_t(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t))\,\Theta_t - \mathbf{f}_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\,,\end{aligned}$$

and, for synchronous virtual velocities:

$$[\boldsymbol{\alpha}^2_{\text{REG}} \cdot (\mathbf{Y}\,, 0) \cdot (\mathbf{X}\,, 1)](\mathbf{v}_t, t) = -\mathbf{f}_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\,.$$

Impulsive forces at singular points $\mathbf{x} \in \Gamma$ are described by one-forms $\mathbf{A}_t(\mathbf{x}) \in \mathbb{T}^*_{\mathbf{x}}\mathbb{C}$ and, on the lifted trajectory in the tangent bundle, by horizontal one-forms

$\boldsymbol{\alpha}^1_{\text{SING}} \in \mathbb{T}^*\mathbb{TC}$ defined by $\boldsymbol{\alpha}^1_{\text{SING}} = T^*\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{A}_t$, that is:

$$\boldsymbol{\alpha}^1_{\text{SING}}(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) = \langle \mathbf{A}_t(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)), T\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{Y}(\mathbf{v}_t) \rangle \,.$$

In the Hamiltonian description, the *force two-form* in the covelocity bundle is defined as

$$\boldsymbol{\alpha}^2_{\text{REG}}(\mathbf{v}^*, t) := -(\mathbf{f} \wedge dt)(\mathbf{v}^*, t)\,, \quad (\mathbf{v}^*\,, t) \in \gamma^*\,,$$

where $\mathbf{f}(\mathbf{v}^*, t) = \mathbf{f}_t(\mathbf{v}^*)$ and, with abuse of notation, $t(\mathbf{v}^*, t) = t$. Given a field of force one-forms $\mathbf{F}_t \in \mathrm{C}^1(\mathbb{C}\,; \mathbb{T}^*\mathbb{C})$ on the configuration manifold, the induced field of force one-forms on the covelocity bundle is defined by

$$\mathbf{f}_t := T^*\boldsymbol{\tau}^*_{\mathbb{C}} \cdot (\mathbf{F}_t \circ \boldsymbol{\tau}^*_{\mathbb{C}}) = \boldsymbol{\theta}\mathbb{C} \cdot (\mathbf{F}_t \circ \boldsymbol{\tau}^*_{\mathbb{C}}) \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,; \mathbb{T}^*\mathbb{T}^*\mathbb{C})\,,$$

where $\boldsymbol{\theta}\mathbb{C} = T^*\boldsymbol{\tau}^*_{\mathbb{C}} = (T\boldsymbol{\tau}^*_{\mathbb{C}})^* \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,; \mathbb{T}^*\mathbb{T}^*\mathbb{C})$. Then

$$\langle \mathbf{f}_t, \mathbf{Y} \rangle := \langle \boldsymbol{\theta}\mathbb{C} \cdot (\mathbf{F}_t \circ \boldsymbol{\tau}^*_{\mathbb{C}}), \mathbf{Y} \rangle = \langle \mathbf{F}_t \circ \boldsymbol{\tau}^*_{\mathbb{C}}, T\boldsymbol{\tau}^*_{\mathbb{C}} \cdot \mathbf{Y} \rangle \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,; \Re)\,,$$

or explicitly

$$\begin{aligned} \mathbf{f}_t(\mathbf{v}^*) \cdot \mathbf{Y}(\mathbf{v}^*) &:= \langle \boldsymbol{\theta}\mathbb{C} \cdot \mathbf{F}_t(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*)), \mathbf{Y}(\mathbf{v}^*) \rangle\,, \\ &= \langle \mathbf{F}_t(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*)), T\boldsymbol{\tau}^*_{\mathbb{C}} \cdot \mathbf{Y}(\mathbf{v}^*) \rangle\,, \quad \forall\, \mathbf{Y}(\mathbf{v}^*) \in \mathbb{T}_{\mathbf{v}^*}\mathbb{T}^*\mathbb{C}\,. \end{aligned}$$

Impulsive forces at singular points are one-forms $\boldsymbol{\alpha}^1_{\text{SING}} \in \mathbb{T}^*\mathbb{T}^*\mathbb{C}$ defined by

$$\boldsymbol{\alpha}^1_{\text{SING}} \cdot \mathbf{Y} := \langle \mathbf{A}_t \circ \boldsymbol{\tau}^*_{\mathbb{C}}, T\boldsymbol{\tau}^*_{\mathbb{C}} \cdot \mathbf{Y} \rangle \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,; \Re)\,,$$

where $\mathbf{A}_t(\mathbf{x}) \in \mathbb{T}^*_{\mathbf{x}}\mathbb{C}$.

2.2.9 Action principle in the covelocity space

Trajectories in the velocity-time state-space and in the covelocity-time state-space are related by: $\mathrm{pr}_{\mathbb{T}^*\mathbb{C}} \circ \boldsymbol{\Gamma}^* := d_{\text{F}}L \circ \mathrm{pr}_{\mathbb{TC}} \circ \boldsymbol{\Gamma}$.

Definition 2.2.1 *The* ***free asynchronous action principle*** *for the trajectory* $\boldsymbol{\Gamma}^* \in \mathrm{C}^1(I\,; \mathbb{T}^*\mathbb{C} \times I)$ *in the covelocity-time state-space, is expressed by the stationarity condition:*

$$\begin{aligned} \partial_{\lambda=0} \int_{(\mathbf{Fl}^{\mathbf{Y}}_\lambda \times \mathbf{Fl}^{\Theta}_\lambda)(\boldsymbol{\Gamma}^*_I)} \boldsymbol{\omega}^1 &= \int_{\partial\boldsymbol{\Gamma}^*_I} \boldsymbol{\omega}^1 \cdot (\mathbf{Y}\,, \Theta) \\ &+ \int_{\boldsymbol{\Gamma}^*_I} \boldsymbol{\alpha}^2_{\text{REG}} \cdot (\mathbf{Y}\,, \Theta) + \int_{\text{SING}(\boldsymbol{\Gamma}^*_I)} \boldsymbol{\alpha}^1_{\text{SING}} \cdot (\mathbf{Y}\,, \Theta)\,. \end{aligned}$$

If the trajectory in the covelocity-time state-space is parametrized with time, we have that

$$\begin{aligned}\boldsymbol{\alpha}^2_{\text{REG}}(\mathbf{v}^*_t, t) \cdot (\mathbf{X}(\mathbf{v}^*_t)\,, 1_t) \cdot (\mathbf{Y}(\mathbf{v}^*_t)\,, \Theta(t)) = &- \mathbf{f}_t(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t)\,\Theta(t) \\ &+ \mathbf{f}_t(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t)\,.\end{aligned}$$

The Euler-Lagrange differential condition of stationarity

$$(d\boldsymbol{\omega}^1 - \boldsymbol{\alpha}^2_{\text{REG}})(\mathbf{v}^*_t, t) \cdot (\mathbf{X}(\mathbf{v}^*_t)\,, 1_t) \cdot (\mathbf{Y}(\mathbf{v}^*_t)\,, \Theta(t)) = 0\,,$$

then becomes

$$\begin{aligned}d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t) = &\, d_{\mathbf{X}(\mathbf{v}^*_t)} H_t(\mathbf{v}^*_t)\,\Theta(t) - d_{\mathbf{Y}(\mathbf{v}^*_t)} H_t(\mathbf{v}^*_t) \\ &- \mathbf{f}_t(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t)\,\Theta(t) + \mathbf{f}_t(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t)\,,\end{aligned}$$

and, splitting, we get

$$\begin{cases} d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t) = (\mathbf{f}_t(\mathbf{v}^*_t) - dH_t(\mathbf{v}^*_t)) \cdot \mathbf{Y}(\mathbf{v}^*_t)\,, \\ d_{\mathbf{X}(\mathbf{v}^*_t)} H_t(\mathbf{v}^*_t) = \mathbf{f}_t(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t)\,, \end{cases}$$

for all $\mathbf{Y}(\mathbf{v}^*_t) \in \mathbb{T}_{\mathbf{v}^*_t}\mathbb{T}^*\mathbb{C}$ such that $\operatorname{sym}\nabla(T\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t)) = 0$.

The former is the general form of Hamilton's canonical law of dynamics while the latter, which expresses the energy conservation law, is a consequence of the former and can be dropped.

Being $T\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t) = d_{\text{F}} H_t(\mathbf{v}^*_t)$, we have that

$$\mathbf{f}_t(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t) = \langle \mathbf{F}_t(\mathbf{v}^*_t), d_{\text{F}} H_t(\mathbf{v}^*_t) \rangle\,.$$

In the case of potential forces, there exists a scalar function $P \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C} \times I\,; \Re)$ such that

$$\mathbf{f}(\mathbf{v}^*, t) = dP(\mathbf{v}^*, t)\,.$$

We define $\boldsymbol{\beta}^1 := -P dt(\mathbf{v}^*, t)$ to get $\boldsymbol{\alpha}^2_{\text{REG}} = d\boldsymbol{\beta}^1$.

Then, setting $P_t(\mathbf{v}^*) := P(\mathbf{v}^*, t)$, Hamilton's canonical law may be written as

$$d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t) = d(P_t - H_t)(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t)\,,$$

for all $\mathbf{Y}(\mathbf{v}^*_t) \in \mathbb{T}_{\mathbf{v}^*_t}\mathbb{T}^*\mathbb{C}$ such that $\operatorname{sym}\nabla(T\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t) \cdot \mathbf{Y}(\mathbf{v}^*_t)) = 0$.

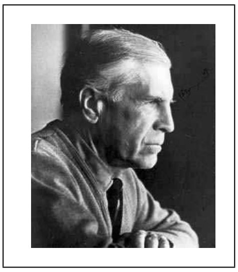

Figure 2.3: Lev Semenovich Pontryagin (1908 - 1988)

2.2.10 Action principle in the Pontryagin bundle

The action principle can be reformulated in terms of both velocity and kinetic momentum by introducing the PONTRYAGIN vector bundle $\boldsymbol{\tau}_{\mathbf{P}} \in \mathrm{C}^1(\mathbb{PC}\,;\mathbb{C})$ which is the WHITNEY sum $\mathbb{PC} := \mathbb{TC} \oplus \mathbb{T}^*\mathbb{C}$ of the tangent and the cotangent bundles, defined as the vector bundle whose fibers are the direct sums of tangent and cotangent spaces:

$$\mathbb{TC} \oplus \mathbb{T}^*\mathbb{C} := \{\mathbf{v_P} := (\mathbf{v}\,,\mathbf{v}^*) \in \mathbb{TC} \oplus \mathbb{T}^*\mathbb{C} \,:\, \boldsymbol{\tau}_{\mathbf{P}}(\mathbf{v_P}) := \boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) = \boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*)\}\,.$$

The canonical one-form $\boldsymbol{\theta}_{\mathbf{P}}(\mathbf{v_P}) \in \mathbb{T}^*_{\mathbf{v_P}}\mathbb{PC}$ is defined by

$$\langle \boldsymbol{\theta}_{\mathbf{P}}, \mathbf{X_P} \rangle(\mathbf{v_P}) := \langle \mathbf{v}^*, T\boldsymbol{\tau}_{\mathbf{P}} \cdot \mathbf{X_P}(\mathbf{v_P}) \rangle\,, \quad \forall\, \mathbf{X_P}(\mathbf{v_P}) \in \mathbb{T}_{\mathbf{v_P}}\mathbb{PC}\,,$$

and the evaluation functional $\mathrm{EVAL} \in \mathrm{C}^1(\mathbb{PC}\,;\Re)$ is given by

$$\mathrm{EVAL}(\mathbf{v_P}) := \langle \mathbf{v}^*, \mathbf{v} \rangle\,,$$

The PONTRYAGIN energy functional $E_{\mathbf{P}} \in \mathrm{C}^1(\mathbb{PC}\,;\Re)$ is then defined by

$$E_{\mathbf{P}}(\mathbf{v_P}) := \mathrm{EVAL}(\mathbf{v_P}) - L(\mathbf{v}) = \langle \mathbf{v}^*, \mathbf{v} \rangle - L(\mathbf{v})\,.$$

We may now state the following result.

Lemma 2.2.4 *The fulfillment of the differential condition*

$$d\boldsymbol{\theta}_{\mathbf{P}}(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P}) \cdot \mathbf{V_P}(\mathbf{v_P}) = -\langle d_{\mathrm{F}} E_{\mathbf{P}}(\mathbf{v_P}), \mathbf{V}(\mathbf{v_P}) \rangle\,,$$

for any vertical vector $\mathbf{V}(\mathbf{v_P}) \in \mathbb{T}_{\mathbf{v_P}}\mathbb{P}_{\boldsymbol{\pi}_{\mathbf{P}}(\mathbf{v_P})}\mathbb{C}$*, is equivalent to require that:*

$$\begin{cases} \mathbf{v}^* = d_{\mathrm{F}} L(\mathbf{v})\,, \\ \mathbf{v} \ = T\boldsymbol{\tau}_{\mathbf{P}}(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P})\,. \end{cases}$$

Proof. Let us consider the extension $\dot{\mathcal{F}}_{\mathbf{P}} := \mathbf{Fl}_\lambda^{\mathbf{V_P}} \uparrow \mathbf{X_P}$ of the vector field $\mathbf{X_P}(\mathbf{v_P}) \in \mathbb{T}_{\mathbf{v_P}}\mathbb{T}^*\mathbb{C}$ along the trajectory by pushing it along the flow $\mathbf{Fl}_\lambda^{\mathbf{V_P}} \in \mathrm{C}^1(\mathbb{PC}\,;\mathbb{PC})$. Then PALAIS formula gives

$$d\boldsymbol{\theta}_{\mathbf{P}}(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P}) \cdot \mathbf{V_P}(\mathbf{v_P}) = d_{\mathbf{X_P}(\mathbf{v_P})}(\boldsymbol{\theta}_{\mathbf{P}} \cdot \mathbf{V_P})(\mathbf{v_P}) - d_{\mathbf{V_P}(\mathbf{v_P})}(\boldsymbol{\theta}_{\mathbf{P}} \cdot \dot{\mathcal{F}}_{\mathbf{P}})(\mathbf{v_P}) \,,$$

with $d_{\mathbf{X}(\mathbf{v_P})}(\boldsymbol{\theta}_{\mathbf{P}} \cdot \mathbf{V_P})(\mathbf{v_P}) = \partial_{\tau=t} \, \langle \mathbf{v}^*(\tau), T\boldsymbol{\tau}_{\mathbf{P}}(\mathbf{v_P}(\tau)) \cdot \mathbf{V_P}(\mathbf{v_P}(\tau)) \rangle = 0$ and

$$\begin{aligned} d_{\mathbf{V_P}(\mathbf{v_P})}(\boldsymbol{\theta}_{\mathbf{P}} \cdot \dot{\mathcal{F}}_{\mathbf{P}})(\mathbf{v_P}) &= \partial_{\lambda=0} \, \langle \boldsymbol{\theta}_{\mathbf{P}}(\mathbf{Fl}_\lambda^{\mathbf{V_P}}(\mathbf{v_P})), (\mathbf{Fl}_\lambda^{\mathbf{V_P}} \uparrow \mathbf{X_P})(\mathbf{Fl}_\lambda^{\mathbf{V_P}}(\mathbf{v_P})) \rangle \\ &= \partial_{\lambda=0} \, \langle \mathbf{Fl}_\lambda^{\mathbf{V_P}}(\mathbf{v}_t^*), T\boldsymbol{\tau}_{\mathbf{P}}(\mathbf{Fl}_\lambda^{\mathbf{V_P}}(\mathbf{v_P})) \cdot T\mathbf{Fl}_\lambda^{\mathbf{V_P}}(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P}) \rangle \\ &= \partial_{\lambda=0} \, \langle \mathbf{Fl}_\lambda^{\mathbf{V_P}}(\mathbf{v_P}), T(\boldsymbol{\tau}_{\mathbf{P}} \circ \mathbf{Fl}_\lambda^{\mathbf{V_P}})(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P}) \rangle \\ &= \partial_{\lambda=0} \, \langle \mathbf{Fl}_\lambda^{\mathbf{V_P}}(\mathbf{v_P}), T\boldsymbol{\tau}_{\mathbf{P}}(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P}) \rangle \\ &= \langle \mathbf{w}^*, T\boldsymbol{\tau}_{\mathbf{P}}(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P}) \rangle \,, \end{aligned}$$

where by verticality $\boldsymbol{\tau}_{\mathbf{P}} \circ \mathbf{Fl}_\lambda^{\mathbf{V_P}} = \boldsymbol{\tau}_{\mathbf{P}}$ and the pair $(\mathbf{w}\,, \mathbf{w}^*) \in \mathbb{PC}$ is defined by the vertical lift:

$$\mathbf{vl}_{\mathbb{PC}}(\mathbf{v_P}) \cdot (\mathbf{w}\,, \mathbf{w}^*) = \mathbf{V_P}(\mathbf{v_P}) \,.$$

On the other hand we have that

$$\langle d_{\mathrm{F}} E_{\mathbf{P}}(\mathbf{v_P}), \mathbf{V_P}(\mathbf{v_P}) \rangle = d\mathrm{EVAL}(\mathbf{v_P}) \cdot \mathbf{V_P}(\mathbf{v_P}) - d_{\mathrm{F}} L(\mathbf{v}) \cdot \mathbf{w} \,.$$

A direct computation gives

$$\begin{aligned} d\mathrm{EVAL}(\mathbf{v}\,, \mathbf{v}^*) \cdot (\mathbf{w}\,, \mathbf{w}^*) &= \lim_{\lambda \to 0} \frac{1}{\lambda} \big[\langle \mathbf{v}^* + \lambda \mathbf{w}^*, \mathbf{v} + \lambda \mathbf{w} \rangle - \langle \mathbf{v}^*, \mathbf{v} \rangle \big] \\ &= \langle \mathbf{v}^*, \mathbf{w} \rangle - \langle \mathbf{w}^*, \mathbf{v} \rangle \,. \end{aligned}$$

Hence the differential condition in the statement may be written as

$$\langle \mathbf{v}^* - d_{\mathrm{F}} L(\mathbf{v}), \mathbf{w} \rangle + \langle \mathbf{w}^*, \mathbf{v} - T\boldsymbol{\tau}_{\mathbf{P}}(\mathbf{v_P}) \cdot \mathbf{X_P}(\mathbf{v_P}) \rangle = 0 \,.$$

By the arbitrarity of $(\mathbf{w}\,, \mathbf{w}^*) \in \mathbb{PC}$ the result follows. ■

The action principle for a trajectory $\mathbf{\Gamma}_{\mathbf{P}I}$ in the extended PONTRYAGIN bundle $\mathbb{PC} \times I$ is expressed by the variational condition:

$$\partial_{\lambda=0} \int_{(\mathbf{Fl}_\lambda^{\mathbf{Y}} \times \mathbf{Fl}_\lambda^{\Theta})(\mathbf{\Gamma}_{\mathbf{P}I})} \boldsymbol{\omega}^1 = \int_{\partial \mathbf{\Gamma}_{\mathbf{P}I}} \boldsymbol{\omega}^1 \cdot (\mathbf{Y}\,, \Theta) \,,$$

for any time-flow $\mathbf{Fl}_\lambda^\Theta \in \mathrm{C}^1(I\,;I)$ with velocity vector field $\Theta \in \mathrm{C}^1(I\,;\mathbb{T}I)$ and any automorphic flow $\mathbf{Fl}_\lambda^\mathbf{Y} \in \mathrm{C}^1(\mathbb{P}\mathbb{C}\,;\mathbb{P}\mathbb{C})$, with projected flow $\varphi_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ defined by the commutative diagram:

$$\begin{array}{ccc} \mathbb{P}\mathbb{C} & \xrightarrow{\mathbf{Fl}_\lambda^\mathbf{Y}} & \mathbb{P}\mathbb{C} \\ \boldsymbol{\tau}_\mathbf{P}\downarrow & & \boldsymbol{\tau}_\mathbf{P}\downarrow \\ \mathbb{C} & \xrightarrow{\varphi_\lambda} & \mathbb{C} \end{array} \quad \iff \quad \boldsymbol{\tau}_\mathbf{P} \circ \mathbf{Fl}_\lambda^\mathbf{Y} = \varphi_\lambda \circ \boldsymbol{\tau}_\mathbf{P}\,.$$

2.2.11 Symplectic and contact manifolds

The basic property of the canonical two-form $d\boldsymbol{\theta}\mathbb{C} \in \Lambda^2(\mathbb{T}\mathbb{C}\,;\Re)$ is its weak nondegeneracy (Theorem 1.8.2, page 189):

$$d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}^*)\cdot\mathbf{X}(\mathbf{v}^*)\cdot\mathbf{Y}(\mathbf{v}^*) = 0\,, \quad \forall\,\mathbf{Y}(\mathbf{v}^*) \in \mathbb{T}_{\mathbf{v}^*}\mathbb{T}^*\mathbb{C} \implies \mathbf{X}(\mathbf{v}^*) = 0\,.$$

Then we say that

Definition 2.2.2 (Exact symplectic manifold) *The velocity phase-space* $\mathbb{T}\mathbb{C}$*, endowed with the exact two-form* $d\boldsymbol{\theta}_{L_t} \in \Lambda^2(\mathbb{T}\mathbb{C}\,;\Re)$*, is an exact symplectic manifold.*

In a symplectic manifold HAMILTON's equation

$$d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}_t^*)\cdot\mathbf{X}(\mathbf{v}_t^*)\cdot\mathbf{Y}(\mathbf{v}_t^*) = (\mathbf{f}_t(\mathbf{v}_t^*) - dH_t(\mathbf{v}_t^*))\cdot\mathbf{Y}(\mathbf{v}_t^*)\,, \quad \forall\,\mathbf{Y}(\mathbf{v}_t^*) \in \mathbb{T}_{\mathbf{v}_t^*}\mathbb{T}^*\mathbb{C}\,,$$

admits a unique solution (if any).

On the other hand, recalling that the action one-form is given by

$$\boldsymbol{\omega}^1(\mathbf{v}_t^*, t) := \mathrm{pr}_{\mathbb{T}^*\mathbb{C}}\!\downarrow\!\boldsymbol{\theta}\mathbb{C}(\mathbf{v}_t^*) - \boldsymbol{\eta}(\mathbf{v}_t^*, t)\,,$$

we have that

$$d\boldsymbol{\omega}^1(\mathbf{v}_t^*, t) := \mathrm{pr}_{\mathbb{T}^*\mathbb{C}}\!\downarrow\! d\boldsymbol{\theta}\mathbb{C}(\mathbf{v}_t^*) - d\boldsymbol{\eta}(\mathbf{v}_t^*, t)\,,$$

so that, normalizing the time speed to the unity, EULER's differential condition of stationarity writes

$$(d\boldsymbol{\omega}^1 - \boldsymbol{\alpha}^2_{\text{REG}})(\mathbf{v}_t^*, t)\cdot(\mathbf{X}(\mathbf{v}_t^*)\,, 1_t)\cdot(\mathbf{Y}(\mathbf{v}_t^*)\,, \Theta(t)) = 0\,,$$

for all $\mathbf{Y}(\mathbf{v}^*) \in \mathbb{T}_{\mathbf{v}^*}\mathbb{T}^*\mathbb{C}$ and all $\Theta(t) \in \mathbb{T}_t I$.

Since the normalized solution $\mathbf{X}(\mathbf{v}_t^*) \in \mathbb{T}_{\mathbf{v}_t^*}\mathbb{T}^*\mathbb{C}$ of EULER's differential condition is also the unique solution of HAMILTON's equation, we infer that the form $d\boldsymbol{\omega}^1(\mathbf{v}^*, t) \in \Lambda^2(\mathbb{T}_{\mathbf{v}^*}\mathbb{T}^*\mathbb{C} \times \mathbb{T}_t I\,;\Re)$ has a 1-D kernel.

Then we have that

Definition 2.2.3 *The covelocity-time state-space* $\mathbb{T}^*\mathbb{C} \times I$ *, endowed with the exact two-form* $d\boldsymbol{\omega}^1 \in \Lambda^2(\mathbb{T}\mathbb{C} \times I\,; \Re)$ *, is an* ***exact contact manifold****.*

If the Lagrangian has a nonsingular fiber derivative, the velocity-time state-space $\mathbb{T}\mathbb{C} \times I$, endowed with the exact two-form $d\boldsymbol{\omega}^1_L \in \Lambda^2(\mathbb{T}\mathbb{C} \times I\,; \Re)$, is also an *exact contact manifold.*

2.2.12 Constrained Hamilton's principle

The proof of the classical Maupertuis' principle given in [2], Theorem 3.8.5 on page 249 , considers a trajectory in the configuration manifold and its asynchronous variations in the configuration manifold in which end-points and instantaneous energy are held fixed while varying start and end-time instants. Asynchronous variations are needed since there could be no path joining the end-points with the same constant energy and the same start and end-time, other than the given trajectory. The treatment in [2] is developed in terms of coordinates.

Our approach provides instead an intrinsic formulation of a constrained Hamilton's principle (**CHP**) in the velocity-time state-space, thus allowing for a direct application of Lagrange's multipliers method to show its equivalence to the geometric form of Hamilton's principle in Proposition 2.2.4 which will be called **UHP** (Unconstrained Hamilton Principle). No asynchronous variations are needed since the **CHP** is formulated as a geometric action principle in the velocity-time state-space. Moreover the energy conservation constraint must be imposed pointwise only on the virtual velocities and not along the varied trajectories. The idea underlying the proof is the following. It is straightforward to see that a trajectory fulfilling the **UHP** is also solution of the **CHP** in which the energy conservation constraint is imposed on test velocity fields. Not trivial is the converse implication, that the geometric trajectory provided by the **CHP** is also solution of the **UHP**. The non-trivial part of the proof is based on Lagrange's multipliers method and this in turn relies upon Banach's closed range theorem in Functional Analysis. In this respect we notice that improper applications of Lagrange's multipliers method outside its range of validity have led, also in recent times, to erroneous statements and results in mechanics, as discussed in [173]. We formulate a variational statement of the **CHP** valid for any dynamical system, including time-dependent lagrangians and non-potential or time-dependent forces. Impulsive forces are not explicitly considered for brevity but could be easily accounted for. The classical Maupertuis' least action principle will be later directly recovered under the special assumption of conservativity. A more general principle which we still call MAUPERTUIS' principle is got under the assumption that the energy and the force do not depend directly on time.

The POINCARÉ-CARTAN one-form $\boldsymbol{\theta}_L \in \mathrm{C}^1(\mathbb{TC} \times I\,; \mathbb{T}^*(\mathbb{TC} \times I))$ in the velocity-time state-space is defined along the trajectory $\mathbf{\Gamma}_I \subset \mathbb{TC} \times I$ by:

$$\langle \boldsymbol{\theta}_L, (\mathbf{Y}\,, \Theta) \rangle(\mathbf{v}_t, t) := \langle \boldsymbol{\theta}_{L_t}, \mathbf{Y} \rangle(\mathbf{v}_t)\,,$$

and the energy functional $E \in \mathrm{C}^1(\mathbf{\Gamma}_I\,; \Re)$ is given by $E(\mathbf{v}_t, t) := E_t(\mathbf{v}_t)$ with $(\mathbf{v}_t\,, t) = \Gamma(t)$.

Lemma 2.2.5 *Defining the energy one-form* $\boldsymbol{\eta} \in \mathrm{C}^1(\mathbb{TC} \times I\,; \mathbb{T}^*(\mathbb{TC} \times I))$ *:*

$$\boldsymbol{\eta}(\mathbf{v}_t, t) := E(\mathbf{v}_t, t)\, dt\,,$$

the exterior derivative yields the formula:

$$[d\boldsymbol{\eta} \cdot (\mathbf{Y}\,, 0) \cdot (\mathbf{X}\,, 1)](\mathbf{v}_t, t) = dE_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\,.$$

Proof. The computation may be performed by PALAIS formula by extending the vector $(\mathbf{X}(\mathbf{v}_t)\,, 1_t) \in \mathbb{T}_{(\mathbf{v}_t\,,t)}\mathbf{\Gamma}_I$ to a field $\mathcal{F} \in \mathrm{C}^1(\mathbb{TC} \times I\,; \mathbb{T}(\mathbb{TC} \times I))$ by pushing it along the flow $\mathbf{Fl}_\lambda^{(\mathbf{Y},0)} \in \mathrm{C}^1(\mathbf{\Gamma}_I\,; \mathbb{TC} \times I)$, according to the relation:

$$\mathcal{F}(\mathbf{Fl}_\lambda^{(\mathbf{Y},0)}(\mathbf{v}_t, t)) := (\mathbf{Fl}_\lambda^{\mathbf{Y}} {\uparrow} \mathbf{X}(\mathbf{v}_t)\,, 1_t)_{(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t)\,,t)}\,.$$

Then PALAIS formula tells us that

$$\begin{aligned} d\boldsymbol{\eta}(\mathbf{v}_t, t) \cdot (\mathbf{Y}(\mathbf{v}_t)\,, 0_t) \cdot (\mathbf{X}(\mathbf{v}_t)\,, 1_t) &= d_{(\mathbf{Y}(\mathbf{v}_t)\,,0_t)} \langle \boldsymbol{\eta}, \mathcal{F} \rangle \\ &- d_{(\mathbf{X}(\mathbf{v}_t)\,,1_t)} \langle \boldsymbol{\eta}, (\mathbf{Y}\,, 0) \rangle + \langle \boldsymbol{\eta}, \mathcal{L}_{(\mathbf{Y}\,,0)} \mathcal{F} \rangle(\mathbf{v}_t, t)\,. \end{aligned}$$

Since, by the chosen extension, the LIE derivative $\mathcal{L}_{(\mathbf{Y}\,,0)}\mathcal{F}$ vanishes, we may evaluate as follows:

$$d_{(\mathbf{X}(\mathbf{v}_t)\,,1_t)} \langle \boldsymbol{\eta}, (\mathbf{Y}\,, 0) \rangle = \partial_{\tau=t} \langle \boldsymbol{\eta}(\mathbf{v}_\tau, \tau), (\mathbf{Y}(\mathbf{v}_\tau)\,, 0) \rangle = \partial_{\tau=t} E_\tau(\mathbf{v}_\tau) \langle\, d\tau, 0 \rangle = 0\,,$$

$$d_{(\mathbf{Y}(\mathbf{v}_t)\,,0_t)} \langle \boldsymbol{\eta}, \mathcal{F} \rangle = \partial_{\lambda=0}\, E_t(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t))\, \langle dt, 1_t \rangle = \partial_{\lambda=0}\, E_t(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t))\,.$$

Summing up, being $\partial_{\lambda=0}\, E_t(\mathbf{Fl}_\lambda^{\mathbf{Y}}(\mathbf{v}_t)) = dE_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)$, we get the result.
■

Theorem 2.2.2 (Constrained Hamilton Principle) *A trajectory* $\mathbf{\Gamma}_I$ *in the velocity-time state-space* $\mathbb{TC} \times I$ *of a dynamical system governed by a time-dependent energy* $E_t \in \mathrm{C}^1(\mathbf{\Gamma}; \Re)$ *and subject to time-dependent forces* $\mathbf{F}_t \in \mathrm{C}^1(\boldsymbol{\gamma}; \mathbb{T}^*\mathbb{C})$, *where* $\boldsymbol{\gamma} = \boldsymbol{\tau}_{\mathbb{C}}(\mathbf{\Gamma})$, *is a* 1*-D submanifold* $\mathbf{\Gamma}_I \subset \mathbb{TC} \times I$ *fulfilling the geometric action principle:*

$$\partial_{\lambda=0} \int_{\mathbf{Fl}_\lambda^{(\mathbf{Y},0)}(\mathbf{\Gamma}_I)} \boldsymbol{\theta}_L = \oint_{\partial \mathbf{\Gamma}_I} \boldsymbol{\theta}_L \cdot \mathbf{Y} \,,$$

for any virtual velocity field fulfilling the energy conservation law:

$$\mathbf{Y}(\mathbf{v}_t) \in \boldsymbol{ker}((dE_t - \mathbf{f}_t)(\mathbf{v}_t)) \subset \mathbb{T}_{\mathbf{v}_t}\mathbb{TC} \,.$$

Proof. Let us prove that the above statement, denoted **CHP** (Constrained Hamilton Principle), is equivalent to the action principle in Proposition 2.2.4, denoted **UHP** (Unconstrained Hamilton Principle). Indeed the latter, in the synchronous case, being $\boldsymbol{\omega}_L^1 = \boldsymbol{\theta}_L - \boldsymbol{\eta}$ may be written as:

$$\partial_{\lambda=0} \int_{\mathbf{Fl}_\lambda^{(\mathbf{Y},0)}(\mathbf{\Gamma}_I)} \boldsymbol{\theta}_L - \oint_{\partial \mathbf{\Gamma}_I} \boldsymbol{\theta}_L \cdot (\mathbf{Y}\,, 0) = \int_{\mathbf{\Gamma}_I} (d\boldsymbol{\eta} + \boldsymbol{\alpha}^2_{\mathrm{REG}}) \cdot (\mathbf{Y}\,, 0) \,,$$

for any field $\mathbf{Y} \in \mathrm{C}^1(\mathbf{\Gamma}; \mathbb{TTC})$. Along a time-parametrized trajectory, by Lemma 2.2.5 we have that

$$(d\boldsymbol{\eta} + \boldsymbol{\alpha}^2_{\mathrm{REG}}) \cdot (\mathbf{Y}\,, 0) \cdot (\mathbf{X}\,, 1)(\mathbf{v}_t, t) = (dE_t - \mathbf{f}_t)(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) \,.$$

Hence clearly the **UHP** implies the **CHP**. The converse implication is proved by comparing EULER's conditions for both action principles. By the extrusion formula, the expression of the **UHP** becomes:

$$\int_{\mathbf{\Gamma}_I} (d\boldsymbol{\theta}_L - d\boldsymbol{\eta} - \boldsymbol{\alpha}^2_{\mathrm{REG}}) \cdot (\mathbf{Y}\,, 0) = 0 \,, \quad \forall\, \mathbf{Y} \in \mathrm{C}^1(\mathbf{\Gamma}; \mathbb{TTC}) \,,$$

and the expression of the **CHP** may be written as:

$$\int_{\mathbf{\Gamma}_I} d\boldsymbol{\theta}_L \cdot (\mathbf{Y}\,, 0) = 0 \,, \quad \forall\, (\mathbf{Y}\,, 0) \in \mathbf{ker}((d\boldsymbol{\eta} + \boldsymbol{\alpha}^2_{\mathrm{REG}}) \cdot (\mathbf{X}\,, 1)) \,.$$

Moreover, being

$$[d\boldsymbol{\theta}_L \cdot (\mathbf{Y}\,, 0) \cdot (\mathbf{X}\,, 1)](\mathbf{v}_t, t) = d\boldsymbol{\theta}_{L_t} \cdot \mathbf{Y}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) \,,$$

the **UHP** and **CHP** are respectively equivalent to the EULER's conditions:

$$d\boldsymbol{\theta}_{L_t} \cdot \mathbf{X}(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) = (\mathbf{f}_t - dE_t)(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) \,, \quad \forall\, \mathbf{Y} \in \mathrm{C}^1(\boldsymbol{\Gamma}\,; \mathbb{TTC}) \,,$$

$$d\boldsymbol{\theta}_{L_t} \cdot \mathbf{X}(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) = 0 \,, \quad \forall\, \mathbf{Y}(\mathbf{v}_t) \in \mathbf{ker}((\mathbf{f}_t - dE_t)(\mathbf{v}_t)) \,.$$

By the non-degeneracy of the two-form $d\boldsymbol{\theta}_{L_t}$ the former equation admits a unique solution $\mathbf{X}(\mathbf{v}_t) \in \mathbb{T}_{\mathbf{v}_t}\mathbb{TC}$. The solution of the latter homogeneous equation is instead definite to within a scalar factor. The former condition clearly implies the latter one, in the sense that the solution of the former is also a solution of the latter. The converse implication, that there is a solution of the latter which is also solution of the former, is proved by LAGRANGE's multiplier method. The argument is as follows. Setting $\mathbf{f}_E(\mathbf{v}_t) := (\mathbf{f}_t - dE_t)(\mathbf{v}_t) \in \mathbb{T}^*_{\mathbf{v}_t}\mathbb{TC} = BL(\mathbb{T}_{\mathbf{v}_t}\mathbb{TC}\,; \Re)$, the subspace $\mathbf{im}(\mathbf{f}_E(\mathbf{v}_t)) = \Re$ is trivially closed and hence $\mathbf{ker}(\mathbf{f}_E(\mathbf{v}_t))^0 = \mathbf{im}(\mathbf{f}_E(\mathbf{v}_t)')$ by BANACH closed range theorem [101]. Here $\mathbf{f}_E(\mathbf{v}_t)' \in BL(\Re\,; \mathbb{T}^*_{\mathbf{v}_t}\mathbb{TC})$ is the dual operator. The latter condition writes $d\boldsymbol{\theta}_{L_t} \cdot \mathbf{X}(\mathbf{v}_t) \in \mathbf{ker}(\mathbf{f}_E(\mathbf{v}_t))^0$ and hence the result above ensures the existence of a $\mu(\mathbf{v}_t) \in \Re$ such that

$$d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t) = \langle \mathbf{f}_E(\mathbf{v}_t)' \cdot \mu(\mathbf{v}_t), \mathbf{Y}(\mathbf{v}_t) \rangle \,, \quad \forall\, \mathbf{Y}(\mathbf{v}_t) \in \mathbb{T}_{\mathbf{v}_t}\mathbb{TC} \,,$$

equivalent to $d\boldsymbol{\theta}_{L_t}(\mathbf{v}_t) \cdot \mathbf{X}(\mathbf{v}_t) = \mu(\mathbf{v}_t)\, \mathbf{f}_E(\mathbf{v}_t)$. Then the field $\mathbf{X}(\mathbf{v}_t)/\mu(\mathbf{v}_t)$ is solution of both EULER's conditions. The LAGRANGE's multipliers provide a field of scaling factors to get the right time schedule along the trajectory. ■

According to Lemma 2.2.2, EULER differential condition ensures that the trajectory $\Gamma \in \mathrm{C}^1(I\,; \mathbb{TC})$ is the lifting to the tangent bundle of the trajectory $\gamma \in \mathrm{C}^1(I\,; \mathbb{C})$, so that $\mathbf{v}_t = \Gamma(t) = \partial_{\tau=t}\, \gamma(\tau)$. Then the virtual flow may be defined as $\mathbf{Fl}^{\mathbf{Y}}_\lambda = T\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,; \mathbb{TC})$ with $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^2(\mathbb{C}\,; \mathbb{C})$ and the virtual velocity is given by $\mathbf{Y} = \mathbf{v}_{T\boldsymbol{\varphi}} = \partial_{\lambda=0}\, T\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{TC}\,; \mathbb{TTC})$.

The variational condition of the constrained principle of Theorem 2.2.2 can then be written explicitly, in terms of the *action functional* $\mathrm{A}_t(\mathbf{v}_t) := \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_t \rangle$ associated with the Lagrangian and of the virtual flow in the configuration manifold as

$$\partial_{\lambda=0} \int_I \mathrm{A}_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t))\, dt = \oint_{\partial I} \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle \, dt \,,$$

with virtual velocities fulfilling conservation of energy, i.e.:

$$dE_t(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = \mathbf{F}_t(\mathbf{v}_t) \cdot \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \,.$$

The action functional is also referred to in the literature as the *reduced action* functional, to underline that the energy term is missing in comparison with HAMILTON's stationarity principle for the lagrangian $L_t := \mathrm{A}_t - E_t$. We underline that, in spite of the explicit appearance of flow and tangent flow in the expression of the principle, only the virtual velocity $\mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))$ along the trajectory $\boldsymbol{\gamma} = \boldsymbol{\tau}_{\mathbb{C}}(\boldsymbol{\Gamma})$ is influent in the formulation of the law of dynamics. In fact virtual flows with coincident initial velocities provide the same test condition. This basic property, which is here hidden by the imposition of the constraint of energy conservation, may be proven on the basis of the equivalent geometric HAMILTON's principle, by introducing a connection in the configuration manifold to get a generalized formulation of LAGRANGE's law of dynamics [174], [175]. The **CHP** may be stated with an equivalent formulation in which the constraint of energy conservation on the virtual velocities is imposed in integral form.

Theorem 2.2.3 (Constrained Hamilton principle: an equivalent form) *A trajectory $\boldsymbol{\Gamma}_I$ of a dynamical system in the velocity-time state-space $\mathbb{TC} \times I$ is a path fulfilling the geometric action principle:*

$$\partial_{\lambda=0} \int_{\mathbf{Fl}_{\lambda}^{(\mathbf{Y},0)}(\boldsymbol{\Gamma}_I)} \boldsymbol{\theta}_L = \oint_{\partial \boldsymbol{\Gamma}_I} \boldsymbol{\theta}_L \cdot \mathbf{Y} \,,$$

for any tangent field $\mathbf{Y} \in \mathrm{C}^1(\boldsymbol{\Gamma}\,; \mathbb{TTC})$ such that

$$\int_I dE_t(\mathbf{v}_t) \cdot \mathbf{Y}(\mathbf{v}_t)\, dt = \int_I \langle \mathbf{F}_t(\mathbf{v}_t), T\boldsymbol{\tau}_{\mathbb{C}} \cdot \mathbf{Y}(\mathbf{v}_t) \rangle \, dt \,.$$

Proof. A trajectory fulfils the action principle of Proposition 2.2.4 and hence a fortiori the constrained principle of Theorem 2.2.3 and then again a fortiori the weaker condition of the principle in Theorem 2.2.2. Since this latter is equivalent to the action principle of Proposition 2.2.4, the circle of implications is closed and the assertion is proven. ■

2.2.13 Maupertuis' least action principle

In the presentation of the least action principle, we shall not follow the standard treatment due to MAUPERTUIS, EULER, LAGRANGE and JACOBI [6], but will instead derive the result by a direct specialization of the constrained geometric HAMILTON principle.

When the lagrangian $L \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,;\Re)$ is time-independent and the system is subject to time-independent forces $\mathbf{F} \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{T}^*\mathbb{C})$, the constraint of energy conservation on the virtual velocity field is independent of time. Then the projected trajectory in the velocity phase-space can be arbitrarily parametrized and the **CHP** directly yields an extended version of MAUPERTUIS' principle in which the dynamical system is not necessarily conservative.

Theorem 2.2.4 (Maupertuis principle) *In a dynamical system governed by a time-independent lagrangian functional* $L \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,;\Re)$ *and subject to time-independent forces* $\mathbf{F} \in \mathrm{C}^1(\boldsymbol{\gamma}\,;\mathbb{T}^*\mathbb{C})$, *the trajectories are* 1*-D submanifolds* $\boldsymbol{\Gamma} \subset \mathbb{T}\mathbb{C}$ *of the velocity phase-space with tangent vectors* $\mathbf{X}(\mathbf{v}) := \partial_{\mu=\lambda}\,\Gamma(\mu) \in \mathbb{T}_{\mathbf{v}}\boldsymbol{\Gamma}$, *with* $\mathbf{v} := \Gamma(\lambda)$, *fulfilling the homogeneous* EULER*'s condition:*

$$d\boldsymbol{\theta}_L(\mathbf{v}) \cdot \mathbf{X}(\mathbf{v}) \cdot \mathbf{Y}(\mathbf{v}) = 0\,, \quad \mathbf{X}(\mathbf{v}) \in \mathbb{T}_{\mathbf{v}}\mathbb{T}\mathbb{C}\,,$$

for any virtual velocity field fulfilling the energy conservation law: $\mathbf{Y}(\mathbf{v}) \in \boldsymbol{ker}((T^*\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{F} - dE)(\mathbf{v})) \subset \mathbb{T}_{\mathbf{v}}\mathbb{T}\mathbb{C}$. *The associated geometric action principle in the phase-space* $\mathbb{T}\mathbb{C}$ *is expressed by the variational condition:*

$$\partial_{\lambda=0} \int_{\mathbf{Fl}^{\mathbf{Y}}_{\lambda}(\boldsymbol{\Gamma})} \boldsymbol{\theta}_L = \oint_{\partial\boldsymbol{\Gamma}} \boldsymbol{\theta}_L \cdot \mathbf{Y}\,,$$

stating the stationarity of the action integral of the POINCARÉ-CARTAN *one-form* $\boldsymbol{\theta}_L = T^*\boldsymbol{\tau}_{\mathbb{C}} \circ d_{\mathrm{F}}L$ *for all virtual flows* $\mathbf{Fl}^{\mathbf{Y}}_{\lambda} \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,;\mathbb{T}\mathbb{C})$ *with an energy conserving virtual velocity* $\mathbf{Y}(\mathbf{v}) \in \boldsymbol{ker}((T^*\boldsymbol{\tau}_{\mathbb{C}} \circ \mathbf{F} - dE)(\mathbf{v}))$.

Figure 2.4: Pierre Louis Moreau de Maupertuis (1698 - 1759)

An alternative statement can be deduced from the one in Theorem 2.2.3. The MAUPERTUIS principle of Theorem 2.2.4 is a geometric action principle whose solutions are determinate to within an arbitrary reparametrization. The relevant

EULER condition is homogeneous in the trajectory speed and hence provides the geometry of the trajectory but not the time law according to which it is travelled by the dynamical system. Anyway, if the dynamical trajectory in the velocity-time state space is projected on the velocity phase-space, both MAUPERTUIS' principle and energy conservation are fulfilled. Therefore the time schedule is recoverable from the initial condition on the velocity by imposing conservation of energy along the geometric trajectory evaluated by MAUPERTUIS' principle. For conservative systems the statement in Theorem 2.2.4 specializes into the classical formulation of the least action principle due to MAUPERTUIS [113, 114], EULER [51], LAGRANGE [89], JACOBI [74, 75] which has been reproduced without exceptions in the literature, see e.g. [90], [2], [6], [3]. The principle deduced from Theorem 2.2.4 is however more general than the classical one because it is formulated without making the standard assumption of fixed end-points of the base trajectory in the configuration manifold and also without assuming that the trajectory developes in a constant energy leaf. Indeed our statement underlines that the constant energy constraint is imposed only on virtual velocities in the velocity phase-space and not on the trajectory speed.

Remark 2.2.9 *In the papers [52] and [53] the authors claim that the classical* MAUPERTUIS*' principle for conservative systems can be given an equivalent formulation by assuming that the trajectory is varied under the assumption of an invariant mean value of the energy (they call this statement the general* MAUPERTUIS*' principle **GMP**). The sketched proof provided in these papers is however inficiated by the misstatement that the fulfilment of the original* MAUPERTUIS*' principle (**MP**), in which the energy is constant under the variations, implies the fulfilment of the **GMP**. But this last variational condition has more variational test fields and hence the converse is true. The implication proved in [52] and [53], that **GMP** implies **MP**, is then trivial and the nontrivial converse implication is missing. Theorem 2.2.3 shows that the pointwise condition:* $(dE \cdot \mathbf{Y})(\mathbf{v}_t) = 0$ *, and the integral condition:*

$$\partial_{\lambda=0} \int_I E(\mathbf{Fl}^{\mathbf{Y}}_\lambda(\mathbf{v}_t))\, dt = \int_I (dE \cdot \mathbf{Y})(\mathbf{v}_t)\, dt = 0\,,$$

on the virtual velocity field lead to equivalent formulations of the classical MAUPERTUIS*' principle.*

The long controversy concerning the least action principle, initiated with the ugly dispute in 1751 between MAUPERTUIS and SAMUEL KÖNIG who claimed that MAUPERTUIS had plagiarized a previous result due to LEIBNIZ who communicated it to JACOB HERMANN in a letter dated 1707. VOLTAIRE, in support of König, d'Alembert and Euler and the king of Prussia Frederick the Great, in support of Maupertuis, where involved at the center of the dispute, but the original of the incriminating letter was not found.

In [6], footnote on page 243, Arnold says:

In almost all textbooks, even the best, this principle is presented so that it is impossible to understand (C. Jacobi, Lectures on Dynamics, 1842 - 1843). I do not choose to break with tradition. A very interesting "proof" of Maupertuis principle is in Section 44 of the mechanics textbook of Landau and Lifšhits (Mechanics, Oxford, Pergamon, 1960).

In [2], footnote on page 249, Abraham & Marsden write:

We thank M. Spivak for helping us to formulate this theorem correctly. The authors, like many others (we were happy to learn), were confused by the standard textbook statements.

The formulation given above should end the long and laborious track followed by this principle. Here a statement of Maupertuis least action principle as a special case of a general variational principle of dynamics has been provided with a simple and clear mathematical proof.

2.3 Dynamics in a manifold with a connection

Let us assume that the configuration manifold $\mathbb{C}$ be endowed with an affine connection ∇ and with the associated parallel transport. We denote by $\mathbf{c}_{\tau,t}\Uparrow$ the parallel transport along a curve $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{C})$ from the point $\mathbf{c}(t) \in \mathbb{C}$ to the point $\mathbf{c}(\tau) \in \mathbb{C}$, setting $\mathbf{c}_{t,\tau}\Downarrow := \mathbf{c}_{\tau,t}\Uparrow$.

A vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ is parallel transported along $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{C})$ if its covariant derivative along the tangent vanishes:

$$\nabla_{\dot{\mathbf{c}}_t}\mathbf{v} = 0\,, \quad \forall\, t \in I\,.$$

Then $\mathbf{v}(\mathbf{c}(\tau)) = \mathbf{c}_{\tau,t}\Uparrow \mathbf{v}(\mathbf{c}(t))\,, \quad \forall\, \tau, t \in I$. The covariant derivative of a vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ may be expressed in terms of parallel transport as:

$$\nabla_{\dot{\mathbf{c}}_t}\mathbf{v} = \partial_{\tau=t}\ \mathbf{c}_{\tau,t}\Downarrow \mathbf{v}(\mathbf{c}(\tau))\,.$$

Indeed, if $\mathbf{v}(\mathbf{c}(\tau)) = \mathbf{c}_{\tau,t}\Uparrow \mathbf{v}(\mathbf{c}(t))$, then $\nabla_{\dot{\mathbf{c}}_t}\mathbf{v} = 0$.

The parallel transport of a covector field $\boldsymbol{\omega} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ is defined by

$$\langle\, \mathbf{c}_{\tau,t}\Uparrow \boldsymbol{\omega}(\mathbf{c}(t)), \mathbf{w}\,\rangle = \langle\, \boldsymbol{\omega}(\mathbf{c}(t)), \mathbf{c}_{\tau,t}\Downarrow \mathbf{w}\,\rangle\,, \quad \forall\, \mathbf{w} \in \mathbb{T}_{\mathbf{c}(\tau)}\mathbb{C}\,,$$

so that the parallel transport of the duality pairing is invariant:

$$\langle \mathbf{c}_{\tau,t}\Uparrow \boldsymbol{\omega}(\mathbf{c}(t)), \mathbf{c}_{\tau,t}\Uparrow \mathbf{v}(\mathbf{c}(t)) \rangle = \langle \boldsymbol{\omega}(\mathbf{c}(t)), \mathbf{v}(\mathbf{c}(t)) \rangle \,, \quad \forall\, \mathbf{v}(\mathbf{c}(t)) \in \mathbb{T}_{\mathbf{c}(t)}\mathbb{C} \,.$$

Accordingly, the covariant derivative of a covector field $\boldsymbol{\omega} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ along the vector $\dot{\mathbf{c}}_t \in \mathbb{T}_{\mathbf{c}_t}\mathbb{C}$ is defined by

$$\begin{aligned} \langle \nabla_{\dot{\mathbf{c}}_t}\boldsymbol{\omega}, \mathbf{v}_t \rangle &= \partial_{\tau=t}\ \langle \mathbf{c}_{\tau,t}\Downarrow \boldsymbol{\omega}(\mathbf{c}(\tau)), \mathbf{v}_t \rangle \\ &= \partial_{\tau=t}\ \langle \boldsymbol{\omega}(\mathbf{c}(\tau)), \mathbf{c}_{\tau,t}\Uparrow \mathbf{v}_t \rangle \,, \quad \forall\, \mathbf{v}_t \in \mathbb{T}_{\mathbf{c}(t)}\mathbb{C} \,. \end{aligned}$$

Let us consider the vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{T}\mathbb{C})$ which is the extension of the velocity $\mathbf{v}_t := \partial_{\tau=t}\,\gamma(\tau)$ of the trajectory performed by dragging it along the flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^2(\mathbb{C}\,;\mathbb{C})$:

$$\mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))) := T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t) \,, \iff \mathbf{v} := \boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t \,,$$

so that $\mathbf{v}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) = \mathbf{v}_t$. Setting $\boldsymbol{\varphi}_\lambda\Uparrow := \boldsymbol{\varphi}_{0,\lambda}\Uparrow$ we have that

$$T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t) = \boldsymbol{\varphi}_\lambda\Uparrow \boldsymbol{\varphi}_\lambda\Downarrow T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t) = \boldsymbol{\varphi}_\lambda\Uparrow \boldsymbol{\varphi}_\lambda\Downarrow \mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))) \,.$$

- The *base derivative* of a functional $f \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,;\Re)$ at $\mathbf{v} \in \mathbb{T}\mathbb{C}$ along a vector $\mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}$ is defined by:

$$\langle d_{\mathrm{B}}f(\mathbf{v}), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})) \rangle := \partial_{\lambda=0}\, f(\boldsymbol{\varphi}_\lambda\Uparrow \mathbf{v}) \,.$$

The definition is well-posed since the r.h.s. depends linearly on $\mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbb{C}$ for any fixed $\mathbf{v} \in \mathbb{T}\mathbb{C}$.

The base derivative provides the rate of change of $f \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,;\Re)$ when the base point $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}) \in \mathbb{C}$ is dragged by the flow while the velocity $\mathbf{v} \in \mathbb{T}\mathbb{C}$ is parallel transported along the flow.

Let $\mathrm{TORS}(\mathbf{v},\mathbf{u}) = \nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v} - [\mathbf{v},\mathbf{u}] \in \mathbb{T}\mathbb{C}$ be the evaluation of the torsion of the connection ∇ on the pair $\mathbf{v},\mathbf{u} \in \mathbb{T}\mathbb{C}$.

The next statement provides the form taken by the law of dynamics in terms of a connection in the configuration manifold.

Proposition 2.3.1 (The law of dynamics with a connection) *In terms of a linear connection* ∇ *on the configuration manifold* $\mathbb{C}$, *the differential law of dynamics*

$$\partial_{\lambda=0}\, L_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)) = \partial_{\tau=t}\ \langle d_{\mathrm{F}}L_\tau(\mathbf{v}_\tau), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau)) \rangle \,,$$

takes the form

$$\langle\, \partial_{\tau=t}\ d_{\mathrm{F}} L_\tau(\mathbf{v}_t) + \nabla_{\mathbf{v}_t}(d_{\mathrm{F}} L_t \circ \mathbf{v}_t) - d_{\mathrm{B}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle \\ = \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathrm{TORS}(\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_t)(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle\,,$$

or

$$\partial_{\tau=t}\ \langle\, d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \gamma_{\tau,t}\!\Uparrow \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle - \langle\, d_{\mathrm{B}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle \\ = \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathrm{TORS}(\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_t)(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle\,,$$

for any virtual velocity field $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ *which is an admissible infinitesimal isometry at the configuration* $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)$ *.*

Proof. Being

$$\partial_{\lambda=0}\, L_t(T\boldsymbol{\varphi}_\lambda(\mathbf{v}_t)) = \partial_{\lambda=0}\, L_t(\mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)))) \\ = \partial_{\lambda=0}\, L_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \boldsymbol{\varphi}_\lambda\!\Downarrow \mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))))\,,$$

by Leibniz rule, we get:

$$\partial_{\lambda=0}\, L_t(\mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)))) = \partial_{\lambda=0}\, L_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}_t) + \partial_{\lambda=0}\, L_t(\boldsymbol{\varphi}_\lambda\!\Downarrow \mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))))\,.$$

By definition of the covariant derivative in terms of the parallel transport:

$$\nabla_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{v}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) := \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda\!\Downarrow \mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)))\,,$$

being $\boldsymbol{\varphi}_\lambda\!\Downarrow \mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\mathbb{C}$, we have that

$$\partial_{\lambda=0}\, L_t(\boldsymbol{\varphi}_\lambda\!\Downarrow \mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)))) = \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), \nabla_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{v}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle\,.$$

Hence, by definition of the *base derivative* $d_{\mathrm{B}} L_t(\mathbf{v}_t) \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{T}^*\mathbb{C})$, we get

$$\partial_{\lambda=0}\, L_t(\mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)))) = \langle\, d_{\mathrm{B}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle \\ + \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), \nabla_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{v}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle\,.$$

On the other hand, denoting by $\gamma_{\tau,t} := \gamma_\tau \circ \gamma_t^{-1} \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{C})$ the flow along the trajectory, we may write

$$\partial_{\tau=t}\ \langle\, d_{\mathrm{F}} \mathrm{L}_\tau(\mathbf{v}_\tau), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau))\,\rangle = \partial_{\tau=t}\ \langle\, d_{\mathrm{F}} \mathrm{L}_\tau(\mathbf{v}_\tau), \gamma_{\tau,t}\!\Uparrow \gamma_{\tau,t}\!\Downarrow \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau))\,\rangle\,,$$

and applying the LEIBNIZ rule:

$$\partial_{\tau=t} \, \langle d_{\mathrm{F}} \mathrm{L}_{\tau}(\mathbf{v}_{\tau}), \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_{\tau})) \rangle = \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \partial_{\tau=t} \, \gamma_{\tau,t}\Downarrow \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_{\tau})) \rangle + \partial_{\tau=t} \, \langle d_{\mathrm{F}} \mathrm{L}_{\tau}(\mathbf{v}_{\tau}), \gamma_{\tau,t}\Uparrow \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle \,.$$

Finally, by definition of the covariant derivatives, we have:

$$\langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \partial_{\tau=t} \, \gamma_{\tau,t}\Downarrow \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_{\tau})) \rangle = \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \nabla_{\mathbf{v}} \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle \,,$$

$$\partial_{\tau=t} \, \langle d_{\mathrm{F}} \mathrm{L}_{\tau}(\mathbf{v}_{\tau}), \gamma_{\tau,t}\Uparrow \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle = \langle \partial_{\tau=t} \, d_{\mathrm{F}} \mathrm{L}_{\tau}(\mathbf{v}_t), \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle + \langle \nabla_{\mathbf{v}_t}(d_{\mathrm{F}} L_t \circ \mathbf{v}_t), \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle \,.$$

The law of dynamics may then be written as

$$\langle \partial_{\tau=t} \, d_{\mathrm{F}} \mathrm{L}_{\tau}(\mathbf{v}_t) + \nabla_{\mathbf{v}_t}(d_{\mathrm{F}} L_t \circ \mathbf{v}_t), \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle + \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \nabla_{\mathbf{v}} \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle$$
$$= \langle d_{\mathrm{B}} L_t(\mathbf{v}_t), \mathbf{v}_{\varphi}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle + \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \nabla_{\mathbf{v}_{\varphi}} \mathbf{v}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle \,.$$

Recalling that $\mathrm{TORS}(\mathbf{v}_{\varphi}, \mathbf{v}_t) := \nabla_{\mathbf{v}_{\varphi}} \mathbf{v} - \nabla_{\mathbf{v}} \mathbf{v}_{\varphi} - [\mathbf{v}_{\varphi}, \mathbf{v}]$, and observing that $[\mathbf{v}_{\varphi}, \mathbf{v}] = 0$ by definition of the vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{C}\,; \mathbb{TC})$, the statement is proven. ■

Remark 2.3.1 *The form taken by the law of dynamics in a configuration manifold endowed with a connection hides the direct implication of* NOETHER*'s theorem. Indeed, to recover the general law of dynamics, one should be able to follow backwards the steps in the proof of proposition 2.3.1 and this is a rather involved path of reasoning to be envisaged.*

2.3.1 Poincare's law of dynamics

A connection on the configuration manifold is induced by a local frame by defining as distant parallel transport the one that leaves invariant the components of a vector in the moving frames (*repére mobile*) while changing the base point. This connection has vanishing curvature and the torsion evaluated on any pair of vectors $\mathbf{u}_{\mathbf{x}}, \mathbf{v}_{\mathbf{x}} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}$ is the negative of the LIE brackets of their extensions by distant parallel transport $\mathbf{u}, \mathbf{v} \in \mathbb{TC}$ such that $\mathbf{u}(\mathbf{x}) = \mathbf{u}_{\mathbf{x}}$ Indeed

$$\mathrm{TORS}(\mathbf{u}_{\mathbf{x}}, \mathbf{v}_{\mathbf{x}}) := \nabla_{\mathbf{u}_{\mathbf{x}}} \mathbf{v} - \nabla_{\mathbf{v}_{\mathbf{x}}} \mathbf{u} - [\mathbf{u}, \mathbf{v}](\mathbf{x}) = -[\mathbf{u}, \mathbf{v}](\mathbf{x}) \,,$$

being $\nabla_{\mathbf{u_x}} \mathbf{v} = \nabla_{\mathbf{v_x}} \mathbf{u} = 0$. Then the Lie bracket $[\mathbf{u}, \mathbf{v}]$ is tensorial and the general formula for the law of dynamics gives:

$$\begin{aligned}\langle \partial_{\tau=t}\, d_{\mathrm{F}} \mathrm{L}_\tau(\mathbf{v}_t) + \nabla_{\mathbf{v}_t}(d_{\mathrm{F}} L_t \circ \mathbf{v}_t) = {} & d_{\mathrm{B}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle \\ & + \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), [\mathbf{v}_t\,, \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))] \rangle\,,\end{aligned}$$

which is the coordinate-free version of the law of dynamics found by Poincaré in 1901, see [?].

2.3.2 Lagrange's law of dynamics

If the connection ∇ is torsion-free, the differential law of dynamics takes the form of Lagrange's differential condition:

$$\langle \partial_{\tau=t}\, d_{\mathrm{F}} \mathrm{L}_\tau(\mathbf{v}_t) + \nabla_{\mathbf{v}_t}(d_{\mathrm{F}} L_t \circ \mathbf{v}_t) - d_{\mathrm{B}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle = 0\,.$$

The Lagrange differential condition holds *a fortiori* in any configuration manifold $\mathbb{C}$ which is a riemannian manifold with the Levi-Civita connection, which is metric-preserving and torsion-free. In particular it holds in any linear configuration manifold $\mathbb{C}$ with the canonical connection by translation and also in the configuration manifold $\mathbb{C}$ when the local connection is induced by a chart (*repére naturel*), since these connections are torsion-free (and curvature-free).

In rigid-body dynamics, or more in general in a perfect dynamical system, the test velocity may be omitted. Then, in a riemannian configuration manifold, we have the following Lagrange*'s equation of perfect dynamics*:

$$\partial_{\tau=t}\, d_{\mathrm{F}} \mathrm{L}_\tau(\mathbf{v}_t) + \nabla_{\mathbf{v}_t}(d_{\mathrm{F}} L_t \circ \mathbf{v}_t) = d_{\mathrm{B}} L_t(\mathbf{v}_t)\,.$$

2.3.3 Hamilton's law of dynamics

Hamilton's law of dynamics is deduced from Lagrange's law by a translation in terms of covectors $\mathbf{v}^* \in \mathbb{T}^*\mathbb{C}$ by means of Legendre's transform. We assume that $L_t \in \mathrm{C}^2(\mathbb{C}\,;\Re)$ is a *regular Lagrangian*, which means that the fiber derivative $d_{\mathrm{F}} L_t \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{T}^*\mathbb{C})$ is a vector bundle isomorphisms. In fact the projected base map, defined by the commutative diagram:

$$\begin{array}{ccc} \mathbb{TC} & \xrightarrow{d_{\mathrm{F}} L_t} & \mathbb{T}^*\mathbb{C} \\ {\scriptstyle \boldsymbol{\tau}_{\mathbb{C}}}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\tau}^*_{\mathbb{C}}} \\ \mathbb{C} & \xrightarrow{\mathbf{id}_{\mathbb{C}}} & \mathbb{C} \end{array} \iff \boldsymbol{\tau}^*_{\mathbb{C}} \circ d_{\mathrm{F}} L_t = \mathbf{id}_{\mathbb{C}} \circ \boldsymbol{\tau}_{\mathbb{C}} = \boldsymbol{\tau}_{\mathbb{C}} \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{C})\,.$$

is the identity on $\mathbb{C}$ and then invertible. The assumption is thus equivalent to require that the fiber derivative is fiberwise bounded and linear, with a bounded linear inverse.

The Hamiltonian $H_t \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\Re)$ is fiberwise defined as the potential of the inverse map $(d_{\mathrm{F}} L_t)^{-1} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}\mathbb{C})$ with the additive constant fixed by the Legendre transformation rule:

$$L_t(\mathbf{v}) + H_t(\mathbf{v}^*) = \langle \mathbf{v}^*, \mathbf{v} \rangle\,, \qquad \begin{cases} \mathbf{v} \;\, = d_{\mathrm{F}} H_t(\mathbf{v}^*) \in \mathbb{T}\mathbb{C}\,, \\ \mathbf{v}^* = d_{\mathrm{F}} L_t(\mathbf{v}) \quad \in \mathbb{T}^*\mathbb{C}\,. \end{cases}$$

The following proposition yields the basic result for the formulation of the canonical Hamilton's law of dynamics. The special case in linear spaces is referred to as Donkin's theorem (1854) in [59].

Proposition 2.3.2 (Base derivatives of Legendre transforms) *In a manifold $\mathbb{C}$ endowed with an affine connection ∇ the following relation holds:*

$$d_{\mathrm{B}} H_t(\mathbf{v}^*) + d_{\mathrm{B}} L_t(d_{\mathrm{F}} H_t(\mathbf{v}^*)) = 0\,, \quad \forall\, \mathbf{v}_t^* \in \mathbb{T}^*\mathbb{C}\,,$$

that is $d_{\mathrm{B}} H_t + d_{\mathrm{B}} L_t \circ d_{\mathrm{F}} H_t = 0$.

Proof. By Legendre transform we have that

$$H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*) + L_t(d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*)) = \langle \boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*, d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*) \rangle\,.$$

Taking the derivative at $\lambda = 0$, we get:

$$\partial_{\lambda=0}\, H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*) = \langle d_{\mathrm{B}} H_t(\mathbf{v}^*), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}_t^*)) \rangle\,.$$

Moreover, by Leibniz rule and the identity $d_{\mathrm{F}} L_t(d_{\mathrm{F}} H_t(\mathbf{v}^*)) = \mathbf{v}^*$, observing that $\boldsymbol{\varphi}_\lambda\!\Downarrow d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*) \in \mathbb{T}_{\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}^*)}\mathbb{C}$ for any $\lambda \in \Re$, we have:

$$\begin{aligned} \partial_{\lambda=0}\, L_t(d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*)) &= \partial_{\lambda=0}\, L_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \boldsymbol{\varphi}_\lambda\!\Downarrow d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*)) \\ &= \langle d_{\mathrm{B}} L_t(d_{\mathrm{F}} H_t(\mathbf{v}^*)), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}^*)) \rangle \\ &\quad + \langle d_{\mathrm{F}} L_t(d_{\mathrm{F}} H_t(\mathbf{v}^*)), \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda\!\Downarrow d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*) \rangle \\ &= \langle d_{\mathrm{B}} L_t(d_{\mathrm{F}} H_t(\mathbf{v}^*)), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}^*(\mathbf{v}^*)) \rangle \\ &\quad + \langle \mathbf{v}^*, \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda\!\Downarrow d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda\!\Uparrow \mathbf{v}^*) \rangle\,. \end{aligned}$$

Since the parallel transport preserves the duality pairing, we infer that

$$\partial_{\lambda=0} \langle \boldsymbol{\varphi}_\lambda \Uparrow \mathbf{v}^*, d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda \Uparrow \mathbf{v}^*) \rangle = \partial_{\lambda=0} \langle \mathbf{v}^*, \boldsymbol{\varphi}_\lambda \Downarrow d_{\mathrm{F}} H_t(\boldsymbol{\varphi}_\lambda \Uparrow \mathbf{v}^*) \rangle .$$

In conclusion:

$$\langle d_{\mathrm{B}} H_t(\mathbf{v}^*), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*)) \rangle + \langle d_{\mathrm{B}} L_t(d_{\mathrm{F}} H_t(\mathbf{v}^*)), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*)) \rangle = 0 ,$$

for any vector $\mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*)) \in \mathbb{T}_{\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*)}\mathbb{C}$, and the result is proven. ■

From propositions 2.3.1 and 2.3.2 we then get:

Proposition 2.3.3 (Hamilton's canonical equations) *If the configuration manifold* $\mathbb{C}$ *is endowed with an affine connection* ∇, *the differential law of dynamics takes the form*

$$\begin{cases} \langle \partial_{\tau=t}\ \mathbf{v}^*_\tau + \nabla_{\mathbf{v}_t}\mathbf{v}^* + d_{\mathrm{B}} H_t(\mathbf{v}^*_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)) \rangle = \langle \mathbf{v}^*_t, \mathrm{TORS}(\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_t)(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)) \rangle , \\ \mathbf{v}_t = d_{\mathrm{F}} H_t(\mathbf{v}^*_t) . \end{cases}$$

If the connection ∇ *is torsion-free, the differential law of dynamics takes the form of* HAMILTON*'s canonical equations:*

$$\begin{cases} \langle \partial_{\tau=t}\ \mathbf{v}^*_\tau + \nabla_{\mathbf{v}_t}\mathbf{v}^* + d_{\mathrm{B}} H_t(\mathbf{v}^*_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*_t)) \rangle = 0 , \\ \mathbf{v}_t = d_{\mathrm{F}} H_t(\mathbf{v}^*_t) , \end{cases}$$

for any virtual velocity field $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{C}\,;\mathbb{TC})$ *which is an admissible infinitesimal isometry at each point of* γ, *and* HAMILTON*'s canonical equations of perfect dynamics are given by:*

$$\begin{cases} \partial_{\tau=t}\ \mathbf{v}^*_\tau + \nabla_{\mathbf{v}_t}\mathbf{v}^* = -d_{\mathrm{B}} H_t(\mathbf{v}^*_t) , \\ \mathbf{v}_t = d_{\mathrm{F}} H_t(\mathbf{v}^*_t) . \end{cases}$$

2.3.4 Action principle in the covelocity-time state-space

The invariance property stated in proposition 2.1.1 shows that HAMILTON's canonical laws of dynamics may be deduced by translating the action principle from the velocity-time state-space into the covelocity-time state-space, by means of the Legendrian functor.

To this end we introduce the one-form $\boldsymbol{\theta} = d_{\mathrm{F}} L_t {\uparrow} \boldsymbol{\theta}_L \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{T}^*\mathbb{C})$ as the LEGENDRE transformed of $\boldsymbol{\theta}_L \in \mathrm{C}^1(\mathbb{TC}\,;\mathbb{T}^*\mathbb{TC})$:

$$\boldsymbol{\theta}(d_{\mathrm{F}} L_t(\mathbf{v})) \cdot d_{\mathrm{F}} L_t {\uparrow} \delta\mathbf{v} := \boldsymbol{\theta}_L(\mathbf{v}) \cdot \delta\mathbf{v} , \quad \forall\, \mathbf{v} \in \mathbb{TC} , \quad \forall\, \delta\mathbf{v} \in \mathbb{T}_{\mathbf{v}}\mathbb{TC} ,$$

where $d_{\rm F}L_t{\uparrow}\delta\mathbf{v} \in \mathbb{T}_{d_{\rm F}L_t(\mathbf{v})}\mathbb{T}^*\mathbb{C}$. Then, being $\boldsymbol{\tau}^*_{\mathbb{C}} = \boldsymbol{\tau}_{\mathbb{C}} \circ d_{\rm F}H_t$, we have that

$$\begin{aligned}\boldsymbol{\theta}(\mathbf{v}^*) \cdot \delta\mathbf{v}^* &= \boldsymbol{\theta}_L(d_{\rm F}H_t(\mathbf{v}^*)) \cdot d_{\rm F}H_t{\uparrow}\delta\mathbf{v}^* \\ &= \langle\, d_{\rm F}L_t(d_{\rm F}H_t(\mathbf{v}^*)), T\boldsymbol{\tau}_{\mathbb{C}}(d_{\rm F}H_t(\mathbf{v}^*)) \cdot d_{\rm F}H_t{\uparrow}\delta\mathbf{v}^* \,\rangle \\ &= \langle\, \mathbf{v}^*, T\boldsymbol{\tau}^*_{\mathbb{C}}(\mathbf{v}^*) \cdot \delta\mathbf{v}^* \,\rangle = \langle\, \mathbf{v}^*, T\boldsymbol{\tau}^*_{\mathbb{C}} \circ \delta\mathbf{v}^* \,\rangle\,.\end{aligned}$$

The one-form $\boldsymbol{\theta} = d_{\rm F}L_t{\uparrow}\boldsymbol{\theta}_L \in {\rm C}^1(\mathbb{T}^*\mathbb{C}\,; \mathbb{T}\mathbb{T}^*\mathbb{C})$ is then independent of the Lagrangian. Accordingly, in the covelocity-time state-space $\mathbb{T}^*\mathbb{C} \times \mathbb{T}I$ we may define the one-form: $\boldsymbol{\omega}^1 = d_{\rm F}L_t{\uparrow}\boldsymbol{\omega}^1_L \in {\rm C}^1(\mathbb{T}^*\mathbb{C} \times I\,; \mathbb{T}\mathbb{T}^*\mathbb{C} \times \mathbb{T}I)$ by

$$\boldsymbol{\omega}^1(d_{\rm F}L_t(\mathbf{v})) \cdot d_{\rm F}L{\uparrow}\delta\mathbf{v} := \boldsymbol{\omega}^1_L(\mathbf{v}) \cdot \delta\mathbf{v}\,, \quad \forall\, \mathbf{v} \in \mathbb{T}\mathbb{C}\,, \quad \forall\, \delta\mathbf{v} \in \mathbb{T}_{\mathbf{v}}\mathbb{T}\mathbb{C}\,,$$

where $d_{\rm F}L_t{\uparrow}\delta\mathbf{v} \in \mathbb{T}_{d_{\rm F}L_t(\mathbf{v})}\mathbb{T}^*\mathbb{C}$, so that

$$\boldsymbol{\omega}^1((\mathbf{v}^*\,, t)) = \boldsymbol{\theta}(\mathbf{v}^*) - H(\mathbf{v}^*, t)dt\,, \quad \forall\, \mathbf{v}^* \in \mathbb{T}^*\mathbb{C}\,.$$

Hamilton's action principle in the velocity-time state-space:

$$\partial_{\lambda=0} \int_{T\boldsymbol{\varphi}_\lambda(\gamma)} \boldsymbol{\omega}^1_L = \oint_{\partial\gamma} \boldsymbol{\omega}^1_L \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}\,, 0)\,,$$

may then be rewritten for the trajectory $d_{\rm F}L_t \circ \gamma \in {\rm C}^1(\mathbb{T}I\,; \mathbb{T}^*\mathbb{C} \times \mathbb{T}I)$ in the covelocity-time state-space, as the variational condition:

$$\partial_{\lambda=0} \int_{d_{\rm F}L_t(T\boldsymbol{\varphi}_\lambda(\gamma))} \boldsymbol{\omega}^1 = \int_{\partial d_{\rm F}L_t(\gamma)} \boldsymbol{\omega}^1 \cdot (\mathbf{v}_{{\rm LEG}(T\boldsymbol{\varphi})}\,, 0)\,,$$

for any flow $\boldsymbol{\varphi}_\lambda \in {\rm C}^1(\mathbb{C}\,; \mathbb{C})$ such that the velocity field $\mathbf{v}_{\boldsymbol{\varphi}} \in {\rm C}^1(\mathbb{C}\,; \mathbb{T}\mathbb{C})$ is an infinitesimal isometry of $\gamma \in \mathbb{C}$.

Localizing, the differential condition reads:

$$d\boldsymbol{\theta}(\mathbf{v}^*_t) \cdot \mathbf{X}(\mathbf{v}^*_t) \cdot \mathbf{v}_{{\rm LEG}(T\boldsymbol{\varphi})}(\mathbf{v}^*_t) = -\langle\, dH_t(\mathbf{v}^*_t), \mathbf{v}_{{\rm LEG}(T\boldsymbol{\varphi})}(\mathbf{v}^*_t) \,\rangle\,,$$

with $\mathbf{v}_{{\rm LEG}(T\boldsymbol{\varphi})} = d_{\rm F}L_t{\uparrow}\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}^*_t) = d_{\rm F}L_t{\uparrow}(\mathbf{k}_{\mathbb{C}} \circ T\mathbf{v}_{\boldsymbol{\varphi}})(\mathbf{v}^*_t) \in \mathbb{T}_{\mathbf{v}^*_t}\mathbb{T}^*\mathbb{C}$, for all $\mathbf{v}_{\boldsymbol{\varphi}} \in {\rm C}^1(\mathbb{C}\,; \mathbb{T}\mathbb{C})$ which is an infinitesimal isometry of $\gamma \in \mathbb{C}$.

2.4 Perfect dynamics

In the context of perfect dynamics, as defined in section 2.1.8 on page 254, some general qualitative properties are available. Some classical results will be illustrated in the sequel on the basis of the previous analysis.

2.4.1 Integral invariants

In the next section it is shown that Hamilton's canonical equations of dynamics may be equivalently enunciated as an invariance property. This invariance property is the key which leads to the definition and the investigations on canonical transformations, which Jacobi has applied as a very effective tool for the closed form solution of problems in perfect dynamics.

2.4.2 Poincaré-Cartan integral invariant

Let us give a preliminary definition.

- The flow associated with the vector field $\mathbf{X}_t^H \in C^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}\mathbb{T}^*\mathbb{C})$ solution of the Hamilton's equation is called the *phase-flow* $\mathbf{Fl}_{\tau,t}^H \in C^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ associated with the hamiltonian $H_t \in C^1(\mathbb{T}^*\mathbb{C}\,;\Re)$.

From the abstract result of proposition 2.1.3 we infer the following statement.

Theorem 2.4.1 (Poincaré-Cartan integral invariant) *In the covelocity-time state-space $\mathbb{T}^*\mathbb{C} \times I$, the integral of the action one-form $\boldsymbol{\omega}^1$ around any loop surrounding a given ray-tube, is invariant, i.e.:*

$$\oint_{\mathbf{l}_1} \boldsymbol{\omega}^1 = \oint_{\mathbf{l}_2} \boldsymbol{\omega}^1 \,,$$

for any two such loops $\mathbf{l}_1, \mathbf{l}_2$.

From theorem 2.1.4 we infer that

- the invariance of the Poincaré-Cartan integral is equivalent to Hamilton's equations.

2.4.3 Poincaré relative integral invariant

By projecting on the covelocity-phase-space, we get the following classical result due to Poincaré.

Theorem 2.4.2 (Poincaré relative integral invariant) *The integral of the canonical one-form $\boldsymbol{\theta}$ around any loop $\mathbf{l} \in \mathbb{T}^*\mathbb{C}$ in the covelocity-phase-space is invariant under the action of the phase-flow $\mathbf{Fl}_{\tau,t}^H \in C^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ associated with any hamiltonian $H_t \in C^1(\mathbb{T}^*\mathbb{C}\,;\Re)$:*

$$\oint_{\mathbf{l}} \boldsymbol{\theta} = \oint_{\mathbf{Fl}_{\tau,t}^H(\mathbf{l})} \boldsymbol{\theta} \,,$$

Proof. A closed loop in the covelocity-phase-space $\mathbb{T}^*\mathbb{C}$ can be seen as the projection of a loop surrounding a ray-tube at a fixed time, so that the hamiltonian one-form

$$\boldsymbol{\omega}^1(\mathbf{v}^*, t) := \boldsymbol{\theta}(\mathbf{v}^*) - H_t(\mathbf{v}^*)\, dt \in \mathbb{T}^*_{(\mathbf{v}^*,t)}(\mathbb{T}^*\mathbb{C} \times I)\,,$$

reduces to $\boldsymbol{\omega}^1 = \boldsymbol{\theta}$. ■

A differential k-form is called a *relative integral invariant* of a phase-flow if its integral on any closed k-chain is invariant under the action of the phase-flow. A differential k-form whose integral on any k-chain is invariant under the action of a phase-flow is said to be an *absolute integral invariant* of the phase-flow.

Poincaré relative integral invariant is a *universal integral invariant* since the invariance property is independent of the hamiltonian and hence holds for any phase-flow.

Theorem 2.4.3 *The canonical two-form* $\boldsymbol{\omega}^2 = -d\boldsymbol{\theta}$ *is an absolute universal integral invariant and its* Lie *derivative along any phase-flow vanishes.*

Proof. The boundary $\partial\mathbf{c}^2$ of any 2-chain $\mathbf{c}^2$ in $\mathbb{T}^*\mathbb{C}$ is closed since $\partial\partial\mathbf{c}^2 = 0$. Then from theorem 2.4.2 and Stokes formula, it follows that, for any phase-flow:

$$\int_{\mathbf{c}^2} \boldsymbol{\omega}^2 = -\oint_{\partial\mathbf{c}^2} \boldsymbol{\theta} = -\oint_{\mathbf{Fl}^H_{\tau,t}(\partial\mathbf{c}^2)} \boldsymbol{\theta} = -\oint_{\partial\mathbf{Fl}^H_{\tau,t}(\mathbf{c}^2)} \boldsymbol{\theta} = \int_{\mathbf{Fl}^H_{\tau,t}(\mathbf{c}^2)} \boldsymbol{\omega}^2\,.$$

By the arbitrarity of the 2-chain, Reynolds' transport formula implies that

$$\partial_{\tau=t} \int_{\mathbf{Fl}^H_{\tau,t}(\mathbf{c}^2)} \boldsymbol{\omega}^2 = \int_{\mathbf{c}^2} \mathcal{L}_{\mathbf{X}^H_t}\, \boldsymbol{\omega}^2 = 0\,, \quad \forall\, \mathbf{c}^2 \iff \mathcal{L}_{\mathbf{X}^H_t}\, \boldsymbol{\omega}^2 = 0\,,$$

and the result is proven. ■

2.4.4 Canonical transformations

We are led to give the following definition.

- A *canonical flow* $\mathbf{Fl}^{\mathbf{X}}_{\tau,t} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ is a flow in the covelocity-phase-space, with associated vector field $\mathbf{X} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}\mathbb{T}^*\mathbb{C})$, which drags the canonical two-form $\boldsymbol{\omega}^2$:

$$\mathcal{L}_{\mathbf{X}_t}\, \boldsymbol{\omega}^2 = 0 \quad \text{or equivalently} \quad F_{\tau,t}{\downarrow}\boldsymbol{\omega}^2 = \boldsymbol{\omega}^2\,.$$

The result in theorem 2.4.3 may then be expressed by stating that

- A hamiltonian phase-flow is a canonical flow.

A basic characterization of canonical transformations is provided by the next proposition.

Theorem 2.4.4 (Canonical transformation of Hamilton's equations) *A transformation is canonical iff it preserves* HAMILTON*'s canonical equations, in the sense that the corresponding pull back yields the same* HAMILTON*'s equations in which both the vector field and the Hamiltonian are pulled back, i.e.*

$$\boldsymbol{\omega}^2 \cdot \mathbf{X}_t = dH_t \iff \boldsymbol{\omega}^2 \cdot \boldsymbol{\varphi}\!\downarrow\!\mathbf{X}_t = d(\boldsymbol{\varphi}\!\downarrow\! H_t)\,.$$

Proof. Let us consider the HAMILTON's equations

$$\boldsymbol{\omega}^2 \cdot \mathbf{X}_t = dH_t\,,$$

rewritten in the form

$$\boldsymbol{\omega}^2 \cdot \mathbf{X}_t \cdot \mathbf{Y} = dH_t \cdot \mathbf{Y}\,, \quad \forall\, \mathbf{Y} \in \mathbb{T}^*\mathbb{C}\,,$$

Performing a pull back by a diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$:

$$\boldsymbol{\varphi}\!\downarrow\!(\boldsymbol{\omega}^2 \cdot \mathbf{X}_t \cdot \mathbf{Y}) = \boldsymbol{\varphi}\!\downarrow\!(dH_t \cdot \mathbf{Y})\,,$$

being

$$\begin{aligned}\boldsymbol{\varphi}\!\downarrow\!(\boldsymbol{\omega}^2 \cdot \mathbf{X} \cdot \mathbf{Y}) &= \boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}^2 \cdot \boldsymbol{\varphi}\!\downarrow\!\mathbf{X}_t \cdot \boldsymbol{\varphi}\!\downarrow\!\mathbf{Y}\,,\\ \boldsymbol{\varphi}\!\downarrow\!(dH_t \cdot \mathbf{Y}) &= \boldsymbol{\varphi}\!\downarrow\!(dH_t) \cdot \boldsymbol{\varphi}\!\downarrow\!\mathbf{Y} = d(\boldsymbol{\varphi}\!\downarrow\! H_t) \cdot \boldsymbol{\varphi}\!\downarrow\!\mathbf{Y}\,,\end{aligned}$$

HAMILTON's equations are transformed into:

$$\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}^2 \cdot \boldsymbol{\varphi}\!\downarrow\!\mathbf{X}_t = d(\boldsymbol{\varphi}\!\downarrow\! H_t)\,.$$

It is then apparent that the hamiltonian structure is preserved if and only if:

$$\boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}^2 = \boldsymbol{\omega}^2\,,$$

that is if the diffeomorphism $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{T}^*\mathbb{S}\,;\mathbb{T}^*\mathbb{S})$ is canonical. ■

2.4.5 Lee Hwa-Chung theorem

The exterior product is natural with respect to a push, i.e.:

$$\boldsymbol{\varphi}\!\downarrow\!(\boldsymbol{\omega}^2 \wedge \boldsymbol{\omega}^2) = \boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}^2 \wedge \boldsymbol{\varphi}\!\downarrow\!\boldsymbol{\omega}^2\,,$$

and hence all exterior powers of $\boldsymbol{\omega}^2$ are dragged by a canonical transformation.

If the configuration manifold is n-dimensional, we get n absolute integral invariants of order $2k$ by taking the integrals of $\{\boldsymbol{\omega}^{2k}\}$ for $k = 1, \dots, n$ over $2k$-dimensional submanifolds and n corresponding relative integral invariants of order k integrating along the boundaries of such manifolds.

These invariants are universal integral invariants since the one-form $\boldsymbol{\theta}$ and the two-form $\boldsymbol{\omega}^2$, and hence the invariance property, do not depend on the particular hamiltonian flow considered.

In 1947 the chinese scientist Lee Hwa-Chung proved the uniqueness of these universal integral invariants [59]. For $k = 1$ his theorem can be stated as follows.

Theorem 2.4.5 (Lee Hwa-Chung theorem) *All* 1*-th order universal relative integral invariants are proportional to* Poincaré *integral invariant.*

2.4.6 Liouville's theorem

If the configuration manifold $\mathbb{C}$ is n-dimensional, the n-th power of $\boldsymbol{\omega}^2$ is a volume $2n$-form on the $2n$-dimensional cotangent bundle $\mathbb{T}^*\mathbb{C}$. Hence we get the following classical result.

Theorem 2.4.6 (Liouville theorem) *The volume of the covelocity-phase-space* $\mathbb{T}^*\mathbb{C}$ *is invariant under the action of a canonical transformation.*

Figure 2.5: Joseph Liouville (1809 - 1882)

A classical application of Liouville's theorem is to ergodic theory.

A nice example of a qualitative description of the properties of motion is provided by the following proposition due to Poincaré. We will denote by $\boldsymbol{\varphi}^k(\mathbf{w}^*)$ the image of $\mathbf{w}^* \in \mathbb{T}^*\mathbb{C}$ thru the k-th iterate of the map $\boldsymbol{\varphi} \in C^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$.

Theorem 2.4.7 (Return theorem) *Let* $\boldsymbol{\varphi} \in C^1(\mathbb{T}^*\mathbb{C}\,;\mathbb{T}^*\mathbb{C})$ *be a volume preserving diffeomophism which maps a bounded open submanifold into itself. Then, given a point* $\mathbf{v}^* \in \mathbb{T}^*\mathbb{C}$ *and a neighbourhood* $U(\mathbf{v}^*)$ *, there exists a point* $\mathbf{w}^* \in U(\mathbf{v}^*)$ *such that* $\boldsymbol{\varphi}^k(\mathbf{w}^*) \in U(\mathbf{v}^*)$ *for some* $k > 0$ *.*

Proof. Let us set $U = U(\mathbf{v}^*)$ for convenience. All the images $\boldsymbol{\varphi}^h(U)$ for any $h \geq 0$ have the same volume by assumption. Then the boundedness of the submanifold requires that $\boldsymbol{\varphi}^h(U) \cap \boldsymbol{\varphi}^k(U) \neq \emptyset$ for some $h > k > 0$, so that $\boldsymbol{\varphi}^{(h-k)}(U) \cap U \neq \emptyset$.

2.4.7 Hamilton-Jacobi equation

The description of the law of dynamics expressed by Hamilton-Jacobi equation stands to the action principle and to the related Euler stationarity condition as Huygens picture of geometrical optics stands to Fermat's least time principle and to the geodesic stationarity condition. The two approach consist respectively in describing the characterisitc property of the ray or trajectories in the state space on one hand, and the evolution of the propagation fronts as hypersurphaces in the state space, on the other.

Figure 2.6: Christiaan Huygens (1629 - 1695)

Let $\gamma \in C^1(I\,;\mathbb{C})$ be the trajectory, in the time interval I, of a dynamical system in the configuration manifold. In the configuration-time state-space the corresponding trajectory is $\gamma := \gamma \times \mathbf{id}_I \in C^1(I\,;\mathbb{C} \times I)$.

- A trajectory between the points $\{\mathbf{x}_0, t_0\} \in \mathbb{C} \times I$ and $\{\mathbf{x}, t\} \in \mathbb{C} \times I$ of the configuration-time state-space, is said to belong to a *central field* if for any $(\boldsymbol{\xi}\,, \tau) \in U_{\mathbf{x}} \times U_t$, with $U_{\mathbf{x}} \times U_t$ open submanifold of $\mathbb{C} \times I$ containing

$(\mathbf{x}\,,t)$, there exists a unique trajectory carrying $(\mathbf{x}_0\,,t_0)$ to $(\boldsymbol{\xi}\,,\tau)$. This assumption is fulfilled if the time interval $(t_0\,,t)$ is sufficiently small [6].

Each trajectory $\gamma \in \mathrm{C}^1(I\,;\mathbb{C}\times I)$ of the central field in the configuration-time state-space, is lifted to a phase-trajectory $\mathbf{\Gamma} := T\gamma \in \mathrm{C}^1(\mathbb{T}I\,;\mathbb{T}\mathbb{C}\times\mathbb{T}I)$ in the velocity-time state-space.

The LEGENDRE transform maps the trajectory $\mathbf{\Gamma} \in \mathrm{C}^1(\mathbb{T}I\,;\mathbb{T}\mathbb{C}\times\mathbb{T}I)$, into a trajectory $\mathbf{\Gamma}^* = d_{\mathrm{F}}L_t \circ \mathbf{\Gamma} \in \mathrm{C}^1(\mathbb{T}I\,;\mathbb{T}^*\mathbb{C}\times\mathbb{T}I)$, in the covelocity-time state-space. Recalling that

$$\begin{aligned}&\boldsymbol{\omega}^1((\mathbf{v}_t^*\,,t))\cdot(\dot{\mathbf{v}}_t^*\,,1_t) = \langle\,\mathbf{v}_t^*, d_{\mathrm{F}}H_t(\mathbf{v}_t^*)\,\rangle - H_t(\mathbf{v}_t^*)\langle\,dt,1_t\,\rangle\\ &=\boldsymbol{\omega}_L^1((\mathbf{v}_t\,,t))\cdot(\mathbf{X}(\mathbf{v}_t)\,,1_t) = \langle\,d_{\mathrm{F}}L_t(\mathbf{v}_t),\mathbf{v}_t\,\rangle - E_t(\mathbf{v}_t)\langle\,dt,1_t\,\rangle = \mathrm{L}_t(\mathbf{v}_t)\,,\end{aligned}$$

by the rule for change of integration domain, the action integral is given by one of the following equivalent expressions:

$$S(\mathbf{\Gamma}^*) := \int_{\mathbf{\Gamma}^*}\boldsymbol{\omega}^1 = S(\mathbf{\Gamma}) := \int_{\mathbf{\Gamma}}\boldsymbol{\omega}_L^1 = S(\gamma) := \int_I L_t(\dot{\gamma}_{0t})\,dt\,.$$

- The *eikonal functional* $J \in \mathrm{C}^1(\mathbb{C}\times I\,;\Re)$, given by $J(\mathbf{x},t) := S(\gamma)$ with $\mathbf{x} = \gamma_t \in \mathbb{C}$, is well-defined by the centrality assumption.
- The *eikonal one-form* $\mathbf{j} \in \mathrm{C}^1(\mathbb{C}\times I\,;\mathbb{T}^*_{(\mathbf{x}\,,t)}(\mathbb{C}\times I))$ is the differential one-form defined by

$$\begin{aligned}\mathbf{j}(\mathbf{x},t) &:= \mathbf{v}_t^*(\mathbf{x}) - H_t(\mathbf{v}_t^*(\mathbf{x}))\,dt\\ &= d_{\mathrm{F}}L_t(\mathbf{v}_t(\mathbf{x})) - E_t(\mathbf{v}_t(\mathbf{x}))\,dt\,,\end{aligned}$$

that is:

$$\mathbf{j}_t \circ \boldsymbol{\tau}_{\mathbb{C}} := d_{\mathrm{F}}L_t - E_t\,dt\,.$$

It is associated with a *central field* of trajectories, in the sense that the momentum $\mathbf{v}_t^* = d_{\mathrm{F}}L_t(\mathbf{v}_t) \in \mathbb{T}^*_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\mathbb{C}$ is the LEGENDRE conjugate to the velocity $\mathbf{v}_t = \dot{\gamma}_{0t} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}$ along the trajectory $\gamma \in \mathrm{C}^1(I\,;\mathbb{C})$ at the point $(\mathbf{x}\,,t)$ with $\mathbf{x} = \boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) = \gamma_t$.

Let us consider an asynchronous flow $\boldsymbol{\varphi}_\lambda \times \theta_\lambda \in \mathrm{C}^1(\mathbb{C}\times I\,;\mathbb{C}\times\mathbb{T}\Re)$ which drags the trajectory $\gamma = \gamma \times \mathbf{id}_I \in \mathrm{C}^1(I\,;\mathbb{C}\times I)$, starting at the point $(\mathbf{x}_0\,,t_0) \in \mathbb{C}\times I$ and ending at the point $(\mathbf{x}\,,t) \in \mathbb{C}\times I$, into a one-parameter family of trajectories $\gamma_\lambda \in \mathrm{C}^1(\mathbb{T}\Re\,;\mathbb{C}\times\mathbb{T}\Re)$ joining the point $(\mathbf{x}_0\,,t_0) \in \mathbb{C}\times\mathbb{T}\Re$ with the point $(\boldsymbol{\varphi}_\lambda(\mathbf{x})\,,\theta_\lambda(t)) \in \mathbb{C}\times\mathbb{T}\Re$, and defined by

$$\gamma_\lambda := \boldsymbol{\varphi}_\lambda \circ \gamma \circ \theta_\lambda^{-1}\,.$$

Then we have the following result.

Theorem 2.4.8 (Integrability of eikonal one-form) *The* ***eikonal one-form*** $\mathbf{j} \in \mathrm{C}^1(\mathbb{C} \times I\,; \mathbb{T}^*_{(\mathbf{x}\,,t)}(\mathbb{C} \times I))$ *associated with a* ***central field*** *of trajectories of a dynamical system, is locally exact and its potential is the eikonal functional:*

$$\mathbf{j} = dJ\,.$$

Proof. Let us denote by $\mathbf{\Gamma}(\lambda) \in \mathrm{C}^1(\mathbb{T}I\,; \mathbb{T}\mathbb{C} \times \mathbb{T}I)$ the trajectory which is the lifting of $\gamma(\lambda) \in \mathrm{C}^1(I\,; \mathbb{C} \times I)$ in the velocity-time phase-space, and by $(\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t)\,, \Theta(t)) \in \mathbb{T}_{\mathbf{v}_t}\mathbb{T}\mathbb{C} \times \mathbb{T}_t I$ the velocity along the asynchronous virtual flow $T\boldsymbol{\varphi}_\lambda \times \theta_\lambda \in \mathrm{C}^1(\mathbb{T}\mathbb{C} \times \mathbb{T}I\,; \mathbb{T}\mathbb{C} \times \mathbb{T}I)$. The variation of the action integral is then given by

$$\begin{aligned}\partial_{\lambda=0}\, S(\mathbf{\Gamma}(\lambda)) = \partial_{\lambda=0} \int_{\mathbf{\Gamma}(\lambda)} \boldsymbol{\omega}^1_L &= \int_{\mathbf{\Gamma}} \mathcal{L}_{(\mathbf{v}_{T\boldsymbol{\varphi}}\,,\Theta)} \boldsymbol{\omega}^1_L \\ &= \int_{\mathbf{\Gamma}} d\boldsymbol{\omega}^1_L \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}\,, \Theta) + \int_{\mathbf{\Gamma}} d(\boldsymbol{\omega}^1_L \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}\,, \Theta))\,.\end{aligned}$$

By Euler's condition for a trajectory in the velocity-time state-space:

$$d\boldsymbol{\omega}^1_L \cdot (\dot{\mathbf{v}}_t\,, 1_t) \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t)\,, \Theta) = 0\,,$$

the first integral vanishes and by Stokes formula we get:

$$\begin{aligned}\partial_{\lambda=0} \int_{\mathbf{\Gamma}(\lambda)} \boldsymbol{\omega}^1_L &= \oint_{\partial\mathbf{\Gamma}} \boldsymbol{\omega}^1_L \cdot (\mathbf{v}_{T\boldsymbol{\varphi}}\,, \Theta) \\ &= \oint_{\partial\mathbf{\Gamma}} \boldsymbol{\theta}_L(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) - E_t(\mathbf{v}_t) \cdot \Theta(t)\,.\end{aligned}$$

Being $T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))$, it is

$$\boldsymbol{\theta}_L(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) = \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), T\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \cdot \mathbf{v}_{T\boldsymbol{\varphi}}(\mathbf{v}_t) \rangle = \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle\,.$$

Hence, taking into account that the initial point $(\mathbf{x}_0\,, t_0) \in \mathbb{C} \times I$ of the trajectories is left fixed by the flow, so that $\mathbf{v}_{\boldsymbol{\varphi}}(\mathbf{x}_0, t_0) = 0$, and observing that

$$\begin{aligned}\langle dJ_t(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle &= \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle\,, \\ \partial_{\tau=t}\, J_\tau(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \cdot \Theta(t) &= -E_t(\mathbf{v}_t) \cdot \Theta(t)\,,\end{aligned}$$

we get:

$$\begin{aligned} dJ_t(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) = d_{\mathrm{F}} L_t(\mathbf{v}_t) &\iff dJ_t(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) = d_{\mathrm{F}} L_t(\mathbf{v}_t)\,, \\ \partial_{\tau=t}\, J_\tau(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) = -E_t(\mathbf{v}_t) &\iff \partial_{\tau=t}\, J_\tau \circ \boldsymbol{\tau}_{\mathbb{C}} = -E_t\,. \end{aligned}$$

Then, by the definition $\mathbf{j}_t \circ \boldsymbol{\tau}_{\mathbb{C}} := d_{\mathrm{F}} L_t - E_t dt$:

$$\mathbf{j} = dJ_t + (\partial_{\tau=t}\, J_\tau) dt\,,$$

which is the result. ■

Remark 2.4.1 *In the literature, the eikonal one-form:*

$$\mathbf{j}(\mathbf{x}, t) = \mathbf{v}_t^*(\mathbf{x}) - H_t(\mathbf{v}_t^*(\mathbf{x}))dt \in \mathbb{T}^*_{(\mathbf{x},t)}(\mathbb{C} \times I)\,,$$

is written in terms of components in a local chart $(\varphi\,, U)$ *as*

$$\mathbf{j}(q, t) = p \cdot dq - (\varphi{\uparrow}H_t)(q, p)dt\,,$$

where $q = \varphi(\mathbf{x})$ *and* $dq \subset \mathbb{T}^*\mathbb{C}$ *is the dual basis of the natural basis* $\partial q \in \mathbb{T}\mathbb{C}$ *.*

The result of theorem 2.4.8 is provided, in terms of coordinates, in [6], chapter IX, section 46 *, subsection* ***C****.*

It is to be undelined that in [6] the product $p \cdot dq$ *appears in the definition of the canonical one-form at the beginning of chapter VIII (Symplectic manifolds) and in chapter IX (Canonical formalism) section* $46 - C$ *in evaluating the differential of the action.*

Hence the expression $p \cdot dq$ *is pretended to denote the component expression of the one-form* $\mathbf{v}_t^* \in \mathbb{T}^*\mathbb{C}$ *as well as the canonical one form* $\boldsymbol{\theta} \in \mathbb{T}^*\mathbb{T}\mathbb{C}$ *. These are, however, fairly distinct objects with completely different properties.*

The one-form $\mathbf{j}(q, t) = p \cdot dq - (\varphi{\uparrow}H_t)(q, p)dt \in \mathbb{T}^*\mathbb{C} \times \mathbb{T}^*I$ *is, according to theorem 2.4.8, an exact form. On the other hand, the canonical one-form* $\boldsymbol{\theta} - H_t dt \in \mathbb{T}^*\mathbb{T}\mathbb{C} \times \mathbb{T}^*I$ *, denoted still by* $p \cdot dq - (\varphi{\uparrow}H_t)(q, p)dt$ *in [6], is not exact, its exterior derivative being the nondegenerate symplectic two-form.*

Remark 2.4.2 *From theorem 2.4.8 we infer the formula*

$$\partial_{\lambda=0}\, S(\gamma(\lambda)) = \oint_{\partial\gamma} \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle - E_t(\mathbf{v}_t)\,\Theta\,.$$

which is referred to as Schwinger*'s principle in [71]. Its equivalence with* Euler *condition can be shown by a proof analogous to the one in theorem 2.1.4.*

Figure 2.7: Julian Schwinger (1918 - 1994)

By STOKES *theorem, given any cycle* $\mathbf{c} \in \mathrm{C}^1(\Re\,; U_{\boldsymbol{\xi}} \times U_\tau)$ *, being* $\partial\mathbf{c} = 0$ *, we have that:*

$$\oint_{\mathbf{c}} \mathbf{j} = \oint_{\mathbf{c}} \mathbf{v}_t^* - H_t(\mathbf{v}_t^*)dt = \int_{\mathbf{c}} dJ = \int_{\partial\mathbf{c}} J = 0\,.$$

This property of path independence is usually stated in the form

$$\oint_{\mathbf{c}} \langle \mathbf{j}, \dot{\mathbf{c}} \rangle dt = \oint_{\mathbf{c}} (\langle \mathbf{v}_t^*, \dot{\mathbf{c}} \rangle - H_t(\mathbf{v}_t^*))dt$$
$$= \oint_{\mathbf{c}} (\langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \dot{\mathbf{c}} - \mathbf{v}_t \rangle + L_t(\mathbf{v}_t))dt = 0\,,$$

and the latter is known as HILBERT*'s path independent integral, see e.g. [176]. Further, a covector field* $\mathbf{v}^* \in \mathrm{C}^1(\mathbb{C} \times I\,; \mathbb{T}^*\mathbb{C})$ *, such that the corresponding eikonal one-form*

$$\mathbf{v}^* - H(\mathbf{v}^*, t)dt \in \mathbb{T}^*\mathbb{C} \times \mathbb{T}^* I\,,$$

is locally exact, is called a MAYER*'s field [176].*

Theorem 2.4.9 (Hamilton-Jacobi equation) *The eikonal functional* $J \in \mathrm{C}^1(U(\boldsymbol{\xi}) \times I\,; \Re)$ *fulfils the* HAMILTON-JACOBI *equation:*

$$\partial_{\tau=t}\, J_\tau(\mathbf{x}) + H_t(dJ_t(\mathbf{x})) = 0\,,$$

that is $\partial_{\tau=t}\, J_\tau + H_t \circ dJ_t = 0$ *.*

Figure 2.8: Christian Gustav Adolph Mayer (1839 - 1907)

Proof. Combining the relations provided in theorem 2.4.8:

$$\begin{cases} dJ_t \circ \boldsymbol{\tau}_{\mathbb{C}} = d_{\mathrm{F}} L_t \,, \\ \partial_{\tau=t} \, J_\tau \circ \boldsymbol{\tau}_{\mathbb{C}} = -E_t \,, \end{cases}$$

and the definition $E_t := H_t \circ d_{\mathrm{F}} L_t$, we get the result. ■

2.5 Geometrical optics

We owe to the greek scientist Heron of Alexandria the first statement about the shortest path followed by reflected light rays.

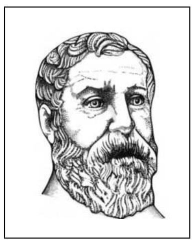

Figure 2.9: Heron of Alexandria (10 - 70)

About one thousand years later, the muslim scientist Ibn al-Haytham, considered the *father of optics* for his book *Book of Optics (Kitab al Manazir)*, extended this principle to refraction of light and provided many extraordinary contributions to optics, calculus, mechanics based on a scientific methodology.

Figure 2.10: Ibn al-Haytham (965 - 1039)

The definitive statement of the action principle of geometrical optics is due to Pierre de Fermat in a letter dated January 1st 1662 to Cureau de la Chambre. This was, with any evidence, the first variational statement of a general physical law.

Fermat's principle is intimately related to the concept of a geodesic and indeed may be enunciated by stating that a ray of light is a geodesic path in the euclidean space endowed with a piecewise regular riemannian metric tensor field, the optical tensor.

The principle provides a formidable motivation for Riemann's idea of a metric field varying from point to point and also undergoing discontinuities across singularity surfaces.

This last situation is similar to the one in which geodesic paths are drawn on the surface of a parallelepiped, as is made to fasten a string around a gift-box.

The calculus of variations of geometrical optics has a peculiar feature in common with geodesics: the Lagrangian functional is a continuous, convex (in fact sublinear) functional on the velocity phase-space which is not fiber-differentiable at the origin.

The conjugacy correspondence between vectors and covectors induced by the fiber derivative of the Lagrangian is no more one-to-one, but rather multivocal and maximal monotone.

These aspects of simple problems in calculus of variations are understated in most treatments, including authoritative articles [76].

To commute from the lagrangian to the hamiltonian description we need the Fenchel transform between convex functionals as described in Chapter 4.

The complementary description of light propagation in terms of wave fronts is based on the *eikonal equation* which is the counterpart of Hamilton-Jacobi partial differential equation of mechanics, when dealing with a Lagrangian functional which is sublinear, hence convex, but non everywhere fiber-differentiable.

2.5.1 Optical index

To illustrate the basics of geometrical optics, let us consider an optical medium in a riemannian manifold $(\mathbb{M}\,, \mathbf{g})$ and denote by $S^1(\mathbb{T}_\mathbf{x}\mathbb{M}, \mathbf{g})$ and by $B^1(\mathbb{T}_\mathbf{x}\mathbb{M}, \mathbf{g})$ the unit sphere and the closed unit ball at $\mathbf{x} \in \mathbb{M}$ according to the metric $\mathbf{g}$.

The fiber subbundle of the tangent bundle whose fibers are the unit spheres (balls) in the tangent spaces, will be accordingly denoted by $S^1(\mathbb{T}\mathbb{M}, \mathbf{g}) \subset \mathbb{T}\mathbb{M}$ ($B^1(\mathbb{T}\mathbb{M}, \mathbf{g}) \subset \mathbb{T}\mathbb{M}$).

The *optical metric* is a square integrable symmetric and positive definite tensor field $\mathbf{n} \in BL\,(\mathbb{T}\mathbb{M}^2\,; \Re)$ whose point-values $\mathbf{n}_\mathbf{x} \in BL\,(\mathbb{T}_\mathbf{x}\mathbb{M}^2\,; \Re)$ describe the light propagation properties. At each point $\mathbf{x} \in \mathbb{M}$, the *optical index* or *index of refraction* $n \in \mathrm{C}^0(S^1(\mathbb{T}\mathbb{M}, \mathbf{g})\,; \Re)$ is the sublinear function

$$n(\mathbf{v}(\mathbf{x})) := \|\mathbf{v}(\mathbf{x})\|_\mathbf{n} = \sqrt{\mathbf{n}_\mathbf{x}(\mathbf{v}(\mathbf{x}), \mathbf{v}(\mathbf{x}))}\,,$$

of the versors $\mathbf{v}(\mathbf{x}) \in S^1(\mathbb{T}_\mathbf{x}\mathbb{M}, \mathbf{g}) \subset \mathbb{T}_\mathbf{x}\mathbb{M}$.

The optical index is the reciprocal of the dimensionless *scalar light speed* $c \in \mathrm{C}^0(S^1(\mathbb{T}\mathbb{M}, \mathbf{g})\,; \Re)$, which is the ratio between the light speed in the vacuum and the one in the optical medium at $\mathbf{x}$:

$$n(\mathbf{v}) = \frac{1}{c(\mathbf{v})} := \|\mathbf{v}\|_\mathbf{n}\,, \quad \forall\, \mathbf{v} \in S^1(\mathbb{T}\mathbb{M}, \mathbf{g})\,.$$

Being a norm associated with a metric tensor field, the optical index is a positive, closed and fiber-sublinear functional on the tangent bundle:

$$\begin{aligned} n(\mathbf{v} + \mathbf{u}) &\le n(\mathbf{v}) + n(\mathbf{u})\,, && \mathbf{v}, \mathbf{u} \in \mathbb{T}\mathbb{M}\,, \\ n(\alpha\mathbf{v}) &= |\alpha| n(\mathbf{v})\,, && \alpha \in \Re\,, \\ n(\mathbf{v}) &\ge 0\,, && \mathbf{v} \in \mathbb{T}\mathbb{M}\,. \end{aligned}$$

The optical metric, being positive definite, is nondegenerate. Considered as a bounded linear operator, the optical index $\mathbf{n} \in BL\,(\mathbb{T}\mathbb{M}\,; \mathbb{T}^*\mathbb{M})$, is invertible to $\mathbf{n}^{-1} \in BL\,(\mathbb{T}^*\mathbb{M}\,; \mathbb{T}\mathbb{M})$. In turn this inverse operator defines a metric in the cotangent space $\mathbf{n}^{-1} \in BL\,(\mathbb{T}^*\mathbb{M}^2\,; \Re)$. We set $q_\mathbf{n}(\mathbf{v}) := \|\mathbf{v}\|^2_\mathbf{n}$.

The *fiber derivative* of the convex *optical index* functional is well-defined for $\mathbf{v} \neq 0$ and is given by the covector field

$$d_{\mathrm{F}} n(\mathbf{v}) = \frac{\mathbf{n}\mathbf{v}}{\|\mathbf{v}\|_\mathbf{n}}\,, \quad \forall\, \mathbf{v} \in S^1(\mathbb{M}, \mathbf{g})\,.$$

The *optical index* functional is everywhere fiber-subdifferentiable and its *fiber-subdifferential* at $\mathbf{v} = 0$ is the unit ball in the optical metric:

$$\partial_{\mathrm{F}} n(0) = B^1(\mathbb{TM}, \mathbf{n}) .$$

From Convex Analysis [149], [72], [73], [155], we know that the optical index is the support functional of the unit ball $B^1(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})$:

$$n(\mathbf{v}) = \sup\{\langle \mathbf{v}^*, \mathbf{v} \rangle - \sqcup_{B^1(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})}(\mathbf{v}^*) \mid \mathbf{v}^* \in \mathbb{T}^*\mathbb{M}\} .$$

Accordingly, its convex conjugate is the indicator of the unit ball $B^1(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})$:

$$\sqcup_{B^1(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})}(\mathbf{v}^*) = \sup\{\langle \mathbf{v}^*, \mathbf{v} \rangle - n(\mathbf{v}) \mid \mathbf{v} \in \mathbb{TM}\} ,$$

which is everywhere fiber-subdifferentiable.

- The*fiber-subdifferential* of the unit ball indicatorat the point $\mathbf{v}^* \in \mathbb{T}^*\mathbb{M}$ is the convex outward normal cone $\mathcal{N}_{B^1(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})}(\mathbf{v}^*)$ to the unit ball $B^1(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})$.

- If $\|\mathbf{v}^*\|_{\mathbf{n}^{-1}} < 1$ then $\mathbf{v}^* \in \mathbb{T}^*\mathbb{M}$ is internal to the unit ball and the normal cone degenerates to the null vector.

- If $\|\mathbf{v}^*\|_{\mathbf{n}^{-1}} = 1$ then $\mathbf{v}^* \in S^1_{\mathbf{x}}(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})$ and the normal cone at $\mathbf{v}^* \in \mathbb{T}^*\mathbb{M}$ is the half-line generated by $\mathbf{n}^{-1}(\mathbf{v}^*) \in \mathbb{TM}$.

The *optical lenght* functional

$$\textsc{Optical Lenght}(\gamma) := \int_I \|\mathbf{v}_t\|_{\mathbf{n}}\, dt ,$$

associated to a regular path $\gamma \in \mathrm{C}^1(I\,; \mathbb{M})$, is the time expended by light in propagating thru the path. The integral is independent of the parametrization of $\gamma \in \mathrm{C}^1(I\,; \mathbb{M})$.

In most physical problems, the path $\gamma \in \mathrm{C}^0(I\,; \mathbb{M})$ is only *piecewise regular*. In optics singularities occur at discontinuity surfaces between two media with different optical indexes.

A *regularity patchwork* $\textsc{Pat}(I)$ is a finite family of open, non overlapping segments such that the union of their closures covers the interval I . If the path is continuously differentiable in each element of the patchwork, we write $\gamma \in \mathrm{C}^1(\textsc{Pat}(I)\,; \mathbb{M})$.

Let us consider a flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ in the optical medium. The flow velocity field $\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ at a point $\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t) \in \mathbb{M}$ is denoted by $\mathbf{v}_{\boldsymbol{\varphi}_t} := \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))$.

The flow is said to be a *virtual flow* according to the patchwork $\mathrm{PAT}(I)$ if its velocity is a *virtual velocity*, that is tangent to the patchwork interelement boundaries.

2.5.2 Fermat's principle

Let us now provide a precise statement of the basic variational principle of geometrical optics.

Definition 2.5.1 (Fermat's principle) *A light ray is a piecewise regular path* $\gamma \in \mathrm{C}^1(\mathrm{PAT}(I)\,;\mathbb{M})$ *with a stationary optical lenght, that is:*

$$\partial_{\lambda=0} \int_{\mathrm{PAT}(I)} \|\boldsymbol{\varphi}_\lambda \uparrow \mathbf{v}_t\|_{\mathbf{n}}\, dt = \int_{\partial\mathrm{PAT}(I)} \langle \frac{\mathbf{n}\mathbf{v}_t}{\|\mathbf{v}_t\|_{\mathbf{n}}}, \mathbf{v}_{\boldsymbol{\varphi}_t} \rangle\, dt\,,$$

for any virtual flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *in the optical medium.*

Figure 2.11: Pierre de Fermat (1601 - 1665)

According to FERMAT's variational principle, the light rays are thus geodesic paths in the riemannian manifold $(\mathbb{M}\,,\mathbf{n})$, (see proposition 1.9.4 on page 206).

Remark 2.5.1 *In the literature (see e.g. [76], [6])* FERMAT*'s principle is usually enunciated by stating that the time expended by light, in propagating thru a* *ray segment joining two given points, is stationary in the class of varied paths sharing the same end points.*

In isotropic optical media, the optical metric is proportional to the euclidean metric, so that $\mathbf{n} = n^2\mathbf{g}$ *and*

$$\|\mathbf{v}_t\|_{\mathbf{n}} = n\,\|\mathbf{v}_t\|_{\mathbf{g}}\,, \qquad \frac{\mathbf{n}\mathbf{v}_t}{\|\mathbf{v}_t\|_{\mathbf{n}}} = n\,\frac{\mathbf{g}\mathbf{v}_t}{\|\mathbf{v}_t\|_{\mathbf{g}}}\,.$$

The stationarity condition in FERMAT*'s principle may then be written as*

$$\partial_{\lambda=0}\int_{\mathrm{PAT}(I)} n\,\|\boldsymbol{\varphi}_\lambda\!\uparrow\!\mathbf{v}_t\|_{\mathbf{g}}\,dt = \int_{\partial\mathrm{PAT}(I)} \langle\, n\,\frac{\mathbf{g}\mathbf{v}_t}{\|\mathbf{v}_t\|_{\mathbf{g}}}, \mathbf{v}_{\boldsymbol{\varphi}_t}\,\rangle\,dt\,.$$

The boundary integral is the product of the slowness times the rate of increase of the lenght of the ray due to the variation induced by the flow. It follows that FERMAT*'s principle may be enunciated by stating that the time expended by light is stationary with respect to any virtual variation of the path. In the general anisotropic case, this interpretation is no more feasible, contrary to the usual claim (see e.g. [6]). In this respect our remark 2.5.3, which deals with to the laws of reflection and refraction at the surfaces of discontinuity for the optical tensor, should also be consulted.*

FERMAT's variational principle may be interpreted in terms of the action principle of dynamics, as stated in Definition 2.1.4, by taking the phase manifold to be the tangent manifold $\mathbb{TM}$ and the Lagrangian $\mathcal{L} \in \mathrm{C}^0(\mathbb{TM}\,;\Re)$ to be the convex optical index functional $n \in \mathrm{C}^0(\mathbb{TM}\,;\Re)$, which is the pointwise support functional of the unit ball in the riemannian manifold $(\mathbb{M}\,,\mathbf{n})$. The Hamiltonian is the *indicator functional* of the closed unit ball $B^1_{\mathbf{x}}(\mathbb{T}^*\mathbb{M},\mathbf{n}^{-1})$:

$$H(\mathbf{v}^*) := \sup\{\langle\mathbf{v}^*,\mathbf{v}\rangle - n(\mathbf{v}) \mid \mathbf{v}\in\mathbb{TM}\} = \sqcup_{B^1(\mathbb{T}^*\mathbb{M},\mathbf{n}^{-1})}(\mathbf{v}^*)\,.$$

The Lagrangian is *positively homogeneous* (and not *homogeneous*, as affirmed in [76]) and has the same differential at all points along the (open) straigth half-lines from the origin (excluded).

The Hamiltonian vanishes in the closed unit ball (in the optical norm) where the covectors are tied to remain (not *identically*, as affirmed in [76]).

At a point $\mathbf{v}^* \in B^1(\mathbb{T}^*\mathbb{M},\mathbf{n}^{-1})$, the evolution-speed is in the closed convex cone $\mathcal{N}_{B^1(\mathbb{T}^*\mathbb{M},\mathbf{n}^{-1})}(\mathbf{v}^*)$ normal to the closed unit ball $B^1(\mathbb{T}^*\mathbb{M},\mathbf{n}^{-1})$.

It follows that the evolution-speed along a ray is either zero or has an undetermined amplitude. No time-evolutive condition then follows from this stationarity condition.

2.5.3 Eikonal equation

The one-form $\boldsymbol{\omega}^1(\mathbf{v}_t^*) = \boldsymbol{\theta}(\mathbf{v}_t^*) - H(\mathbf{v}_t^*)\,dt$ is given by

$$\boldsymbol{\omega}^1(\mathbf{v}_t^*) = \boldsymbol{\theta}(\mathbf{v}_t^*)\,, \quad \forall\, \mathbf{v}_t^* \in B^1(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})\,,$$

The action functional $S \in \mathrm{C}^1(\mathbb{TM}\,;\Re)$ is such that $dS = \boldsymbol{\theta} \in \mathbb{T}^*\mathbb{T}^*\mathbb{M}$. The related eikonal functional $J_t \in \mathrm{C}^1(\mathbb{M}\,;\Re)$ is such that $dJ = \mathbf{v}_t^* \in \mathbb{T}^*\mathbb{M}$.

The Hamilton-Jacobi equation for the eikonal functional gives:

$$\partial_{\tau=t}\; J_\tau(\mathbf{x}) + H(dJ(\mathbf{x})) = 0\,.$$

Since $H = \sqcup_{B^1(\mathbb{T}^*\mathbb{M},\mathbf{n})}$, this equation splits into

$$\begin{aligned} \partial_{\tau=t}\; J_\tau(\mathbf{x}) &= 0\,, \\ dJ_t(\mathbf{x}) &\in B^1_{\mathbf{x}}(\mathbb{T}^*\mathbb{M}, \mathbf{n}^{-1})\,. \end{aligned}$$

By the former condition, the eikonal functional does not depend explicitly on the evolution parameter $t \in I$ and, by the latter condition, its derivative belongs to the unit ball, in the cotangent bundle, according to the optical metric. This property is expressed by the *eikonal inequality* $\|dJ\|_{\mathbf{n}^{-1}} \le 1$. The discussion at the end of section 2.5.1, shows that during light propagation the eikonal functional is solution of the nonlinear partial differential equation

$$\|dJ\|_{\mathbf{n}^{-1}} = 1\,,$$

which is called the *eikonal* equation.

Let us set $\mathbf{n} = \mathbf{g}\mathbf{N}$, with $\mathbf{N} \in BL(\mathbb{TM}\,;\mathbb{TM})$ so that

$$dJ = \mathbf{n}\nabla_{\mathbf{n}} J = \mathbf{g}\mathbf{N}\nabla_{\mathbf{n}} J = \mathbf{g}\nabla_{\mathbf{g}} J\,.$$

The vector $\nabla_{\mathbf{g}} J \in \mathbb{TM}$ was called by Hamilton the *normal slowness* of the wave front. Indeed, in terms of the gradient $\nabla_{\mathbf{g}} J \in \mathbb{TM}$ it is

$$\|dJ\|^2_{\mathbf{n}^{-1}} = \|\nabla_{\mathbf{n}} J\|^2_{\mathbf{n}} = \mathbf{n}\,(\nabla_{\mathbf{n}} J, \nabla_{\mathbf{n}} J) = \mathbf{g}\,(\mathbf{N}^{-1}\nabla_{\mathbf{g}} J, \nabla_{\mathbf{g}} J) = 1\,.$$

In isotropic optical media, being $\mathbf{n} = n^2\mathbf{g}$, that is $\mathbf{N} = n^2\mathbf{I}$, we have that

$$\|dJ\|^2_{\mathbf{n}^{-1}} = \|\nabla_{\mathbf{n}} J\|^2_{\mathbf{n}} = n^{-2}\,\mathbf{g}\,(\nabla_{\mathbf{g}} J, \nabla_{\mathbf{g}} J) = n^{-2}\,\|\nabla_{\mathbf{g}} J\|^2_{\mathbf{g}} = 1\,,$$

so that $\|\nabla_{\mathbf{g}} J\|_{\mathbf{g}} = n = 1/c$.

2.5.4 Light evolution

To find an ordinary differential equation for the light propagation along a ray, we may modify the statement of FERMAT principle in order to deal with a fiber-differentiable Lagrangian.

Proposition 2.5.1 (A light evolution principle) *A ray of light is a path* $\gamma \in \mathrm{C}^1(\mathrm{PAT}(I)\,;\mathbb{M})$ *which, when the speed of its parametrization is proportional to the speed of light, fulfils the variational condition:*

$$\partial_{\lambda=0} \int_{\mathrm{PAT}(I)} \tfrac{1}{2}\, \|\boldsymbol{\varphi}_\lambda \!\uparrow\! \mathbf{v}_t\|^2_{\mathbf{n}}\, dt = \int_{\partial\mathrm{PAT}(I)} \mathbf{n}(\mathbf{v}_t, \mathbf{v}_{\boldsymbol{\varphi}_t})\, dt\,,$$

whose EULER *differential condition is:*

$$\tfrac{1}{2}\, (\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{n})(\mathbf{v}_t, \mathbf{v}_t) = \partial_{\tau=t}\; \mathbf{n}(\mathbf{v}_\tau, \mathbf{v}_{\boldsymbol{\varphi}_\tau})\,,$$

with the jump conditions

$$\langle\, [[\mathbf{n}(\mathbf{v}_t)]], \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))\,\rangle = 0\,,$$

for any virtual flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ *.*

Proof. Firstly we remark that, by definition:

$$\|\boldsymbol{\varphi}_\lambda \!\uparrow\! \mathbf{v}_t\|^2_{\mathbf{n}} = (\boldsymbol{\varphi}_\lambda \!\downarrow\! \mathbf{n})(\mathbf{v}_t, \mathbf{v}_t)\,,$$

and that, by assumption, $\|\mathbf{v}_t\|_{\mathbf{n}} := \sqrt{\mathbf{n}(\mathbf{v}_t, \mathbf{v}_t)} = n(\mathbf{v}_t) = \alpha > 0$. Then

$$\partial_{\lambda=0} \sqrt{(\boldsymbol{\varphi}_\lambda \!\downarrow\! \mathbf{n})(\mathbf{v}_t, \mathbf{v}_t)} = \frac{\partial_{\lambda=0}\, (\boldsymbol{\varphi}_\lambda \!\downarrow\! \mathbf{n})(\mathbf{v}_t, \mathbf{v}_t)}{2\, \sqrt{\mathbf{n}(\mathbf{v}_t, \mathbf{v}_t)}} = \tfrac{1}{2\alpha}\, (\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{n})(\mathbf{v}_t, \mathbf{v}_t)\,,$$

and the variational condition in the statement of FERMAT principle may be written

$$\tfrac{1}{2} \int_{\mathrm{PAT}(I)} (\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{n})(\mathbf{v}_t, \mathbf{v}_t)\, dt = \int_{\partial\mathrm{PAT}(I)} \mathbf{n}(\mathbf{v}_t, \mathbf{v}_{\boldsymbol{\varphi}_t})\, dt\,,$$

which is equivalent to

$$\tfrac{1}{2} \int_{\mathrm{PAT}(I)} (\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}} \mathbf{n})(\mathbf{v}_t, \mathbf{v}_t)\, dt = \int_{\mathrm{PAT}(I)} \partial_{\tau=t}\; \mathbf{n}(\mathbf{v}_\tau, \mathbf{v}_{\boldsymbol{\varphi}_\tau})\, dt\,,$$

and, by the arbitrarity of the virtual flow, to the differential and the jump conditions in the statement. ■

Proposition 2.5.2 (Differential equation of light rays) *In an optical medium* $(\mathbb{M}, \mathbf{n})$ *with a connection* ∇, *a path* $\gamma \in C^1(I; \mathbb{M})$, *whose parametrization-speed is proportional to the speed of light, is a light-ray if and only if it fulfills the differential equation:*

$$\partial_{\tau=t}\, \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \boldsymbol{\chi}_{\tau,t}\Uparrow \mathbf{v}_{\boldsymbol{\varphi}_t}) = \tfrac{1}{2}\langle d_{\mathrm{B}} q_{\mathbf{n}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle + \langle (\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\mathbf{v}_t)\mathrm{TORS}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle,$$

which may be also written as

$$\nabla_{\mathbf{v}_t}(\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbf{v}) = \tfrac{1}{2}\, d_{\mathrm{B}} q_{\mathbf{n}}(\mathbf{v}_t) + (\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}\mathbf{v}_t)\mathrm{TORS}(\mathbf{v}_t).$$

Proof. We have that

$$\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{n})(\mathbf{v}_t, \mathbf{v}_t) = \tfrac{1}{2}\partial_{\lambda=0}\, \mathbf{n}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))}(\boldsymbol{\varphi}_\lambda\uparrow \mathbf{v}_t, \boldsymbol{\varphi}_\lambda\uparrow \mathbf{v}_t).$$

The velocity along the path may be extended to a vector field $\mathbf{v} \in C^1(\mathbb{M}; \mathbb{TM})$ by pushing it along the flow $\boldsymbol{\varphi}_\lambda \in C^1(\mathbb{M}; \mathbb{M})$ according to the relation:

$$\mathbf{v}(\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))) := \boldsymbol{\varphi}_\lambda\uparrow \mathbf{v}_t.$$

Then, writing $\boldsymbol{\varphi}_\lambda\uparrow \mathbf{v}_t = \boldsymbol{\varphi}_\lambda\Uparrow \boldsymbol{\varphi}_\lambda\Downarrow \boldsymbol{\varphi}_\lambda\uparrow \mathbf{v}_t$ and applying LEIBNIZ rule, we get

$$\begin{aligned}
\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{n})(\mathbf{v}_t, \mathbf{v}_t) &= \tfrac{1}{2}\partial_{\lambda=0}\, \mathbf{n}_{\boldsymbol{\varphi}_\lambda(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t))}(\boldsymbol{\varphi}_\lambda\Uparrow \mathbf{v}_t, \boldsymbol{\varphi}_\lambda\Uparrow \mathbf{v}_t) \\
&\quad + \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}(\partial_{\lambda=0}\, \boldsymbol{\varphi}_\lambda\Downarrow \boldsymbol{\varphi}_\lambda\uparrow \mathbf{v}_t, \mathbf{v}_t) \\
&= \tfrac{1}{2}\langle d_{\mathrm{B}} q_{\mathbf{n}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle + \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_{\boldsymbol{\varphi}_t}}\mathbf{v}, \mathbf{v}_t).
\end{aligned}$$

Similarly, defining the trajectory-flow $\boldsymbol{\chi}_{\tau,t} \in C^1(\mathbb{M}; \mathbb{M})$ by $\boldsymbol{\chi}_{\tau,t} \circ \gamma_t = \gamma_\tau$, we have that

$$\begin{aligned}
\partial_{\tau=t}\, \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \mathbf{v}_{\boldsymbol{\varphi}_\tau}) &= \partial_{\tau=t}\, \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \boldsymbol{\chi}_{\tau,t}\Uparrow \boldsymbol{\chi}_{\tau,t}\Downarrow \mathbf{v}_{\boldsymbol{\varphi}_\tau}) \\
&= \partial_{\tau=t}\, \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_\tau)}(\mathbf{v}_\tau, \boldsymbol{\chi}_{\tau,t}\Uparrow \mathbf{v}_{\boldsymbol{\varphi}_t}) + \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_t) \\
&= \langle \nabla_{\mathbf{v}_t}(\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbf{v}), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle + \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}_t).
\end{aligned}$$

By definition of the vector field $\mathbf{v} \in C^1(\mathbb{M}; \mathbb{TM})$ we have that $[\mathbf{v}_{\boldsymbol{\varphi}}, \mathbf{v}] = 0$ and hence

$$\mathrm{TORS}(\mathbf{v}) \cdot \mathbf{v}_{\boldsymbol{\varphi}} = \mathrm{tors}(\mathbf{v}, \mathbf{v}_{\boldsymbol{\varphi}}) = \nabla_{\mathbf{v}}\mathbf{v}_{\boldsymbol{\varphi}} - \nabla_{\mathbf{v}_{\boldsymbol{\varphi}}}\mathbf{v}.$$

The differential condition of proposition 2.5.1 may then be written as

$$\tfrac{1}{2}\langle d_{\mathrm{B}} q_{\mathbf{n}}(\mathbf{v}_t), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle = \langle \nabla_{\mathbf{v}_t}(\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbf{v}), \mathbf{v}_{\boldsymbol{\varphi}_t}\rangle + \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}(\mathbf{v}_t, \mathrm{TORS}(\mathbf{v}_t) \cdot \mathbf{v}_{\boldsymbol{\varphi}_t}).$$

and the statement is proven. ■

Remark 2.5.2 *In the riemannian manifold* $(\mathbb{M}, \mathbf{n})$*, endowed with the connection* ∇ *induced by local charts, the torsion vanishes and the differential equations of a light-ray becomes*

$$\nabla_{\mathbf{v}_t}(\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbf{v}) = \tfrac{1}{2}\, d_{\mathrm{B}} q_{\mathbf{n}}(\mathbf{v}_t)\,.$$

In a connection which preserves the optical metric, we have that $d_{\mathrm{B}} q_{\mathbf{n}} = 0$ *and* $\nabla_{\mathbf{v}_t}\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})} = 0$*, so that*

$$\langle \nabla_{\mathbf{v}_t}(\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}\mathbf{v}), \mathbf{w} \rangle = d_{\mathbf{v}_t}\mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v})}(\mathbf{v}, \mathbf{w}) - \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}(\mathbf{v}_t, \nabla_{\mathbf{v}_t}\mathbf{w}) = \mathbf{n}_{\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)}(\nabla_{\mathbf{v}_t}\mathbf{v}, \mathbf{w})\,,$$

for any $\mathbf{w} \in \mathrm{C}^1(\mathbb{M}; \mathbb{TM})$*, Hence, in the* LEVI-CIVITA *connection associated with the optical metric, which is torsion-free and optical-metric preserving, the differential equations of a light-ray becomes* $\nabla_{\mathbf{v}_t}\mathbf{v} = 0$*.*

Remark 2.5.3 *In isotropic optical media, the jump condition*

$$\langle [[\mathbf{n}(\mathbf{v}_t)]], \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle = 0\,,$$

reads

$$\langle [[n\,\mathbf{g}(\mathbf{v}_t)]], \mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\tau}_{\mathbb{C}}(\mathbf{v}_t)) \rangle = 0\,,$$

for any virtual flow $\boldsymbol{\varphi}_\lambda \in \mathrm{C}^1(\mathbb{M}; \mathbb{M})$*. Since the velocities of virtual flows are tangent to the discontinuity surface, the law of reflection and* SNELL*'s law of refraction in isotropic media are immediately deduced. In general anisotropic optical media,* SNELL*'s law is not adequate to describe the refraction properties.*

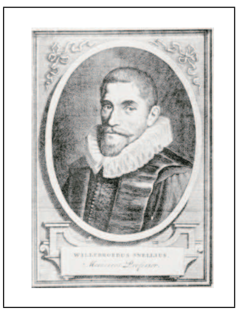

Figure 2.12: Willebrord Snellius (1580 - 1626)

2.5.5 Dynamics vs Optics

FERMAT's principle in optics postulates that the integral of the optical lenght density along a ray is an extremal with respect to the variation induced by any virtual flow.

HAMILTON's action principle in dynamics postulates that the integral of the Lagrangian along a trajectory is an extremal with respect to the variation induced by any virtual flow.

A basic difference is that the optical lenght of a path is independent of its parametrization, while the HAMILTON's action integral depends on the parametrization of the trajectory.

This is quite natural since FERMAT's principle for optical rays was not intended to evaluate the speed of light along a ray, but only the image of the ray. In fact the speed of light is considered as a constitutive property of the optical medium.

In mechanics, on the contrary, both the image of the trajectory and its time-law are governed by HAMILTON's action principle.

There are two main ways to provide a parametrization independent formulation of dynamics.

The older way was first formulated by MAUPERTUIS and then made precise by EULER, LAGRANGE and JACOBI. It is concerned with the case in which the energy of the system is constant along the trajectory. The idea is to consider a constant energy submanifold of the tangent bundle and to restrict to it the differental condition of stationarity. As a consequence the canonical two-form becomes a contact form with a one-dimensional kernel made of characteristic vectors. The corresponding integral line provides the geometric description of the trajectory in the velocity phase-space. A suitable reparametrization permits to recover a full description of the trajectory. Arbitrary variations of the velocity in the constant energy submanifold, are allowed for in the action principle.

The other more recent way is due to E. CARTAN and has been reproduced as an action principle by ARNOLD. The underlying idea is to enlarge the velocity phase-space to a velocity-time state-space. As a consequence the canonical two-form becomes a contact form once more and variations in velocity and time are considered in the action principle. In our formulation no end-point conditions are appended and velocity variations are assumed to be projectable vector fields.

2.6 Symplectic structure

The peculiar form of HAMILTON's system of ordinary differential equation for the momentum $\dot{\mathbf{v}}^* \in \mathbb{T}_{\mathbf{v}^*}\mathbb{T}^*\mathbb{C}$, suggests to endow the covelocity-phase-space $\mathbb{T}^*\mathbb{C}$ of a special kind of geometry in which the role of the symmetric and positive definite metric tensor $\mathbf{g}$ of riemannian geometry, is played by a skew-symmetric, closed and weakly non-degenerate differential two-form $\boldsymbol{\omega}^2$: A detailed account of these geometrical structures can be found in [6] for the finite dimensional case. Symplectic infinite dimensional spaces are dealt with in [106].

We will not treat this topic in detail here. Instead we will show how some basic results of classical mechanics may be directly inferred from the skew-symmetric structure of HAMILTON's equations.

Let us consider a differentiable manifold $\mathbb{M}$ and a differential two-form $\boldsymbol{\omega}^2$ on $\mathbb{M}$ such that:

- the form $\boldsymbol{\omega}^2 \in \mathrm{C}^1(\mathbb{M}\,;\Lambda^2(\mathbb{M}))$ is closed:

$$d\boldsymbol{\omega}^2 = 0\,,$$

- the form $\boldsymbol{\omega}^2 \in \mathrm{C}^1(\mathbb{M}\,;\Lambda^2(\mathbb{M}))$ is non degenerate:

$$\boldsymbol{\omega}^2 \cdot \mathbf{X} = 0 \iff \mathbf{X} = 0\,, \qquad \mathbf{X} \in \mathbb{T}\mathbb{C}\,.$$

The pair $\{\mathbb{M}, \boldsymbol{\omega}^2\}$ is called a *symplectic manifold*.

- We say that a time-dependent vector field $\mathbf{X}_{H_t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ admits an hamiltonian functional $H_t \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ if

$$\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} = dH_t\,.$$

The non degeneracy of $\boldsymbol{\omega}^2$ ensures that *hamiltonian vector field* corresponding to a given hamiltonian functional is unique.

A necessary condition in order that the vector field $\mathbf{X}_{H_t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{T}\mathbb{M})$ be hamiltonian is that

$$d(\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t}) = ddH_t = 0\,.$$

If the manifold is *star shaped* the previous condition is also sufficient by POINCARÉ lemma (see section 1.6.12 on page 154).

2.6.1 Poisson brackets

Let $\boldsymbol{\varphi}^K_{t,s} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{M})$ be the flow of a vector field $\mathbf{X}_{K_t} \in \mathrm{C}^1(\mathbb{M}\,;\mathbb{TM})$ and $H_t \in \mathrm{C}^2(\mathbb{C}\,;\Re)$ be a time dependent functional. Then

- The *time-convective derivative* of $H_t \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ along the flow generated by $\mathbf{X}_{K_t}$ is given by

$$(\mathcal{L}_{t,\mathbf{X}_{K_t}} H_t)_s(\mathbf{x}) = \partial_{t=s}\; H_t(\boldsymbol{\varphi}^K_{t,s}(\mathbf{x})) = \partial_{t=s}\; (\boldsymbol{\varphi}^K_{t,s}{\downarrow}H_t)(\mathbf{x})\,, \quad \mathbf{x} \in \mathbb{M}\,.$$

We have that

$$\mathcal{L}_{t,\mathbf{X}_{K_t}} H_t = \mathcal{L}_{\mathbf{X}_{K_t}} H_t + \partial_{\tau=t}\; H_\tau\,,$$

where $\partial_{\tau=t}\; H_\tau$ is the *partial time derivative* and $\mathcal{L}_{\mathbf{X}_{K_t}} H_t$ is the spatial directional derivative of the functional K_t along the vector $\mathbf{X}_{K_t}$, also called the *autonomous* LIE *derivative* along the time-dependent flow generated by $\mathbf{X}_{K_t}$ (see section 1.3.7 on page 77).

Figure 2.13: Siméon Denis Poisson (1781 - 1840)

- The POISSON *bracket* of two time dependent functionals $H_t, K_t \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ is the functional $[K_t, H_t] \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ defined by

$$[K_t, H_t] := \boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} \cdot \mathbf{X}_{K_t} = dH_t \cdot \mathbf{X}_{K_t} = \mathcal{L}_{\mathbf{X}_{K_t}} H_t = \mathcal{L}_{t,\mathbf{X}_{K_t}} H_t - \partial_{\tau=t}\; H_\tau\,,$$

  which is skew-symmetric in H_t and K_t.

As a direct consequence we derive an invariance result which is a special extension of NOETHER's theorem [6]:

For time dependent fields we have that:

- The POISSON bracket of two time dependent functionals vanishes iff each one of them is dragged along the flow generated by the hamiltonian vector field corresponding to the other:

$$[K_t, H_t] = -[H_t, K_t] = 0 \iff \mathcal{L}_{\mathbf{X}_{H_t}} K_t = -\mathcal{L}_{\mathbf{X}_{K_t}} H_t = 0$$

$$\iff \begin{cases} \mathcal{L}_{t,\mathbf{X}_{K_t}} K_t = \partial_t K_t \\ \mathcal{L}_{t,\mathbf{X}_{K_t}} H_t = \partial_t H_t \,. \end{cases}$$

Hence in particular (*drag of the energy*):

- Any time dependent functional is dragged along the flow generated by its hamiltonian vector field:

$$[H_t, H_t] = 0 \iff \mathcal{L}_{\mathbf{X}_{H_t}} H_t = \mathcal{L}_{t,\mathbf{X}_{K_t}} H_t - \partial_{\tau=t}\ H_\tau = 0$$

$$\iff \mathcal{L}_{t,\mathbf{X}_{K_t}} H_t = \partial_{\tau=t}\ H_\tau \,.$$

For time independent fields the previous result may be stated as follows.

- The vanishing of the POISSON bracket of two time independent functionals is necessary and sufficient in order that each one of them be constant along the flow generated by the hamiltonian vector field corresponding to the other:

$$[H, K] = -[K, H] = 0 \iff \mathcal{L}_{\mathbf{X}_H} K = -\mathcal{L}_{\mathbf{X}_K} H = 0\,.$$

From this result we infer that (*conservation of energy*):

- Any time independent functional is constant along the flow generated by its hamiltonian vector field:

$$[H, H] = 0 \iff \mathcal{L}_{\mathbf{X}_H} H = 0\,.$$

2.6.2 Canonical transformations

Until now we have made no use of the closedness of the symplectic two-form $\boldsymbol{\omega}^2 \in \mathrm{C}^1(\mathbb{M}\,; \Lambda^2(\mathbb{M}))$. The reason why this assumption is made will be clarified hereafter. To this end we recall the definition of canonical flow (see section 2.4.4).

- A flow $\mathbf{Fl}^{\mathbf{X}}_{t,s} \in \mathrm{C}^1(\mathbb{M} \times I\,; \mathbb{M} \times I)$ is said to be *canonical* if it drags the symplectic two-form $\boldsymbol{\omega}^2 \in \mathrm{C}^1(\mathbb{M}\,; \Lambda^2(\mathbb{M}))$:

$$\mathcal{L}_{\mathbf{X}_t}\boldsymbol{\omega}^2 = 0 \quad \text{or equivalently} \quad \mathbf{Fl}^{\mathbf{X}}_{t,s}\!\downarrow\!\boldsymbol{\omega}^2 = \boldsymbol{\omega}^2 \,.$$

The closedness of the symplectic two-form $\boldsymbol{\omega}^2 \in \mathrm{C}^1(\mathbb{M}\,; \Lambda^2(\mathbb{M}))$ opens the way to a proof of the next theorem which does not make direct recourse to POINCARÉ's relative integral invariant.

Theorem 2.6.1 *The flow of a time dependent hamiltonian vector field is canonical.*

Proof. By the homotopy formula and the closedness of the symplectic two-form, we have that:

$$\mathcal{L}_{\mathbf{X}_{H_t}}\boldsymbol{\omega}^2 = d\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} + d(\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t}) = d(\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t})\,.$$

Hence, if $\mathbf{X}_{H_t} \in \mathrm{C}^1(\mathbb{M}\,; \mathbb{TC})$ is a hamiltonian vector field: $\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} = dH_t$, we infer that

$$ddH_t = d(\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t}) = \mathcal{L}_{\mathbf{X}_{H_t}}\boldsymbol{\omega}^2 = 0\,.$$

As a consequence of this result we have the following important property.

Theorem 2.6.2 *The* LIE *bracket of two hamiltonian vector fields is an hamiltonian vector field and its hamiltonian is the* POISSON *bracket of the two hamiltonians:*

$$\boldsymbol{\omega}^2 \cdot [\mathbf{X}_{K_t}, \mathbf{X}_{H_t}] = d\,[K_t, H_t]\,.$$

Proof. The result is a direct consequence of the following equality between one-forms:

$$\begin{aligned} d\,[K_t, H_t] = d(\mathcal{L}_{\mathbf{X}_{K_t}} H_t) = \mathcal{L}_{\mathbf{X}_{K_t}}(dH_t) &= \mathcal{L}_{\mathbf{X}_{K_t}}(\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t}) \\ &= \boldsymbol{\omega}^2 \cdot \mathcal{L}_{\mathbf{X}_{K_t}}\mathbf{X}_{H_t} = \boldsymbol{\omega}^2 \cdot [\mathbf{X}_{K_t}, \mathbf{X}_{H_t}]\,. \end{aligned}$$

The fourth equality holds since, by the property $\mathcal{L}_{\mathbf{X}_{K_t}}\boldsymbol{\omega}^2 = 0$ and LEIBNIZ rule, we have that

$$\begin{aligned} \mathcal{L}_{\mathbf{X}_{K_t}}(\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t}) \cdot \mathbf{X} &= \mathcal{L}_{\mathbf{X}_{K_t}}(\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} \cdot \mathbf{X}) - \boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} \cdot \mathcal{L}_{\mathbf{X}_{K_t}}\mathbf{X} \\ &= \mathcal{L}_{\mathbf{X}_{K_t}}\boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} \cdot \mathbf{X} + \boldsymbol{\omega}^2 \cdot \mathcal{L}_{\mathbf{X}_{K_t}}\mathbf{X}_{H_t} \cdot \mathbf{X} \\ &\quad + \boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} \cdot \mathcal{L}_{\mathbf{X}_{K_t}}\mathbf{X} - \boldsymbol{\omega}^2 \cdot \mathbf{X}_{H_t} \cdot \mathcal{L}_{\mathbf{X}_{K_t}}\mathbf{X} \\ &= \boldsymbol{\omega}^2 \cdot \mathcal{L}_{\mathbf{X}_{K_t}}\mathbf{X}_{H_t} \cdot \mathbf{X} = \boldsymbol{\omega}^2 \cdot [\mathbf{X}_{K_t}, \mathbf{X}_{H_t}] \cdot \mathbf{X}\,. \end{aligned}$$

As a corollary we may state that:

- The flows of two time dependent hamiltonian vector fields commute if and only if the POISSON bracket of the two hamiltonians is locally constant on $\mathbb{M}$. Indeed the commutation of flows of two vector fields is equivalent to the vanishing of their LIE bracket and, by the non degeneracy of the symplectic form, we have that:

$$[\mathbf{X}_{H_t}, \mathbf{X}_{K_t}] = 0 \iff \boldsymbol{\omega}^2 \cdot [\mathbf{X}_{H_t}, \mathbf{X}_{K_t}] = 0 \iff d\,[H_t, K_t] = 0\,.$$

From this result we get another extension of NOETHER's theorem [6].

Theorem 2.6.3 *The* POISSON *brackets of any triplet of possibly time dependent functionals* $H_t, K_t, L_t \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ *fulfil the* JACOBI*'s identity:*

$$[[H_t, K_t], L_t] + [[L_t, H_t], K_t] + [[K_t, L_t], H_t] = 0\,.$$

Proof. We have that

$$\begin{aligned} &[[H_t, K_t], L_t] + [[L_t, H_t], K_t] = \\ &= [[L_t, H_t], K_t] - [[K_t, H_t], L_t] = \\ &= (\mathcal{L}_{\mathbf{X}_{K_t}} \mathcal{L}_{\mathbf{X}_{L_t}} - \mathcal{L}_{\mathbf{X}_{L_t}} \mathcal{L}_{\mathbf{X}_{K_t}})\, H_t \\ &= [\mathbf{X}_{K_t}, \mathbf{X}_{L_t}]\, H_t\,. \end{aligned}$$

Then, summing up twice the JACOBI triplet, we get an equality between a sum of second derivatives and a sum of first derivatives of the three functionals. The equality implies that the triplet must vanish.

As a direct consequence, we get the POISSON theorem.

Theorem 2.6.4 (Poisson theorem) *Let us assume that two time dependent functionals* $K_t, L_t \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ *are dragged by the flow generated by the hamiltonian vector field associated with a time dependent functional* $H_t \in \mathrm{C}^2(\mathbb{M}\,;\Re)$. *Then their* POISSON *bracket is also dragged by the flow.*

Proof. We have that: $[[K_t, L_t], H_t] = -[[H_t, K_t], L_t] - [[L_t, H_t], K_t] = 0$ and hence

$$[[K_t, L_t], H_t] = \mathcal{L}_{\mathbf{X}_{H_t}}[K_t, L_t] = \mathcal{L}_{t,\mathbf{X}_{K_t}}[K_t, L_t] - \partial_{\tau=t}\,[K_\tau, L_\tau] = 0\,.$$

The classical POISSON theorem for time independent functionals reads:

Theorem 2.6.5 *If two time independent functionals* $K, L \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ *are invariant along the flow generated by the hamiltonian vector field associated with a time dependent functional* $H_t \in \mathrm{C}^2(\mathbb{M}\,;\Re)$ *, then their* POISSON *bracket is also invariant.*

2.6.3 Integral invariants

Time independent forms

Let us recall the following definitions:

- A k-form $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is said to be an *integral invariant* of a transformation $\boldsymbol{\varphi} \in \mathrm{C}(\mathbb{M}\,;\mathbb{M})$ if the integral of $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ on any k-dimensional manifold $\mathbb{N} \subseteq \mathbb{M}$ is not changed by the transformation:

$$\int_{\mathbb{N}} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}(\mathbb{N})} \boldsymbol{\omega}^k = \int_{\mathbb{N}} \boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega}^k\,, \quad \forall\, \mathbb{N} \subseteq \mathbb{M} \iff \boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega}^k = \boldsymbol{\omega}^k\,.$$

The treatment developed in section 2.4.1 shows that the symplectic two-form $\boldsymbol{\omega}^2 \in \mathrm{C}^1(\mathbb{M}\,;\Lambda^2(\mathbb{M}))$ is an integral invariant of any hamiltonian flow, that is a *universal integral invariant*.

- A k-form $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is said to be a *relative integral invariant* of a transformation $\boldsymbol{\varphi} \in \mathrm{C}(\mathbb{M}\,;\mathbb{M})$ if the integral of $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ on any *closed* k-dimensional manifold $\mathbb{N} \subseteq \mathbb{M}$ is not changed by the transformation:

$$\int_{\mathbb{N}} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}(\mathbb{N})} \boldsymbol{\omega}^k = \int_{\mathbb{N}} \boldsymbol{\varphi}{\downarrow}\boldsymbol{\omega}^k\,, \quad \forall\, \mathbb{N} \subseteq \mathbb{M} \quad \text{such that} \quad \partial\mathbb{N} = 0\,.$$

We have that

- If a k-form $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is a relative integral invariant of $\boldsymbol{\varphi} \in \mathrm{C}(\mathbb{M}\,;\mathbb{M})$, then the $(k+1)$-form of $d\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^{k+1}(\mathbb{M}))$ is an integral invariant of the transformation.

Indeed for any k-dimensional submanifold $\mathbb{N} \subseteq \mathbb{M}$ we have

$$\int_{\mathbb{N}} d\boldsymbol{\omega}^k = \int_{\partial\mathbb{N}} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}(\partial\mathbb{N})} \boldsymbol{\omega}^k = \int_{\partial\boldsymbol{\varphi}(\mathbb{N})} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}(\mathbb{N})} d\boldsymbol{\omega}^k \,.$$

The converse statement holds only if any k-dimensional submanifold is the boundary of a $(k+1)$-dimensional submanifold.

- The one-form dH is an integral invariant for the flow of the time independent hamiltonian H, since the zero-form H is an integral invariant (and hence *a fortiori* a relative integral invariant).

- A time independent k-form $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is an *integral invariant* of a flow $\boldsymbol{\varphi}_{t,s} \in \mathrm{C}(\mathbb{M}\,;\mathbb{M})$ if the time derivative along the flow of its integral on any dragged k-dimensional manifold $\mathbb{N} \subseteq \mathbb{M}$ is equal to zero

$$\begin{aligned} \partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,s}(\mathbb{N})} \boldsymbol{\omega}^k &= \int_{\boldsymbol{\varphi}_{t,s}(\mathbb{N})} \mathcal{L}_{\mathbf{X}_t}\, \boldsymbol{\omega}^k = \int_{\mathbb{N}} \boldsymbol{\varphi}_{t,s}{\downarrow}\mathcal{L}_{\mathbf{X}_t}\, \boldsymbol{\omega}^k \\ &= \int_{\mathbb{N}} \partial_{\tau=t}\; \boldsymbol{\varphi}_{\tau,s}{\downarrow}\boldsymbol{\omega}^k = 0 \,, \end{aligned}$$

  where we have recalled the formula

$$\boldsymbol{\varphi}_{t,s}{\downarrow}\mathcal{L}_{\mathbf{X}_t}\boldsymbol{\omega}^k = \partial_{\tau=t}\; \boldsymbol{\varphi}_{\tau,s}{\downarrow}\boldsymbol{\omega}^k \,.$$

Hence a time independent k-form $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is an *integral invariant* of a flow $\boldsymbol{\varphi}_{t,s} \in \mathrm{C}(\mathbb{M}\,;\mathbb{M})$ iff its Lie derivative vanishes identically along the flow or equivalently if the time derivative of its pull back vanishes identically (the form is dragged by the flow):

$$\begin{aligned} \mathcal{L}_{\mathbf{X}_t}\, \boldsymbol{\omega}^k = 0 &\iff \partial_{\tau=t}\; \boldsymbol{\varphi}_{\tau,\mathbb{M}}{\downarrow}\boldsymbol{\omega}^k = 0 \\ &\iff \boldsymbol{\varphi}_{t,\mathbb{M}}{\downarrow}\boldsymbol{\omega}^k = \boldsymbol{\omega}^k \,. \end{aligned}$$

Time dependent forms

- A time dependent k-form $\boldsymbol{\omega}_t^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is a *dragged integral* of a flow $\boldsymbol{\varphi}_{t,s} \in \mathrm{C}(\mathbb{M}\,;\mathbb{M})$ with velocity field $\mathbf{X}_t \in \mathrm{C}(\mathbb{M}\,;\mathbb{TM})$ if the time derivative along the flow of its integral on any dragged k-dimensional manifold $\mathbb{N} \subseteq \mathbb{M}$ is equal the integral of its partial time derivative:

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,s}(\mathbb{N})} \boldsymbol{\omega}^k_\tau = \int_{\boldsymbol{\varphi}_{t,s}(\mathbb{N})} \mathcal{L}_{t,\mathbf{X}_t}\, \boldsymbol{\omega}^k_t = \int_{\mathbb{N}} \boldsymbol{\varphi}_{t,s}\!\downarrow\!(\mathcal{L}_{t,\mathbf{X}_t}\, \boldsymbol{\omega}^k_t)$$
$$= \int_{\mathbb{N}} \partial_{\tau=t}\, (\boldsymbol{\varphi}_{\tau,s}\!\downarrow\!\boldsymbol{\omega}^k_\tau)\,,$$

where we have made use of the transport formula:

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,s}(\mathbb{N})} \boldsymbol{\omega}^k_\tau = \int_{\boldsymbol{\varphi}_{t,s}(\mathbb{N})} \mathcal{L}_{t,\mathbf{X}_t}\, \boldsymbol{\omega}^k_t\,,$$

and have recalled that

$$\boldsymbol{\varphi}_{t,\mathbb{M}}\!\downarrow\!\mathcal{L}_{\mathbf{X}_t}\boldsymbol{\omega}^k_t = \partial_{\tau=t}\, \boldsymbol{\varphi}^*_{\tau,s}\boldsymbol{\omega}^k_\tau\,.$$

If a dragged integral k-form $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is time independent we have that

$$\partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,s}(\mathbb{N})} \boldsymbol{\omega}^k = \int_{\boldsymbol{\varphi}_{t,s}(\mathbb{N})} \mathcal{L}_{\mathbf{X}_t}\, \boldsymbol{\omega}^k\,, \qquad \forall\, \mathbb{N} \subseteq \mathbb{M}\,,$$

and hence the k-form $\boldsymbol{\omega}^k \in \mathrm{C}(\mathbb{M}\,;\Lambda^k(\mathbb{M}))$ is an integral invariant iff

$$\mathcal{L}_{\mathbf{X}_t}\, \boldsymbol{\omega}^k = 0\,.$$

2.7 Conclusions

About two centuries after LAGRANGE's and HAMILTON's genial discoveries and almost one century after EMMY NOETHER's masterpiece have passed away. In the meantime a simple extension of HAMILTON's action principle was at hand waiting to be discovered. This extension reveals that NOETHER's celebrated result is a direct consequence of a more general way of stating the law of dynamics.

We would like to feel that HILBERT's and EINSTEIN's praises for NOETHER's contribution of an invariant result in dynamics are also of support for the ideas presented in this paper. The extendend version of HAMILTON's action principle applyies in a natural way to piecewice regular paths and yields the corresponding jump conditions at singular points. The simple treatment based on standard calculus, was only achieved afted a translation of HAMILTON's action principle in

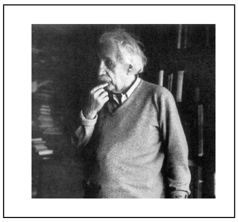

Figure 2.14: Albert Einstein (1879 - 1955)

geometrical differential terms and a subsequent analysis performed by the tools of calculus on manifolds. This revealed how to rewrite HAMILTON's principle and opened the way for the direct proof of the law of dynamics. Remarkably, the proof of this more general result is definitely simpler than the special, classical one of LAGRANGE's law of dynamics in a manifold with torsionless connection. The result is directly extendable to other problems of calculus of variations and in particular to the analysis of the properties of geodesic paths on a manifold. The dynamics of deformable bodies has been discussed in detail by a direct application of the general results and by pointing out some peculiar issues which deserve special attention. The principles of elastodynamics have been derived by a simple introduction of a hyperelastic constitutive law.

Chapter 3

Continuum Mechanics

In this chaper is devoted to an introduction of basic principles of nonlinear Continuum Mechanics. A geometric description of CAUCHY's model of a continuous body is provided as the tangent bundle associated to a 3D compact and connected embedded submanifold of the euclidean space. The rigidity condition and the relevant axiomatic definitions of static and dynamic equilibrium, in the actual and in the reference placement, are provided.

3.1 Bodies and deformations

According to CAUCHY's model, a *continuous material body*, briefly a *continuum* is a set of particles identified with the points $\mathbf{x} \in \mathbb{B}$ of a differentiable submanifold, referred to as the *reference placement*, embedded in the ambient euclidean space $\{\mathbb{S}\,,\mathbf{g}\}$.

The euclidean space is endowed with the standard metric tensor field $\mathbf{g}(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{S}^2\,;\Re)$ which is constant according to the standard connection induced by the distant parallel transport by translation.

We will denote by $\mathbb{TS}$ the tangent bundle to the euclidean space, in which each linear tangent space $\mathbb{T}_{\mathbf{x}}\mathbb{S}$ may be identified with the linear space of translations V.

In a mechanical theory, experimental tests provide measurements of the lenght of the material fibers (tangent vectors) at the points of a placement $\boldsymbol{\varphi}(\mathbb{B}) \subset \mathbb{S}$ of the body in the ambient space, described by a smooth configuration map $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ which is assumed to be a diffeomorphism between $\mathbb{B}$

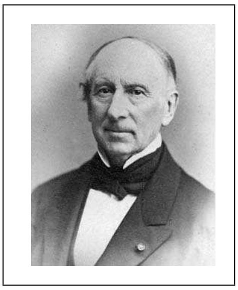

Figure 3.1: Augustin Louis Cauchy (1789 - 1857)

and $\boldsymbol{\varphi}(\mathbb{B})$.

The results of metric measurements can be interpreted by substituting the standard metric tensor $\mathbf{g}(\mathbf{x})$ at $\mathbf{x} \in \mathbb{B}$ with a configuration-induced metric tensor $(\boldsymbol{\varphi}{\downarrow}\mathbf{g})(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{B}^2\,;\Re)$ defined, at any $\mathbf{x} \in \mathbb{B}$, by:

$$(\boldsymbol{\varphi}{\downarrow}\mathbf{g})_{\mathbf{x}}(\mathbf{a},\mathbf{b}) := \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{x})}(T_{\mathbf{x}}\boldsymbol{\varphi}\cdot\mathbf{a}, T_{\mathbf{x}}\boldsymbol{\varphi}\cdot\mathbf{b})\,, \quad \forall\,\mathbf{a},\mathbf{b} \in \mathbb{T}_{\mathbf{x}}\mathbb{B}\,.$$

Here the differential $T_{\mathbf{x}}\boldsymbol{\varphi} \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{B}\,;\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ at $\mathbf{x} \in \mathbb{B}$ of the configuration map is the linear map which transforms a vector $\mathbf{h} \in \mathbb{T}_{\mathbf{x}}\mathbb{B}$ into the corresponding vector $T_{\mathbf{x}}\boldsymbol{\varphi}\cdot\mathbf{h} \in \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}$.

The tangent map $T\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{TB}\,;\mathbb{TS})$ is accordingly defined by

$$(T\boldsymbol{\varphi}\circ\mathbf{v})(\mathbf{x}) := T_{\mathbf{x}}\boldsymbol{\varphi}\cdot\mathbf{v}(\mathbf{x}) \in \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,,$$

for any vector field $\mathbf{v} \in \mathrm{C}^1(\mathbb{B},\mathbb{TB})$.

In differential geometric terms, the tensor field $\boldsymbol{\varphi}{\downarrow}\mathbf{g}$ on $\mathbb{B}$ is called the *pull-back* of the metric tensor field $\mathbf{g}$ on $\boldsymbol{\varphi}(\mathbb{B})$ according to the map $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$. In terms of the tangent map it is defined as

$$(\boldsymbol{\varphi}{\downarrow}\mathbf{g})_{\mathbf{x}}(\mathbf{u}(\mathbf{x}),\mathbf{v}(\mathbf{x})) := \mathbf{g}_{\boldsymbol{\varphi}(\mathbf{x})}((T\boldsymbol{\varphi}\circ\mathbf{u})(\mathbf{x}),(T\boldsymbol{\varphi}\circ\mathbf{v})(\mathbf{x}))\,,$$

for any pair $\mathbf{u},\mathbf{v} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{TB})$ of tangent vector fields.

The metric tensor $\mathbf{g}(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{B}^2\,;\Re)$ may be considered as a linear isomorphism $\mathbf{g}^\flat(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{B}\,;\mathbb{T}^*_{\mathbf{x}}\mathbb{B})$ defined by

$$\langle\,\mathbf{g}^\flat_{\mathbf{x}}(\mathbf{a}),\mathbf{b}\,\rangle := \mathbf{g}_{\mathbf{x}}(\mathbf{a},\mathbf{b})\,, \quad \forall\,\mathbf{a},\mathbf{b} \in \mathbb{T}_{\mathbf{x}}\mathbb{B}\,.$$

A linear operator $\mathbf{A}(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{B}\,;\mathbb{T}_{\mathbf{x}}\mathbb{B})$ is then associated with the tensor $(\mathbf{gA})(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{B}^2\,;\Re)$, defined, at each $\mathbf{x} \in \mathbb{B}$, by the identity:

$$(\mathbf{gA})(\mathbf{a},\mathbf{b}) := \langle\,(\mathbf{g}^\flat\circ\mathbf{A})(\mathbf{a}),\mathbf{b}\,\rangle = \mathbf{g}(\mathbf{Aa},\mathbf{b})\,, \quad \forall\,\mathbf{a},\mathbf{b} \in \mathbb{T}_{\mathbf{x}}\mathbb{B}\,.$$

Accordingly, we have that $(\boldsymbol{\varphi}{\downarrow}\mathbf{g})_{\mathbf{x}} = \mathbf{g}((T_{\mathbf{x}}\boldsymbol{\varphi})^T \cdot T_{\mathbf{x}}\boldsymbol{\varphi})$. The metric change

$$\tfrac{1}{2}\left((\boldsymbol{\varphi}{\downarrow}\mathbf{g})_{\mathbf{x}} - \mathbf{g}_{\mathbf{x}}\right),$$

at a point $\mathbf{x} \in \mathbb{B}$ due to the configuration map $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ is called the Green*'s strain* at that point. It is then defined as (onehalf of) the gap between the $\mathbf{g}$-symmetric operator $(T_{\mathbf{x}}\boldsymbol{\varphi})^T \cdot T_{\mathbf{x}}\boldsymbol{\varphi} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{B}\,;\mathbb{T}_{\mathbf{x}}\mathbb{B})$ and the identity. The reason why it is convenient to adopt a factor $\frac{1}{2}$ will be apparent later on when dealing with equilibrium boundary conditions.

If only lenght measurements are available, the configuration-induced metric tensor $\boldsymbol{\varphi}{\downarrow}\mathbf{g}$ may be evaluated as follows. Firstly we remark that what are needed are the values of the metric tensor $\boldsymbol{\varphi}{\downarrow}\mathbf{g}$ on pairs of vectors taken from a basis, to get the corresponding symmetric Gram matrix:

$$\mathrm{G\scriptstyle RAM}_{\boldsymbol{\varphi}\downarrow\mathbf{g}}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3) = \begin{vmatrix} \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_1,\mathbf{e}_1) & \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_1,\mathbf{e}_2) & \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_1,\mathbf{e}_3) \\ \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_2,\mathbf{e}_1) & \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_2,\mathbf{e}_2) & \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_2,\mathbf{e}_3) \\ \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_3,\mathbf{e}_1) & \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_3,\mathbf{e}_2) & \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_3,\mathbf{e}_3) \end{vmatrix}$$

whose diagonal elements are the squared lenghts of the transformed basis vectors while elements out of diagonal are the inner products between pairs of transformed basis vectors.

All the elements of the Gram matrix may be evaluated by considering the tetrahedron with sides $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ $\mathbf{e}_3 - \mathbf{e}_2$, $\mathbf{e}_3 - \mathbf{e}_1$ and $\mathbf{e}_2 - \mathbf{e}_1$, generated by the basis vectors, and measuring the squared lenghts of the sides of the transformed tetrahedron. Indeed, the parallelogram formula yields:

$$\boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_i+\mathbf{e}_j,\mathbf{e}_i+\mathbf{e}_j) = 2\left(\boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_i,\mathbf{e}_i) + \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_j,\mathbf{e}_j)\right) - \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_i-\mathbf{e}_j,\mathbf{e}_i-\mathbf{e}_j),$$

and the polarization formula gives:

$$4\,\boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_i,\mathbf{e}_j) = \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_i+\mathbf{e}_j,\mathbf{e}_i+\mathbf{e}_j) - \boldsymbol{\varphi}{\downarrow}\mathbf{g}(\mathbf{e}_i-\mathbf{e}_j,\mathbf{e}_i-\mathbf{e}_j).$$

The volume change due to the configuration map $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ is expressed by the jacobian determinant which is the ratio between the configuration-induced volume form and the standard one: $\boldsymbol{\varphi}{\downarrow}\boldsymbol{\mu}_{\mathbf{g}} = J_{\boldsymbol{\varphi}}\,\boldsymbol{\mu}_{\mathbf{g}}$, with $J_{\boldsymbol{\varphi}}(\mathbf{x}) = \det(T_{\mathbf{x}}\boldsymbol{\varphi})$.

Given a basis $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$, we have that

$$\boldsymbol{\mu}^2_{\mathbf{g}}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3) = \det(\mathrm{G\scriptstyle RAM}_{\mathbf{g}}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3))$$
$$(\boldsymbol{\varphi}{\downarrow}\boldsymbol{\mu}_{\mathbf{g}})^2(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3) = \det(\mathrm{G\scriptstyle RAM}_{\boldsymbol{\varphi}\downarrow\mathbf{g}}(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)),$$

see chapter 1, section 1.1.3 on page 9.

Then the absolute value of the jacobian determinant is equal to the square root of the ratio between the determinants of the Gram matrix of any basis with respect to the metrics $\boldsymbol{\varphi}{\downarrow}\mathbf{g}$ and $\mathbf{g}$:

$$J_{\boldsymbol{\varphi}}^2 = \frac{\det(\mathrm{G{\scriptsize RAM}}_{\boldsymbol{\varphi}\downarrow\mathbf{g}}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3))}{\det(\mathrm{G{\scriptsize RAM}}_{\mathbf{g}}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3))}\,.$$

The metric changes at a point of a n-dimensional manifold are described by lenght measurements along the $(n+1)n/2$ sides of a non-degenerated simplex, i.e. a convex polyhedron with $n+1$-vertices in the n-dimensional tangent space. In the 3D euclidean space the simplex is a tetrahedron.

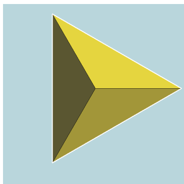

Figure 3.2: tetrahedron

3.2 Kinematics and Equilibrium

A precise statement of the axiom of dynamical equilibrium requires to define in a proper way the linear kinematical space made up of the velocities of the virtual motions that the body is allowed to undergo at any fixed instant of time.

The very concept of *force system* is based on the specification of the kinematical space and on its topological properties since force systems are work-conjugate to the virtual velocities and the relevant the duality pairing is called the *virtual work*.

The ideas underlying the definition of the kinematical space are twofold. From a physical point of view we must recognize that the body under investigation is chosen in an arbitrary way and hence any kinematical definition must be reproducible on any part of a given one.

On the mathematical side the requirement is that the topological properties must ensure the existence of boundary traces and a basic closedness property.

Dually force systems are assumed to be bounded linear functionals over the fields of the topological kinematical space. This means that the virtual work of a given force system may be made as small as desired by taking the virtual velocity field in a sufficiently small neighbourhood of the null field.

Mathematical minded people will find a brief but precise account of the relevant aspects in section 3.5.2 and in the references quoted therein.

Preliminarily, in section 3.2.1, we will adopt a heuristic approach to provide the basic ideas without the burden of functional analysis concepts and tools that are needed to appreciate the mathematical treatment.

3.2.1 Basic ideas

The reproducibility requirement is fulfilled by allowing for virtual velocities to be discontinuous on the borders of a patchwork made of an arbitrary but finite number of sub-bodies. In this way it is possible to apply the equilibrium condition to any part of any body.

This is the kinematic counterpart of the well-known EULER-CAUCHY principle stating that, if a body is in equilibrium, then any of its parts is also in equilibrium.

Real bodies may usually be considered as composed by a finite number of continuous simple sub-bodies in which the admissible velocities are required to have no discountinuity surfaces. Moreover, on the boundary of these simple sub-bodies, the admissible velocities are subject to prescribed linear or affine conditions. All these are called constraint conditions.

More complex, nonlinear conditions are also considered and imposed as relations between dual entities described by multivalued maps. A well-developed theory exists for multivalued maps with maximal monotone graphs. Linear or affine relations are described by constant-valued monotone multivalued maps. These more general conditions are called constitutive laws.

The velocities which meet the continuity constraint and homogeneous boundary constraint are assumed to belong to a linear space, the space of conforming velocities. If this space is finite dimensional, any basis is called a set of degrees of freedom.

Force system are defined as dual entities of the virtual velocities performing virtual power in a linear fashion. They can be added one another and multiplied by reals, thus forming a linear space.

The physical idea of frictionless, firm and bilateral constraints, is modeled by requiring that the reactive force systems exerted by the constraints must perform a null virtual power for any conforming virtual velocity field.

In imposing the equilibrium condition on a system of forces, we may consider both conforming or non-conforming virtual velocities. To detect and evaluate a reactive force system we must consider a non-conforming rigid virtual velocity field and impose that the virtual power performed by active and reactive force systems vanishes. This approach provides sufficient informations on reactive force systems only for some special one-dimensional structural models composed by beam elements, referred to as non-redundant structural models.

In the general case, constitutive laws describing the material behavior must be provided to get further informations able to detect the reactive force systems.

Mechanics is founded on the concept of equilibrium first enunciated in variational terms by JOHANN BERNOULLI in 1717 in a letter to PIERRE VARIGNON. This could be considered at right the cornerstone for the beginning of a matematical theory of mechanics.

Figure 3.3: Johann Bernoulli (1667 - 1748)

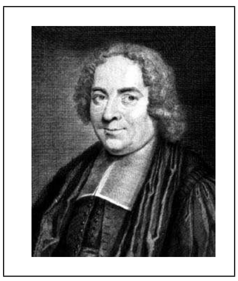

Figure 3.4: Pierre Varignon (1654 - 1722)

In its modern formulation, the *axiom of equilibrium* states that:

- At any configuration $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ of a body $\mathbb{B}$, a system of forces acting on it is in equilibrium if it performs a null virtual power for any virtual motion of the body which starts as an *infinitesimal isometry*.

A virtual motion is called an infinitesimal isometry if it causes no rate of change of the metric properties of the body, that is the lenght of any path drawn in the body has a vanishing time-rate of variation.

Let us denote by $\mathrm{RIG}(\boldsymbol{\varphi}(\mathbb{B}))$, or simply RIG, the linear space of virtual infinitesimal isometries, also called rigid-body virtual velocities, and by $\mathbf{f}$ the force system acting on the body, at the current placement $\boldsymbol{\varphi}(\mathbb{B})$.

A formal statement of the axiom of dynamical equilibrium is then expressed by the variational condition

$$\langle \mathbf{f}, \mathbf{v} \rangle = 0\,, \quad \forall\, \mathbf{v} \in \mathrm{RIG}(\boldsymbol{\varphi}(\mathbb{B}))\,.$$

A celebrated kinematical result, stated, in the context of euclidean space, by LEONHARD EULER in the middle of the XVII century and extended to riemannian manifolds by WILHELM KILLING in the last decades of the XIX century, shows that infinitesimal isometries of a body are velocity fields characterized by the vanishing of the symmetric part of their spatial derivative in the a body. This implies that every connected component of the body undergoes a motion with a constant spatial derivative. The issue is discussed in detail in the next section.

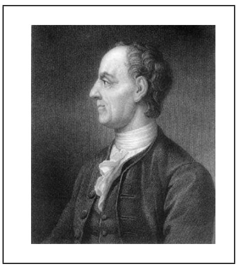

Figure 3.5: Leonhard Euler (1707 - 1783)

3.2.2 Kelvin, Helmholtz and Lagrange's theorems

To provide a mathematical definition of a virtual infinitesimal isometry, let us consider a virtual spatial flow $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$ dragging the body $\mathbb{B}$ in the ambient space.

The virtual velocity field $\mathbf{v} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{TS})$ of the body at the placement $\boldsymbol{\varphi}(\mathbb{B})$ under the virtual flow $\mathbf{Fl}^{\mathbf{v}}_{\lambda}$, is given by: $\mathbf{v} = \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda}$.

A virtual infinitesimal isometry is characterized by the vanishing of the LIE derivative of the metric tensor along the virtual spatial flow:

$$\mathcal{L}_{\mathbf{v}}\mathbf{g} := \partial_{\lambda=0} \, \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow \mathbf{g} = 0 \, .$$

In a riemannian manifold $\{\mathbb{S}\,,\mathbf{g}\}$ with the LEVI-CIVITA connection ∇, the LIE derivative of the metric tensor, along a vector field $\mathbf{v} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{TS})$, is provided by EULER*'s distorsion rate formula*:

$$\tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g} := \tfrac{1}{2}\partial_{\lambda=0} \, \mathbf{Fl}^{\mathbf{v}}_{\lambda}\!\downarrow \mathbf{g} = \mathbf{g} \circ (\mathrm{sym}\, \nabla\mathbf{v}) \, ,$$

(see secton 1.9.4 on page 202). In particular this formula holds in the euclidean space $\{\mathbb{S}\,,\mathrm{CAN}\}$ with the canonical connection induced by translations.

In a ambient manifold $\mathbb{S}$ endowed with an affine connection ∇ let us consider a motion described by a flow $\mathbf{Fl}^{\mathbf{v}}_{\tau,t}\mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$ associated with a time-dependent velocity vector field $\mathbf{v}_t \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$.

- The *acceleration field* is defined, according to EULER*'s formula for the acceleration* (1770), by the *material time derivative* of the velocity vector field along its flow:

$$\mathbf{a}_t = \nabla_{t,\mathbf{v}_t}\mathbf{v}_t := \partial_{\tau=t}\, \mathbf{v}_\tau + \nabla_{\mathbf{v}_t}\mathbf{v}_t \, .$$

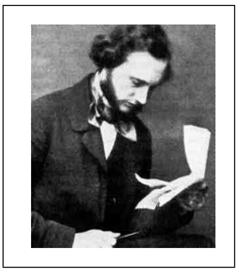

Figure 3.6: William Thomson, lord Kelvin (1824 - 1907)

Theorem 3.2.1 (Kelvin's kinematical theorem) *Let* $\{\mathbb{S}\,,\mathbf{g}\}$ *be a riemannian manifold with the* LEVI-CIVITA *connection. In a time-dependent flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$ *the time rate of the circuitation of the velocity field around any loop* $\mathbf{c} \in \mathrm{C}^1(I\,;\mathbb{S})$ *is equal to the circuitation of the acceleration:*

$$\partial_{\tau=t} \oint_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\mathbf{c})} \mathbf{g}\mathbf{v}_t = \oint_{\mathbf{c}} \mathbf{g}\mathbf{a}_t \, .$$

In particular, if the acceleration is the gradient of a potential: $\mathbf{a}_t = \mathbf{g}^\sharp \circ df_t$, *then* $\mathbf{g}\mathbf{a}_t = \mathbf{g}^\flat \circ \mathbf{a}_t = df_t$ *and the circuitous integral at the r.h.s. vanishes. Thus the circuitation of the velocity field, around any closed loop dragged by the flow, is a costant of the motion.*

Proof. By REYNOLDS*'s transport theorem* we get:

$$\partial_{\tau=t} \oint_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\mathbf{c})} \mathbf{g}\mathbf{v}_t = \oint_{\mathbf{c}} \mathcal{L}_{t,\mathbf{v}_t}(\mathbf{g}\mathbf{v}_t) = \oint_{\mathbf{c}} \mathbf{g}(\partial_{\tau=t}\,\mathbf{v}_\tau) + \oint_{\mathbf{c}} \mathcal{L}_{\mathbf{v}_t}(\mathbf{g}\mathbf{v}_t)\,.$$

Moreover, EULER*'s distorsion rate formula* tells us that

$$\mathcal{L}_{\mathbf{v}_t}(\mathbf{g}\mathbf{v}_t) = (\mathcal{L}_{\mathbf{v}_t}\mathbf{g})\mathbf{v}_t + \mathbf{g}(\mathcal{L}_{\mathbf{v}_t}\mathbf{v}_t) = (\mathcal{L}_{\mathbf{v}_t}\mathbf{g})\mathbf{v}_t = \mathbf{g}(\nabla\mathbf{v}_t)\cdot\mathbf{v}_t + \mathbf{g}(\nabla\mathbf{v}_t)^T\cdot\mathbf{v}_t\,,$$

and being $\nabla\mathbf{g} = 0$, we have:

$$\mathbf{g}((\nabla\mathbf{v}_t)^T\cdot\mathbf{v}_t,\mathbf{w}) = \mathbf{g}(\mathbf{v}_t,\nabla_{\mathbf{w}}\mathbf{v}_t) = \tfrac{1}{2}d_{\mathbf{w}}(\mathbf{g}\mathbf{v}_t\cdot\mathbf{v}_t)\,,$$

so that

$$\oint_{\mathbf{c}} \mathbf{g}((\nabla\mathbf{v}_t)^T\cdot\mathbf{v}_t) = \tfrac{1}{2}\oint_{\mathbf{c}} d(\mathbf{g}(\mathbf{v}_t,\mathbf{v}_t)) = 0\,,$$

and the result follows. ■

Corollary 3.2.1 (Flux of the vorticity) *Let* $\{\mathbb{S}\,,\mathbf{g}\}$ *be a riemannian manifold with the* LEVI-CIVITA *connection. In a time-dependent flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$, *if the acceleration is the gradient of a potential, the flux of the vorticity* $\operatorname{rot}\mathbf{v}_t$ *through any surface dragged by the flow, is a costant of the motion.*

Proof. Applying Theorem 3.2.1 to the boundary of a surface Σ , by STOKES formula we get

$$\begin{aligned}\int_\Sigma \boldsymbol{\mu}\cdot\operatorname{rot}\mathbf{a}_t &= \int_\Sigma d(\mathbf{g}\mathbf{a}_t) = \oint_{\partial\Sigma} \mathbf{g}\mathbf{a}_t = \partial_{\tau=t}\oint_{\partial\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \mathbf{g}\mathbf{v}_t \\ &= \partial_{\tau=t}\int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} d(\mathbf{g}\mathbf{v}_\tau) = \partial_{\tau=t}\int_{\mathbf{Fl}^{\mathbf{v}}_{\tau,t}(\Sigma)} \boldsymbol{\mu}\cdot(\operatorname{rot}\mathbf{v}_\tau)\,,\end{aligned}$$

and the result follows. ■

In a motion described by a flow $\mathbf{Fl}^{\mathbf{v}}_{\tau,t}\mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$ associated with a time-dependent velocity vector field $\mathbf{v}_t \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$, the material time derivative of a vector field $\mathbf{u}_t \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$ along the flow is defined by:

$$\dot{\mathbf{u}}_t := \partial_{\tau=t}\,\mathbf{u}_\tau + \nabla_{\mathbf{v}_t}\mathbf{u}_t\,.$$

Definition 3.2.1 (Material line) *The integral curve of a time-dependent vector field $\mathbf{u}_t \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$ is a material line if the vector field is dragged by the flow describing the motion, to within a proportionality, due to the arbitrarity of the parametrization, that is:*

$$\boldsymbol{\mu}\cdot\mathbf{Fl}^{\mathbf{v}}_{t,\tau}\!\uparrow\!\mathbf{u}_\tau\cdot\mathbf{u}_t = 0\,,$$

which expresses the proportionality between the vector field and its pull-back along the flow.

Figure 3.7: Hermann Ludwig Ferdinand von Helmholtz (1821 - 1894)

Theorem 3.2.2 (Helmholtz's kinematical theorem) *Let $\mathbb{S}$ be a configuration manifold endowed with a torsionless connection ∇. In a motion described by the flow $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$, the integral curve of a time-dependent vector field $\mathbf{u}_t \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$ is a material line if and only if*

$$\boldsymbol{\mu}\cdot\mathcal{L}_{t,\mathbf{v}_t}\mathbf{u}_t\cdot\mathbf{u}_t = \boldsymbol{\mu}\cdot(\dot{\mathbf{u}}_t - \nabla_{\mathbf{u}_t}\mathbf{v}_t)\cdot\mathbf{u}_t = 0\,.$$

Proof. Taking the time derivative, the materiality condition becomes

$$\partial_{\tau=t}\,\boldsymbol{\mu}\cdot\mathbf{Fl}^{\mathbf{v}}_{t,\tau}\!\uparrow\!\mathbf{u}_\tau\cdot\mathbf{u}_t = \boldsymbol{\mu}\cdot\mathcal{L}_{t,\mathbf{v}_t}\mathbf{u}_t\cdot\mathbf{u}_t = 0\,.$$

and the vanishing of the torsion tells us that $\mathcal{L}_{\mathbf{v}_t}\mathbf{u}_t = \nabla_{\mathbf{v}_t}\mathbf{u}_t - \nabla_{\mathbf{u}_t}\mathbf{v}_t$. Then

$$\mathcal{L}_{t,\mathbf{v}_t}\mathbf{u}_t = \partial_{\tau=t}\,\mathbf{u}_\tau + \nabla_{\mathbf{v}_t}\mathbf{u}_t - \nabla_{\mathbf{u}_t}\mathbf{v}_t = \dot{\mathbf{u}}_t - \nabla_{\mathbf{u}_t}\mathbf{v}_t\,.$$

and the result is proven. ■

Corollary 3.2.2 (Materiality of vortex lines) *If the acceleration is the gradient of a potential, the vortex lines are material lines.*

Proof. By noting that

$$\boldsymbol{\mu} \cdot \operatorname{rot} \mathbf{a}_t = \mathcal{L}_{t,\mathbf{v}_t}(\boldsymbol{\mu} \cdot \operatorname{rot} \mathbf{v}_t) = (\mathcal{L}_{\mathbf{v}_t}\boldsymbol{\mu}) \cdot \operatorname{rot} \mathbf{v}_t + \boldsymbol{\mu} \cdot \mathcal{L}_{t,\mathbf{v}_t} \operatorname{rot} \mathbf{v}_t \,,$$

by the skew-symmetry of $\mathcal{L}_{\mathbf{v}_t}\boldsymbol{\mu}$ we infer that $(\mathcal{L}_{\mathbf{v}_t}\boldsymbol{\mu}) \cdot \operatorname{rot} \mathbf{v}_t \cdot \operatorname{rot} \mathbf{v}_t = 0$ and hence that

$$\boldsymbol{\mu} \cdot \operatorname{rot} \mathbf{a}_t \cdot \operatorname{rot} \mathbf{v}_t = \boldsymbol{\mu} \cdot \mathcal{L}_{t,\mathbf{v}_t} \operatorname{rot} \mathbf{v}_t \cdot \operatorname{rot} \mathbf{v}_t \,,$$

and the result follows from theorem 3.2.2. ■

Definition 3.2.2 (Material surface) *A surface* $\Sigma \subset \mathbb{S}$ *described by a time-dependent condition* $f(\mathbf{x}, t) = 0$ *is a material surface if the scalar function* $f \in \mathrm{C}^1(\mathbb{S}\,;\Re)$ *is an invariant of the flow, that is:*

$$f_t = \mathbf{Fl}^{\mathbf{v}}_{t,\tau}\!\uparrow\! f_\tau = 0 \,.$$

Theorem 3.2.3 (Lagrange's kinematical theorem) *Let* $\mathbb{S}$ *be the configuration manifold. In a motion described by the flow* $\mathbf{Fl}^{\mathbf{v}}_{\tau,t} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$*, the surface described by the time-dependent condition* $f(\mathbf{x}, t) = 0$ *is a material surface iff*

$$\mathcal{L}_{t,\mathbf{v}_t} f = \partial_{\tau=t}\, f + \mathcal{L}_{\mathbf{v}_t} f = 0 \,.$$

Proof. Taking the time derivative, the materiality condition becomes

$$\partial_{\tau=t}\, \mathbf{Fl}^{\mathbf{v}}_{t,\tau}\!\uparrow\! f_\tau = \partial_{\tau=t}\, f_\tau + \mathcal{L}_{\mathbf{v}_t} f = 0 \,.$$

and the result is proven. ■

3.2.3 Euler's kinematical theorem

Euler's condition for an infinitesimal isometry is that:

$$\textsc{Eul}(\mathbf{v}) := \operatorname{sym} \nabla \mathbf{v} = 0 \,.$$

In the euclidean space $\{\mathbb{S}\,, \mathbf{g}\}$ Euler's condition implies that the skew-symmetric part of the derivative $\nabla \mathbf{v}(\mathbf{x}) \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{S}\,;\mathbb{T}_{\mathbf{x}}\mathbb{S})$ is constant in each connected body, a consequence of the following pointwise result.

Theorem 3.2.4 (Euler's kinematical theorem) *The vanishing, at a point* $\mathbf{x} \in \boldsymbol{\varphi}(\mathbb{B})$ *, of the derivative of the symmetric part* $\mathrm{sym}\, \nabla\mathbf{v}$ *of the gradient of a vector field* $\mathbf{v} \in \mathrm{C}^2(\boldsymbol{\varphi}(\mathbb{B})\,; \mathbb{TS})$ *implies the vanishing of the derivative of the gradient at the same point, i.e.:*

$$\nabla\, (\mathrm{sym}\, \nabla\mathbf{v})(\mathbf{x}) = 0 \implies \nabla^2\mathbf{v}(\mathbf{x}) = 0\,.$$

Proof. Let $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h} \in \mathbb{TS}$ be arbitrary constant vector fields and denote by $d_{\mathbf{h}}$ the directional derivative along $\mathbf{h} \in \mathbb{TS}$. By assumption:

$$d_{\mathbf{h}}\mathbf{g}\,(d_{\mathbf{h}_1}\mathbf{v}, \mathbf{h}_2) + \mathbf{g}\,(d_{\mathbf{h}_2}\mathbf{v}, \mathbf{h}_1) = \mathbf{g}\,(d^2_{\mathbf{h}\mathbf{h}_1}\mathbf{v}, \mathbf{h}_2) + \mathbf{g}\,(d^2_{\mathbf{h}\mathbf{h}_2}\mathbf{v}, \mathbf{h}_1) = 0\,.$$

By substituting $\mathbf{h}_1$ with $\mathbf{h}$ and $\mathbf{h}_2$ with $\mathbf{h}$, we get two more relations, so that

$$\begin{aligned} i) \quad & \mathbf{g}\,(d^2_{\mathbf{h}\mathbf{h}_1}\mathbf{v}, \mathbf{h}_2) + \mathbf{g}\,(d^2_{\mathbf{h}\mathbf{h}_2}\mathbf{v}, \mathbf{h}_1) = 0\,, \\ ii) \quad & \mathbf{g}\,(d^2_{\mathbf{h}_1\mathbf{h}}\mathbf{v}, \mathbf{h}_2) + \mathbf{g}\,(d^2_{\mathbf{h}_1\mathbf{h}_2}\mathbf{v}, \mathbf{h}) = 0\,, \\ iii) \quad & \mathbf{g}\,(d^2_{\mathbf{h}_2\mathbf{h}_1}\mathbf{v}, \mathbf{h}) + \mathbf{g}\,(d^2_{\mathbf{h}_2\mathbf{h}}\mathbf{v}, \mathbf{h}_1) = 0\,. \end{aligned}$$

Since the second directional derivative is symmetric, it follows that

$$\mathbf{g}\,(d^2_{\mathbf{h}_1\mathbf{h}_2}\mathbf{v}, \mathbf{h}) = 0\,, \quad \forall\, \mathbf{h}_1, \mathbf{h}_2, \mathbf{h} \in \mathbb{TS}\,,$$

and hence $\nabla^2\mathbf{v} = 0$. ■

Euler's kinematical theorem provides a simple representation formula for infinitesimal isometries, as illustrated below.

Let the speed $\mathbf{v}$ be regular (say in $\mathrm{C}^2(\boldsymbol{\varphi}(\mathbb{B})\,; \mathbb{TS})$) in a connected body $\boldsymbol{\varphi}(\mathbb{B})$. Then, from the condition $\mathrm{sym}\, \nabla\mathbf{v}(\mathbf{x}) = 0$ for any $\mathbf{x} \in \boldsymbol{\varphi}(\mathbb{B})$, we infer that $\nabla\mathbf{v}(\mathbf{x}) = \mathbf{W}$, with $\mathbf{W}$ a skew-symmetric operator. An infinitesimal isometry $\mathbf{v}$ is then characterized by the following equivalent properties:

$$\begin{aligned} i) \quad & \mathbf{v}(\mathbf{x}) - \mathbf{v}(\mathbf{y}) = \mathbf{W}\,(\mathbf{x} - \mathbf{y})\,, \quad \forall\, \mathbf{x}, \mathbf{y} \in \boldsymbol{\varphi}(\mathbb{B})\,, \\ ii) \quad & \mathbf{g}\,(\mathbf{v}(\mathbf{x}) - \mathbf{v}(\mathbf{y}), \mathbf{x} - \mathbf{y}) = 0\,, \quad \forall\, \mathbf{x}, \mathbf{y} \in \boldsymbol{\varphi}(\mathbb{B})\,, \qquad \text{equiprojectivity}\,, \\ iii) \quad & \mathbf{v}(\mathbf{x}) = \mathbf{v}_0 + \mathbf{W}\,(\mathbf{x} - \mathbf{x}_0) = \mathbf{v}_0 + \boldsymbol{\omega} \times (\mathbf{x} - \mathbf{x}_0)\,, \quad \forall\, \mathbf{x} \in \boldsymbol{\varphi}(\mathbb{B})\,. \end{aligned}$$

To show that $ii)$ implies $i)$, we rewrite it as

$$\mathbf{g}\,(\mathbf{v}(\mathbf{x} + \lambda\mathbf{h}) - \mathbf{v}(\mathbf{x}), \mathbf{h}) = 0\,, \quad \forall\, \mathbf{h} \in \boldsymbol{\varphi}(\mathbb{B})\,, \quad \lambda \in \Re\,,$$

then take the derivative $\partial_{\lambda=0}$ to get

$$\mathbf{g}\,(\nabla\mathbf{v}(\mathbf{x}) \cdot \mathbf{h}, \mathbf{h}) = 0\,, \quad \forall\, \mathbf{h} \in \boldsymbol{\varphi}(\mathbb{B}) \iff \mathrm{sym}\, \nabla\mathbf{v}(\mathbf{x}) = 0\,.$$

The last formula provides the classical representation of a simple infinitesimal isometry as the sum of two vector fields:

- a *translational velocity* field with speed $\mathbf{v}_0$, characterized by the linear operator $\mathrm{TRA} \in BL(\mathbb{TS}\,; \mathrm{C}^\infty(\boldsymbol{\varphi}(\mathbb{B})\,; \mathbb{TS}))$ defined by

$$\mathrm{TRA}(\mathbf{v}_0)(\mathbf{x}) = \mathbf{v}_0\,, \quad \forall\, \mathbf{x} \in \boldsymbol{\varphi}(\mathbb{B})\,,$$

- plus a *rotational velocity* field about the pole $\mathbf{x}_0$ with angular speed $\boldsymbol{\omega}$, characterized by the linear operator $\mathrm{ROT}_{\mathbf{x}_0} \in BL(\mathbb{TS}\,; \mathrm{C}^\infty(\boldsymbol{\varphi}(\mathbb{B})\,; \mathbb{TS}))$ defined by

$$\mathrm{ROT}_{\mathbf{x}_0}(\boldsymbol{\omega})(\mathbf{x}) = \boldsymbol{\omega} \times (\mathbf{x} - \mathbf{x}_0)\,, \quad \forall\, \mathbf{x} \in \boldsymbol{\varphi}(\mathbb{B})\,.$$

The *angular speed* $\boldsymbol{\omega}$ is in a one-to-one relation with skew-symmetric tensor $\mathbf{W}$ by the formula $\mathbf{W}\mathbf{h} = \boldsymbol{\omega} \times \mathbf{h}\,, \quad \forall\, \mathbf{h} \in \mathbb{TS}$ which is equivalent to

$$\boldsymbol{\mu}_{\mathbf{g}} \cdot \boldsymbol{\omega} = \mathbf{g} \cdot \mathbf{W}\,.$$

To prove this, recall that the cross product $\times$ between vectors is defined by the identity

$$\boldsymbol{\mu}_{\mathbf{g}}(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{g}(\mathbf{a} \times \mathbf{b}, \mathbf{c})\,, \quad \forall\, \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{TS} \iff (\boldsymbol{\mu}_{\mathbf{g}} \cdot \mathbf{a}) \cdot \mathbf{b} = \mathbf{g} \cdot (\mathbf{a} \times \mathbf{b})\,,$$

so that, $\forall\, \mathbf{h}, \mathbf{c} \in \mathbb{TS}$, we have

$$(\boldsymbol{\mu}_{\mathbf{g}}\boldsymbol{\omega})(\mathbf{h}, \mathbf{c}) = \boldsymbol{\mu}_{\mathbf{g}}(\boldsymbol{\omega}, \mathbf{h}, \mathbf{c}) = \mathbf{g}(\boldsymbol{\omega} \times \mathbf{h}, \mathbf{c}) = \mathbf{g}(\mathbf{W}\mathbf{h}, \mathbf{c}) = (\mathbf{g}\mathbf{W})(\mathbf{h}, \mathbf{c})\,,$$

which ends the proof.

3.2.4 Cardinal equations of statics

The virtual work, performed by a system of forces acting on a body undergoing a simple infinitesimal isometry, can be expressed in terms of two characteristic vectors.

To see this, we define the linear operators, RES, adjoint of TRA, and $\mathrm{MOM}_{\mathbf{x}_0}$, adjoint of $\mathrm{ROT}_{\mathbf{x}_0}$, by the identities:

$$\begin{aligned} \langle \mathbf{f}, \mathrm{TRA}(\mathbf{v}_0) \rangle &= \mathbf{g}\,(\mathrm{RES}(\mathbf{f}), \mathbf{v}_0)\,, && \forall\, \mathbf{v}_0 \in \mathbb{TS}\,, \\ \langle \mathbf{f}, \mathrm{ROT}_{\mathbf{x}_0}(\boldsymbol{\omega}) \rangle &= \mathbf{g}\,(\mathrm{MOM}_{\mathbf{x}_0}(\mathbf{f}), \boldsymbol{\omega})\,, && \forall\, \boldsymbol{\omega} \in \mathbb{TS}\,. \end{aligned}$$

The vectors $\text{RES}(\mathbf{f})$ and $\text{MOM}_{\mathbf{x}_0}(\mathbf{f})$ are respectively called the *resultant force* and the *resultant moment* of the force system $\mathbf{f}$. Then, being

$$\langle \mathbf{f}, \mathbf{v} \rangle = \langle \mathbf{f}, \text{TRA}(\mathbf{v}_0) \rangle + \langle \mathbf{f}, \text{ROT}_{\mathbf{x}_0}(\boldsymbol{\omega}) \rangle = \mathbf{g}\,(\text{RES}(\mathbf{f}), \mathbf{v}_0) + \mathbf{g}\,(\text{MOM}_{\mathbf{x}_0}(\mathbf{f}), \boldsymbol{\omega})\,,$$

the vanishing of the virtual work for any simple infinitesimal isometry is equivalent to require that

$$\text{RES}(\mathbf{f}) = 0\,, \quad \text{MOM}_{\mathbf{x}_0}(\mathbf{f}) = 0\,.$$

These are called the *cardinal equations of statics.*

3.3 Conservation of mass

Let us consider in the euclidean space $\mathbb{S}$ a continuous body, whose reference placement is an embedded submanifold $\mathbb{B} \subset \mathbb{S}$, undergoing a motion $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ with $t \in I$, an open time interval, and $\boldsymbol{\varphi}_0$ the identity.

The evolution of the body in space, defined by $\boldsymbol{\varphi}_{\tau,t} = \boldsymbol{\varphi}_\tau \circ \boldsymbol{\varphi}_t^{-1}$, maps the position of a particle at time t into its position at time τ.

The corresponding *trajectory* tracked by the body in the time interval I is the dragged submanifold

$$\text{TRA}_I(\boldsymbol{\varphi}) := \bigcup_{t \in I} \boldsymbol{\varphi}_t(\mathbb{B})\,.$$

The inertial and gravitational properties of the body are described by a time-dependent positive scalar field, the mass-density per unit volume of the current placement $\boldsymbol{\varphi}_t(\mathbb{B})$.

The spatial description of the mass-density along the trajectory is a scalar field $\rho_t \in \mathrm{C}^1(\text{TRA}_I(\boldsymbol{\varphi})\,;\Re)$. The corresponding material description is the scalar field provided by the composition $\rho_{0t} = \rho_t \circ \boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{B}\,;\Re)$.

The total mass of the body at time t is given by

$$M_t = \int_{\boldsymbol{\varphi}_t(\mathbb{B})} \rho_t\, \boldsymbol{\mu}\,,$$

where $\boldsymbol{\mu}$ is the standard volume form in the euclidean space $\mathbb{S}$.

The *principle of conservation of mass* states that for all bodies

$$M_\tau = M_t\,, \qquad \forall\, \tau, t \in I\,.$$

Let $\mathbf{\Omega} = \boldsymbol{\varphi}_t(\mathbb{B}) \subset \mathbb{S}$ be the placement of the body at time $t \in I$.

Introducing the time-dependent mass-form $\mathbf{m}_t = \rho_t\, \boldsymbol{\mu}$ and recalling the formula relating the integrals over diffeomorphic manifolds, we express the principle of conservation of mass as:

$$\int_{\mathbf{\Omega}} \mathbf{m}_t = \int_{\boldsymbol{\varphi}_{\tau,t}(\mathbf{\Omega})} \boldsymbol{\varphi}_{\tau,t}\!\uparrow\!\mathbf{m}_t = \int_{\boldsymbol{\varphi}_{\tau,t}(\mathbf{\Omega})} \mathbf{m}_\tau \,.$$

Being valid for all bodies, the principle of conservation of mass can be localized as follows. Since

$$\begin{aligned} \boldsymbol{\varphi}_{\tau,t}\!\uparrow\!(\rho_t\boldsymbol{\mu}) &= (\boldsymbol{\varphi}_{\tau,t}\!\uparrow\!\rho_t)(\boldsymbol{\varphi}_{\tau,t}\!\uparrow\!\boldsymbol{\mu}) \\ \boldsymbol{\varphi}_{\tau,t}\!\uparrow\!\boldsymbol{\mu} &= \det(d\boldsymbol{\varphi}_{t,\tau})\, \boldsymbol{\mu} \\ \boldsymbol{\varphi}_{\tau,t}\!\uparrow\!\rho_t &= \rho_t \circ \boldsymbol{\varphi}_{t,\tau} \,, \end{aligned}$$

we get

$$\boldsymbol{\varphi}_{\tau,t}\!\uparrow\!\mathbf{m}_t = \mathbf{m}_\tau \iff \rho_t \circ \boldsymbol{\varphi}_{t,\tau} = \det(d\boldsymbol{\varphi}_{\tau,t})\, \rho_\tau \,.$$

The principle of conservation of mass may then be formulated by stating that

- the mass-form is dragged by the flow.

In terms of time-rates the principle of conservation of mass states that, along any motion of any body at any instant, the time derivative of the total mass must vanish.

Let $\mathbf{v}_t \in \mathrm{C}^1(\boldsymbol{\varphi}_t(\mathbb{B})\,; \mathbb{TS})$ be the velocity of the motion. By Reynolds transport theorem we get

$$\partial_{\tau=t}\; M_\tau = \partial_{\tau=t} \int_{\boldsymbol{\varphi}_{\tau,t}(\mathbf{\Omega})} \mathbf{m}_\tau = \int_{\mathbf{\Omega}} \mathcal{L}_{t,\mathbf{v}}\mathbf{m}_t = \int_{\mathbf{\Omega}} (\partial_{\tau=t}\; \mathbf{m}_\tau + \mathcal{L}_{\mathbf{v}}\mathbf{m}_t) = 0 \,,$$

where $\boldsymbol{\varphi}_{t,\tau}\!\uparrow\!\mathbf{m}_t$ denotes the pull-back of the mass-form, with the mass density frozen at time t, and

- $\mathcal{L}_{\mathbf{v}}\, \mathbf{m}_t := \partial_{\tau=t}\; \boldsymbol{\varphi}_{t,\tau}\!\uparrow\!\mathbf{m}_t$ is the convective (or Lie) derivative along the flow. For scalar spatial fields it coincides with the directional derivative along the flow.

- $\mathcal{L}_{t,\mathbf{v}}\, \mathbf{m}_t := \partial_{\tau=t}\; \mathbf{m}_\tau + \mathcal{L}_{\mathbf{v}}\mathbf{m}_t$ is the convective time-derivative of the mass-form. For scalar spatial fields it coincides with the material time-derivative.

The local version of the principle of conservation of mass in rate form, amounts to require that, along any trajectory of the body $\mathbb{B}$ the convective time-derivative of the mass-form vanishes at any time:

$$\mathcal{L}_{t,\mathbf{v}}\, \mathbf{m}_t = \mathcal{L}_{t,\mathbf{v}}\, (\rho_t\, \boldsymbol{\mu}) = 0\,.$$

To express the principle in terms of the scalar mass-density, we recall that the convective time-derivative (or material time-derivative) of the mass-density is given by

$$\mathcal{L}_{t,\mathbf{v}}\, \rho_t := \partial_{\tau=t}\ \rho_\tau + \mathcal{L}_{\mathbf{v}} \rho_t = \partial_{\tau=t}\ \rho_\tau + d_{\mathbf{v}} \rho_t\,.$$

Then, being by definition $\mathcal{L}_{\mathbf{v}}\, \boldsymbol{\mu} = (\operatorname{div} \mathbf{v})\, \boldsymbol{\mu}$, the principle of conservation of mass is written as:

$$\begin{aligned}
\partial_{\tau=t}\ \int_{\boldsymbol{\varphi}_{\tau,t}(\boldsymbol{\Omega})} \rho_\tau\, \boldsymbol{\mu} &= \int_{\boldsymbol{\Omega}} \mathcal{L}_{t,\mathbf{v}}\, (\rho_t\, \boldsymbol{\mu}) \\
&= \int_{\boldsymbol{\Omega}} \partial_{\tau=t}\, \rho_\tau\, \boldsymbol{\mu} + \int_{\boldsymbol{\Omega}} \mathcal{L}_{\mathbf{v}}\, (\rho_t\, \boldsymbol{\mu}) \\
&= \int_{\boldsymbol{\Omega}} (\partial_{\tau=t}\ \rho_\tau + \mathcal{L}_{\mathbf{v}}\, \rho_t)\, \boldsymbol{\mu} + \int_{\boldsymbol{\Omega}} \rho_t\, \mathcal{L}_{\mathbf{v}}\, \boldsymbol{\mu} \\
&= \int_{\boldsymbol{\Omega}} (\mathcal{L}_{t,\mathbf{v}}\, \rho_t + \rho_t \operatorname{div} \mathbf{v})\, \boldsymbol{\mu}\,.
\end{aligned}$$

or, recalling that $\mathcal{L}_{\mathbf{v}}\, (\rho\, \boldsymbol{\mu}) = \mathcal{L}_{(\rho \mathbf{v})}\, \boldsymbol{\mu}$, as

$$\begin{aligned}
\partial_{\tau=t}\ \int_{\boldsymbol{\varphi}_{\tau,t}(\boldsymbol{\Omega})} \rho_\tau\, \boldsymbol{\mu} &= \int_{\boldsymbol{\Omega}} \mathcal{L}_{t,\mathbf{v}}\, (\rho_t\, \boldsymbol{\mu}) \\
&= \int_{\boldsymbol{\Omega}} (\partial_{\tau=t}\ \rho_\tau + \mathcal{L}_{(\rho_t \mathbf{v})})\, \boldsymbol{\mu} \\
&= \int_{\boldsymbol{\Omega}} (\partial_{\tau=t}\ \rho_\tau + \operatorname{div} (\rho_t\, \mathbf{v}))\, \boldsymbol{\mu} \\
&= \int_{\boldsymbol{\Omega}} \partial_{\tau=t}\ \rho_\tau\, \boldsymbol{\mu} + \int_{\partial\boldsymbol{\Omega}} \rho_t\, \mathbf{g}(\mathbf{v}, \mathbf{n})\ (\boldsymbol{\mu}\mathbf{n}) = 0\,.
\end{aligned}$$

Again, by localizing, we infer the following equivalent forms of the differential law of mass conservation:

$$\mathcal{L}_{t,\mathbf{v}}\,\rho + \rho\,\mathrm{div}\,\mathbf{v} = 0 \iff \partial_t\rho + \mathrm{div}\,(\rho\,\mathbf{v}) = 0\,.$$

By taking account of the positivity of the mass-density, it follows that the velocity field is solenoidal at an instant of time iff the total time-derivative of the mass-density along the motion vanishes at that time i.e.

$$\mathcal{L}_{t,\mathbf{v}}\,\rho = 0 \iff \mathrm{div}\,\mathbf{v} = 0\,.$$

3.3.1 Mass flow thru a control volume

Let $\mathbf{C} \subset \mathrm{TRA}_I(\boldsymbol{\varphi},\mathbb{B}) \subset \mathbb{S}$ be a control-volume travelling, in the trajectory tracked by a body $\mathbb{B}$ in a time interval I, according to a flow $\mathbf{Fl}^{\mathbf{u}}_{\tau,t} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$, with velocity field $\mathbf{u}_t \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$. The time rate of change of the mass included in the travelling control-volume is provided by the transport formula

$$\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{u}}_{\tau,t}(\mathbf{C})} \rho_\tau\,\boldsymbol{\mu} = \int_{\mathbf{C}} \partial_{\tau=t}\,\rho_\tau\,\boldsymbol{\mu} + \int_{\mathbf{C}} \mathcal{L}_{\mathbf{u}}\,(\rho_t\,\boldsymbol{\mu})\,.$$

By the principle of conservation of mass, in the motion of a body with velocity $\mathbf{v} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{TS})$ we have that

$$\mathcal{L}_{t,\mathbf{v}}\,(\rho_t\boldsymbol{\mu}) = \partial_{\tau=t}\,\rho_\tau\,\boldsymbol{\mu} + \mathcal{L}_{\mathbf{v}}\,(\rho_t\,\boldsymbol{\mu}) = 0\,.$$

Hence

$$\begin{aligned}
\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{u}}_{\tau,t}(\mathbf{C})} \mathbf{m}_\tau &= \int_{\mathbf{C}} \partial_{\tau=t}\,\rho_\tau\,\boldsymbol{\mu} + \int_{\mathbf{C}} \mathcal{L}_{\mathbf{u}}\,(\rho_t\,\boldsymbol{\mu}) \\
&= -\int_{\mathbf{C}} \mathcal{L}_{\mathbf{v}}\,(\rho_t\,\boldsymbol{\mu}) + \int_{\mathbf{C}} \mathcal{L}_{\mathbf{u}}\,(\rho_t\,\boldsymbol{\mu}) \\
&= \int_{\mathbf{C}} \mathcal{L}_{\mathbf{u}-\mathbf{v}}\,(\rho_t\,\boldsymbol{\mu}) = \int_{\mathbf{C}} \mathcal{L}_{\rho_t(\mathbf{u}-\mathbf{v})}\,\boldsymbol{\mu} = \int_{\mathbf{C}} \mathrm{div}\,(\rho_t\,(\mathbf{u}-\mathbf{v}))\,\boldsymbol{\mu} \\
&= -\oint_{\partial\mathbf{C}} \mathbf{g}(\rho_t\,(\mathbf{v}-\mathbf{u}),\mathbf{n})\,(\boldsymbol{\mu}\mathbf{n}) = -\oint_{\partial\mathbf{C}} \mathbf{m}_t \cdot (\mathbf{v}-\mathbf{u})\,.
\end{aligned}$$

Since $\mathbf{n}$ is the outward normal to the boundary $\partial\mathbf{C}$ of the control-volume and $\mathbf{v}-\mathbf{u}$ is the relative velocity of the motion of the body with respect to the travelling control-volume, we may state that:

- The time rate of change of the mass included in a control-volume, travelling in the trajectory of a body, is equal to the inflow of mass-density thru the surface bounding the control-volume.

This alternative form of the principle of conservation of mass has the typical aspect of a balance law.

3.4 Euler equations of dynamics

According to JEAN D'ALEMBERT's point of view, the equations of dynamics, for a continuous body in motion in the euclidean space, are recovered from the cardinal equations of statics by adding, to the applied forces, the inertial term due to the field of momentum rate that the body undergoes in its motion with respect to an inertial reference system.

Figure 3.8: sir Isaac Newton (1643 - 1727)

The acceleration of a material particle is the time derivative of its speed. If the ambient space is a manifold $\mathbb{S}$ with an affine connection ∇ and the spatial velocity field of a particle along the trajectory is given, the acceleration is evaluated by taking the material time-derivative, according to EULER's formula:

$$\mathbf{a}_t = \nabla_{t,\mathbf{v}_t}\mathbf{v}_t := \partial_{\tau=t}\,\mathbf{v}_\tau + \nabla_{\mathbf{v}_t}\mathbf{v}_t\,.$$

In the usual euclidean setting, the connection is the one induced by the parallel transport by translation.

Following D'ALEMBERT's idea, the original statement of NEWTON's law of particle dynamics may be rewritten in variational terms as

$$\int_{\boldsymbol{\varphi}_t(\mathbb{B})} \mathbf{g}(\mathbf{a}_t, \delta\mathbf{v}_t)\,\mathbf{m}_t = \langle\, \mathbf{f}_t, \delta\mathbf{v}_t \,\rangle\,,$$

for any rigid virtual velocity field $\delta \mathbf{v}_t \in \mathrm{C}^0(\boldsymbol{\Omega}_t \,; \mathbb{T}_{\boldsymbol{\Omega}_t}\mathbb{S})$ where $\boldsymbol{\Omega}_t = \boldsymbol{\varphi}_t(\mathbb{B})$. Note that the symbol δ has no meaning by itself, it is the composed symbol $\delta\mathbf{v}$ that denotes a virtual velocity field.

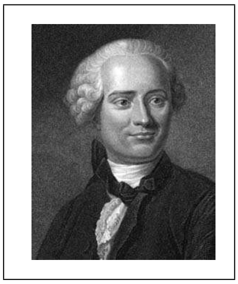

Figure 3.9: Jean Le Rond d'Alembert (1717 - 1783)

A more general way to state the law of dynamics for a continuous body with a variable mass was envisaged by LEONHARD EULER. Hereafter we state a variational formulation of EULER's law in the general setting of a riemannian ambient manifold $\{\mathbb{S}\,,\mathbf{g}\}$ with $\mathbf{g} \in \mathrm{C}^1(\mathbb{S}^2\,;\Re)$ the metric tensor field.

If not otherwise specified, the connection will be assumed to be the LEVI-CIVITA connection induced by the metric field.

Let us denote by $\boldsymbol{\varphi}_{\tau,t} := \boldsymbol{\varphi}_\tau \circ \boldsymbol{\varphi}_t^{-1} \in \mathrm{C}^1(\mathrm{TRA}_I(\boldsymbol{\varphi},\mathbb{B})\,;\mathrm{TRA}_I(\boldsymbol{\varphi},\mathbb{B}))$ the flow along the trajectory, and by $\boldsymbol{\varphi}_{\tau,t}\Uparrow$ the parallel transport by translation along the trajectory, from the placement $\boldsymbol{\Omega}_t = \boldsymbol{\varphi}_t(\mathbb{B})$ to the placement $\boldsymbol{\Omega}_\tau = \boldsymbol{\varphi}_\tau(\mathbb{B})$.

In variational form, EULER's law of motion may be stated as

$$\partial_{\tau=t} \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{v}_\tau, \boldsymbol{\varphi}_{\tau,t}\Uparrow \delta\mathbf{v}_t)\,\mathbf{m}_\tau = \langle\, \mathbf{f}_t, \delta\mathbf{v}_t \,\rangle\,,$$

for any rigid velocity field $\delta\mathbf{v}_t \in \mathrm{C}^0(\boldsymbol{\Omega}_t\,;\mathbb{T}_{\boldsymbol{\Omega}_t}\mathbb{S})$. Being $\mathbf{m}_t = \rho_t\,\boldsymbol{\mu}$, EULER's law may be restated as

$$\partial_{\tau=t} \int_{\boldsymbol{\Omega}_\tau} \mathbf{g}(\rho_\tau\mathbf{v}_\tau, \boldsymbol{\varphi}_{\tau,t}\Uparrow \delta\mathbf{v}_t)\,\boldsymbol{\mu} = \langle\, \mathbf{f}_t, \delta\mathbf{v}_t \,\rangle\,,$$

where the vector field $\rho_t\mathbf{v}_t \in \mathrm{C}^0(\boldsymbol{\Omega}_t\,;\mathbb{T}_{\boldsymbol{\Omega}_t}\mathbb{S})$ is the *kinetic momentum* per unit volume at time $t \in I$.

The statement of EULER's law requires to extend, the virtual velocity field $\delta\mathbf{v}_t \in \mathrm{C}^0(\boldsymbol{\Omega}_t\,;\mathbb{T}_{\boldsymbol{\Omega}_t}\mathbb{S})$ at the placement $\boldsymbol{\Omega}_t$, to a virtual velocity field $\delta\mathbf{v}_{\boldsymbol{\varphi}} \in$

$C^1(\mathrm{TRA}_I(\boldsymbol{\varphi}, \mathbb{B}); \mathbb{TS})$ defined along the trajectory, according to the translation rule:

$$\delta\mathbf{v}_{\boldsymbol{\varphi}}(\boldsymbol{\varphi}_{\tau,t}(\mathbf{x})) := \boldsymbol{\varphi}_{\tau,t}\Uparrow \delta\mathbf{v}_t(\mathbf{x}) , \quad \forall\, \mathbf{x} \in \boldsymbol{\varphi}_t(\mathbb{B}) ,$$

so that $\delta\mathbf{v}_{\boldsymbol{\varphi}}(\mathbf{x}) = \delta\mathbf{v}_t(\mathbf{x})$ and

$$\nabla_{\mathbf{v}_t}\delta\mathbf{v}_{\boldsymbol{\varphi}} = \partial_{\tau=t}\ \boldsymbol{\varphi}_{\tau,t}\Downarrow \boldsymbol{\varphi}_{\tau,t}\Uparrow \delta\mathbf{v}_t = \partial_{\tau=t}\ \delta\mathbf{v}_t = 0 .$$

EULER's and D'ALEMBERT's laws of dynamics are equivalent if conservation of mass holds, as stated by the next proposition.

Theorem 3.4.1 *In a riemannian configuration manifold* $\{\mathbb{S}, \mathbf{g}\}$ *endowed with a metric connection,* EULER*'s law of dynamics, is equivalent to* D'ALEMBERT*'s law of dynamics by conservation of mass:* $\mathcal{L}_{t,\mathbf{v}_t}\, \mathbf{m} = 0$ *in* $\boldsymbol{\varphi}_t(\mathbb{B})$ *.*

Proof. Let us recall that, for a scalar field $f \in C^1(\boldsymbol{\varphi}_t(\mathbb{B}); \Re)$ convective and the material time derivatives coincide, that is: $\mathcal{L}_{t,\mathbf{v}_t} f = \nabla_{t,\mathbf{v}_t} f$. Moreover we have that

$$\begin{aligned}
&\nabla_{t,\mathbf{v}_t}\mathbf{g} = \nabla_{\mathbf{v}_t}\mathbf{g} = 0 , && \text{metric connection,}\\
&\nabla_{t,\mathbf{v}_t}\delta\mathbf{v}_{\boldsymbol{\varphi}} = \nabla_{\mathbf{v}_t}\delta\mathbf{v}_{\boldsymbol{\varphi}} = 0 , && \text{parallel transport,}\\
&\mathbf{a}_t := \nabla_{t,\mathbf{v}_t}\mathbf{v}_t , && \text{material time derivative,}
\end{aligned}$$

and hence

$$\begin{aligned}
\mathcal{L}_{t,\mathbf{v}_t}\, \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_{\boldsymbol{\varphi}}) &= \nabla_{t,\mathbf{v}_t}\mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_{\boldsymbol{\varphi}})\\
&= (\nabla_{\mathbf{v}_t}\mathbf{g})(\mathbf{v}_t, \delta\mathbf{v}_{\boldsymbol{\varphi}}) + \mathbf{g}(\nabla_{t,\mathbf{v}_t}\mathbf{v}_t, \delta\mathbf{v}) + \mathbf{g}(\mathbf{v}_t, \nabla_{t,\mathbf{v}_t}\delta\mathbf{v}_{\boldsymbol{\varphi}})\\
&= \mathbf{g}(\mathbf{a}_t, \delta\mathbf{v}) .
\end{aligned}$$

By the transport theorem we get:

$$\begin{aligned}
\partial_{\tau=t} \int_{\boldsymbol{\Omega}_\tau} \mathbf{g}(\mathbf{v}_\tau, \boldsymbol{\varphi}_{\tau,t}\Uparrow \delta\mathbf{v})\, \mathbf{m}_\tau &= \int_{\boldsymbol{\Omega}_t} \mathcal{L}_{t,\mathbf{v}_t}\, (\mathbf{g}(\mathbf{v}_t, \boldsymbol{\varphi}_{\tau,t}\Uparrow \delta\mathbf{v})\, \mathbf{m})\\
&= \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_{\boldsymbol{\varphi}})\, (\mathcal{L}_{t,\mathbf{v}_t}\, \mathbf{m}) + (\mathcal{L}_{t,\mathbf{v}_t}\, \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_{\boldsymbol{\varphi}}))\, \mathbf{m}_t\\
&= \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{a}_t, \delta\mathbf{v})\, \mathbf{m}_t ,
\end{aligned}$$

and the result follows. ■

3.4.1 Gauss principle for constrained motions

Let us consider a riemannian ambient manifold $\{\mathbb{S}\,,\mathbf{g}\}$ with a metric tensor field $\mathbf{g} \in \mathrm{C}^1(\mathbb{T}\mathbb{S}^2\,;\Re)$ and the motion of a constrained continuous dynamical systems. This means that admissible velocity fields at any placement $\boldsymbol{\Omega}_t$ are bound to take values into subspaces $L_t(\mathbf{x})$ of the tangent spaces $\mathbb{T}_{\mathbf{x}}\mathbb{S}$ that is $\delta\mathbf{v}_t(\mathbf{x}) \in L_t(\mathbf{x})$ for $\mathbf{x} \in \boldsymbol{\Omega}_t$. The constraint $\mathcal{A}$ is the linear space of velocity fields on $\boldsymbol{\Omega}_t$ fulfilling this constraint. Then, denoting by $\mathcal{A}_0$ a less restrictive constraint space, i.e. such that $L_t(\mathbf{x}) \subset L_0(\mathbf{x})$ for $\mathbf{x} \in \boldsymbol{\Omega}_t$, we consider the mean square inner product

$$\textsc{Gauss}(\mathbf{u},\mathbf{v}) := \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{u},\mathbf{v})\,\mathbf{m}_t\,,$$

and the induced squared norm

$$\textsc{Gauss}(\mathbf{u}) := \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{u},\mathbf{u})\,\mathbf{m}_t\,,$$

Then we have the following property.

Theorem 3.4.2 (Gauss principle) *The acceleration field of a dynamical system subject to a kinematical constraint minimizes the mean square deviation from the acceleration field corresponding to a less stringent constraint.*

Proof. By imposing D'ALEMBERT's law of dynamics for the two constraints:

$$\int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{a}_{0t},\delta\mathbf{v}_t)\,\mathbf{m}_t = \langle\,\mathbf{f}_t,\delta\mathbf{v}_t\,\rangle\,, \quad \forall\,\delta\mathbf{v}_t \in \mathcal{A}_0\,, \quad \mathbf{a}_{0t} \in \mathcal{A}_0\,,$$

and

$$\int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{a}_t,\delta\mathbf{v}_t)\,\mathbf{m}_t = \langle\,\mathbf{f}_t,\delta\mathbf{v}_t\,\rangle\,, \quad \forall\,\delta\mathbf{v}_t \in \mathcal{A}\,, \quad \mathbf{a}_t \in \mathcal{A}\,,$$

assuming that $\mathcal{A} \subset \mathcal{A}_0$ and subtracting we get

$$\int_{\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{a}_t - \mathbf{a}_{0t},\delta\mathbf{v}_t)\,\mathbf{m}_t = 0\,, \quad \forall\,\delta\mathbf{v}_t \in \mathcal{A}\,.$$

Since $\mathbf{a}_t \in \mathcal{A}$, this is equivalent to $d\textsc{Gauss}(\mathbf{a}_t - \mathbf{a}_{0t}) \cdot \delta\mathbf{v}_t = 0\,,\ \ \forall\,\delta\mathbf{v}_t \in \mathcal{A}$, where d is the directional derivative, and the result follows by the strict convexity of the map $\textsc{Gauss}(\mathbf{a}_t - \mathbf{a}_{0t})$ with $\mathbf{a}_{0t}$ fixed. ■

Theorem 3.4.3 (Gibbs-Appell equations) GAUSS *principle is equivalent to the* GIBBS-APPELL *equations of dynamics:*

$$d\tfrac{1}{2}\text{GAUSS}(\mathbf{a}_t) \cdot \delta\mathbf{v}_t = \langle \mathbf{f}_t, \delta\mathbf{v}_t \rangle\,, \quad \forall\, \delta\mathbf{v}_t \in \mathcal{A}\,.$$

Proof. Being $\text{GAUSS}(\mathbf{a}_t - \mathbf{a}_{0t}) = \text{GAUSS}(\mathbf{a}_t) + \text{GAUSS}(\mathbf{a}_{0t}) - 2\,\text{GAUSS}(\mathbf{a}_t, \mathbf{a}_{0t})$, taking the derivative along any $\delta\mathbf{v}_t \in \mathcal{A}$ and recalling the D'ALEMBERT's law of dynamics the result follows. ■

3.4.2 Dynamics of a travelling control volume

The case of a variable mass can be conveniently dealt with by writing the equations of dynamics in terms of a control volume travelling along the trajectory. To this end, it is expedient to preliminarily remark that, being $\mathbf{m}_t = \rho_t \boldsymbol{\mu}$, we have:

$$\begin{aligned}
\partial_{\tau=t} \int_{\boldsymbol{\Omega}_\tau} \mathbf{g}(\mathbf{v}_\tau, \delta\mathbf{v}_t)\, \mathbf{m}_\tau &= \int_{\boldsymbol{\Omega}_t} \mathcal{L}_{t,\mathbf{v}_t} \left(\mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t\right) \\
&= \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\partial_{\tau=t}\ \rho_\tau \mathbf{v}_\tau, \delta\mathbf{v}_t)\, \boldsymbol{\mu} + \mathcal{L}_{\mathbf{v}_t}(\mathbf{g}(\rho_t \mathbf{v}_t, \delta\mathbf{v}_t)\, \boldsymbol{\mu}) \\
&= \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\partial_{\tau=t}\ \rho_\tau \mathbf{v}_\tau, \delta\mathbf{v}_t)\, \boldsymbol{\mu} + \mathcal{L}_{\mathbf{g}(\rho_t \mathbf{v}_t, \delta\mathbf{v}_t)\mathbf{v}_t}\, \boldsymbol{\mu} \\
&= \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\partial_{\tau=t}\ \rho_\tau \mathbf{v}_\tau, \delta\mathbf{v}_t)\, \boldsymbol{\mu} \\
&\quad + \oint_{\partial\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t \cdot \mathbf{v}_t\,.
\end{aligned}$$

The last equality follows from the divergence theorem:

$$\begin{aligned}
\int_{\boldsymbol{\Omega}_t} \mathcal{L}_{\mathbf{g}(\rho_t \mathbf{v}_t, \delta\mathbf{v}_t)\mathbf{v}_t}\, \boldsymbol{\mu} &= \int_{\boldsymbol{\Omega}_t} \operatorname{div}\left(\mathbf{g}(\rho_t \mathbf{v}_t, \delta\mathbf{v}_t)\mathbf{v}_t\right) \boldsymbol{\mu} \\
&= \oint_{\partial\boldsymbol{\Omega}_t} \mathbf{g}(\rho_t \mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{g}(\mathbf{v}_t, \mathbf{n})\, \partial\boldsymbol{\mu} \\
&= \oint_{\partial\boldsymbol{\Omega}_t} \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t \cdot \mathbf{v}_t\,,
\end{aligned}$$

since, for all $\mathbf{a}, \mathbf{b} \in \mathbb{T}_{\mathbf{x}}\boldsymbol{\Omega}_t$:

$$\mathbf{g}(\mathbf{v}_t, \mathbf{n})\, \partial\boldsymbol{\mu} \cdot \mathbf{a} \cdot \mathbf{b} = \mathbf{g}(\mathbf{v}_t, \mathbf{n})\, \boldsymbol{\mu} \cdot \mathbf{n} \cdot \mathbf{a} \cdot \mathbf{b} = \boldsymbol{\mu} \cdot \mathbf{v}_t \cdot \mathbf{a} \cdot \mathbf{b} \,.$$

Let $\mathbf{C} \in \mathbb{S}$ be a control-volume travelling, in the trajectory $\mathrm{TRA}_I(\boldsymbol{\varphi}) \subset \mathbb{S}$ tracked by a body in a time interval I, according to a flow $\mathbf{Fl}^{\mathbf{u}}_{\tau,t} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$, with time dependent velocity field $\mathbf{u}_t \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{TS})$.

We set $\mathbf{C} = \boldsymbol{\Omega}_t$. Then

$$\begin{aligned} \partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{u}}_{\tau,t}(\mathbf{C})} \mathbf{g}(\mathbf{v}_\tau, \delta\mathbf{v}_t)\, \mathbf{m}_\tau &= \int_{\mathbf{C}} \mathcal{L}_{t,\mathbf{u}_t} \left(\mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t\right) \\ &= \int_{\mathbf{C}} \mathbf{g}(\partial_{\tau=t}\ \rho_\tau \mathbf{v}_\tau, \delta\mathbf{v}_t)\, \boldsymbol{\mu} \\ &\quad + \oint_{\partial\mathbf{C}} \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t \cdot \mathbf{u}_t \,. \end{aligned}$$

By comparing the two previous expressions, we get the following relation

$$\begin{aligned} \partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{u}}_{\tau,t}(\mathbf{C})} \mathbf{g}(\mathbf{v}_\tau, \delta\mathbf{v}_t)\, \mathbf{m}_\tau - \partial_{\tau=t} \int_{\boldsymbol{\Omega}_\tau} \mathbf{g}(\mathbf{v}_\tau, \delta\mathbf{v}_t)\, \mathbf{m}_\tau &= \\ = - \oint_{\partial\mathbf{C}} \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t \cdot (\mathbf{v}_t - \mathbf{u}_t) & \\ = - \oint_{\partial\mathbf{C}} \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t \cdot \mathbf{w}_t \,. & \end{aligned}$$

EULER's law of motion for the travelling control-volume may then be stated as

$$\partial_{\tau=t} \int_{\mathbf{Fl}^{\mathbf{u}}_{\tau,t}(\mathbf{C})} \mathbf{g}(\mathbf{v}_\tau, \delta\mathbf{v}_t)\, \mathbf{m}_\tau = \langle \mathbf{f}_t, \delta\mathbf{v}_t \rangle - \oint_{\partial\mathbf{C}} \mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t \cdot \mathbf{w}_t \,,$$

for any rigid virtual velocity field $\delta\mathbf{v}_t \in \mathrm{C}(\boldsymbol{\Omega}_t\,;\mathbb{T}_{\boldsymbol{\Omega}_t}\mathbb{S})$ extended along the trajectory $\mathrm{TRA}_I(\boldsymbol{\varphi}, \mathbb{B})$ by pointwise translation.

The boundary integral provides the virtual work of the equivalent boundary force system (*thrust*) acting on a travelling control volume due to the momentum-loss per unit time. Here $\mathbf{v}_t$ is the absolute velocity of the material particles crossing the boundary of the control volume and $\mathbf{w}_t := \mathbf{v}_t - \mathbf{u}_t$ is the relative velocity of the material particles with respect to the crossed boundary points of the control volume.

We have that $\mathbf{m}_t \cdot \mathbf{w}_t = \mathbf{g}(\rho_t \mathbf{w}_t, \mathbf{n})\, \partial\boldsymbol{\mu}$ where the term $\mathbf{g}(\rho_t \mathbf{w}_t, \mathbf{n})$ is the mass leaving the control volume per unit time and unit surface area (*surfacial mass-loss rate*).

It is apparent that the thrust vanishes if either the absolute velocity of the particles at the control-volume boundary vanishes, i.e. $\mathbf{v}_t := \mathbf{w}_t + \mathbf{u}_t = 0$, or the surfacial mass-loss rate vanishes i.e. $\mathbf{g}(\rho_t \mathbf{w}_t, \mathbf{n}) = 0$.

3.5 The stress fields

In general, we need to know how the material body changes locally its shape under the action of a force system (by means of constitutive relations) and to impose that the local changes of shape be compatible with the kinematics of the body as a whole.

The problem so posed is a very hard one to be solved in general, also with computational approachs based on suitable discretizations of the continuous problem (i.e. finite element methods and similar ones).

Effective methods are now at hand for bodies whose geometry doesn't change significantly during the dynamical evolution. For such problems a linearized analysis may lead to satisfactory results with a comparatively low computational effort and provides an iterative tool in nonlinear solution algorithms. Basic to the theory, is however the mathematical representation, of a force system acting in dynamical equilibrium on the body, as a field of pointwise stresses in the body.

Figure 3.10: Jacob (Jacques) Bernoulli (1654 - 1705)

This representation was envisaged for non-viscous fluids by Jacob, Johann and Daniel Bernoulli and by Euler during the course of the XVIII century. The complete characterization for continuous bodies, including solids, is essentially due to Cauchy in 1827.

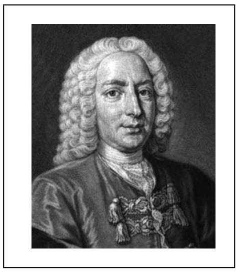

Figure 3.11: Daniel Bernoulli (1700 - 1782)

CAUCHY's geometrical approach was based on imposing the translational equilibrium to a *coin-shaped* and to a *tetrahedral domain* suitably contracting to a point. A similar, elegant approach more recently proposed by WALTER NOLL is based on the equilibrium of a *triangular prisma*, also suitably contracting to a point [163].

We shall not follow these approaches because, although fascinating for simplicity and skillfulness, they do not provide a satisfactory scenic view of the matter, since no explicit reference is made to basic duality arguments. Moreover, their application requires more regularity assumptions than needed.

The modern point of view, which has been first stressed in recent times by the author, is based on the application of a reasoning that we owe to LAGRANGE, a master of CAUCHY, and is well-known as the *method of Lagrangian multipliers*. According to TRUESDELL and TOUPIN [187], the idea of applying this method to the definition of the stress field in a body is due to GABRIO PIOLA as fas as 1833 [144]. Anyway, the LAGRANGE's multipliers method has gained its full soundness, about one century later (1934), by the tools of functional analysis that we owe mainly to the genius of STEFAN BANACH [12].

3.5.1 Lagrange multipliers

The method of LAGRANGE's multipliers was originarily envisioned to deal with the problem of finding the solution to an extremality problem of a functional under constraint's condition on its argument. We need a most general version of the method and a precise mathematical formulation is provided hereafter. We preliminarily recall that a BANACH space is a linear topological space endowed with a norm and complete in the induced norm-topology: each CAUCHY convergent sequence of elements converges towards an element of the space.

Figure 3.12: Joseph-Louis Lagrange (1736 - 1813)

In most applications BANACH spaces are indeed HILBERT spaces since the norm derives from an inner product, a positive definite symmetric bilinear form.

LAGRANGE's multipliers provide a tool to express the value of a continuous linear functional over a BANACH space of test fields when its value is known to vanish on a linear subspace of admissible test fields. which are in the kernel of a continuous linear operator providing an implicit representation of the linear constraints.

Theorem 3.5.1 (Lagrange's multipliers) *Let* $\boldsymbol{\alpha}^1 \in \mathbb{T}^*_{\mathbf{x}}\mathbb{C}$ *be a one-form at a point* $\mathbf{x} \in \mathbb{C}$ *of a manifold* $\mathbb{C}$ *modeled on a* BANACH *space, and let* $\boldsymbol{\pi}_{\mathcal{A}} : \mathcal{A} \mapsto \mathbb{C}$ *be a vector subbundle of the tangent bundle* $\mathbb{TC}$ *. Let us assume that the linear fiber* $\mathcal{A}_{\mathbf{x}} := \{\mathbf{v} \in \mathcal{A} \mid \boldsymbol{\pi}_{\mathcal{A}}(\mathbf{v}) = \mathbf{x}\}$ *at* $\mathbf{x} \in \mathbb{C}$ *is implicitly represented as* $Ker\, \mathbf{G}_{\mathbf{x}}$ *where* $\mathbf{G}_{\mathbf{x}} \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{C}\,; E)$ *is a bounded linear operator with closed range in a* BANACH *space* E *. Then, the orthogonality condition*

$$\langle \boldsymbol{\alpha}^1, \mathbf{v} \rangle = 0\,, \quad \forall\, \mathbf{v} \in \mathcal{A}_{\mathbf{x}} = Ker\, \mathbf{G}_{\mathbf{x}}\,,$$

is equivalent to require that there exists a $\boldsymbol{\lambda}_{\boldsymbol{\alpha}} \in E^*$ *such that*

$$\langle \boldsymbol{\alpha}^1, \mathbf{v} \rangle = \langle \boldsymbol{\lambda}_{\boldsymbol{\alpha}}, \mathbf{G}_{\mathbf{x}} \cdot \mathbf{v} \rangle\,, \quad \forall\, \mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}\,.$$

Proof. Let us denote by $\mathbf{G}^*_{\mathbf{x}} \in BL(E^*\,; \mathbb{T}^*_{\mathbf{x}}\mathbb{C})$ the bounded linear map defined by the duality condition:

$$\langle \mathbf{G}^*_{\mathbf{x}} \cdot \boldsymbol{\lambda}, \mathbf{v} \rangle_{\mathbb{T}^*_{\mathbf{x}}\mathbb{C} \times \mathbb{T}_{\mathbf{x}}\mathbb{C}} = \langle \boldsymbol{\lambda}, \mathbf{G} \cdot \mathbf{v} \rangle_{E^* \times E}\,, \quad \forall\, \mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}, \quad \forall\, \boldsymbol{\lambda} \in E^*\,,$$

which implies that

$$Ker\, \mathbf{G}_{\mathbf{x}} = (\mathrm{Im}\, \mathbf{G}^*_{\mathbf{x}})^0 \subseteq \mathbb{T}_{\mathbf{x}}\mathbb{C}\,.$$

By assumption $\mathrm{Im}\,\mathbf{G}_{\mathbf{x}}$ is a closed linear subspace of E and then BANACH's closed range theorem ensures that $\mathrm{Im}\,\mathbf{G}_{\mathbf{x}}^{*} \subset \mathbb{T}_{\mathbf{x}}^{*}\mathbb{C}$ is a closed linear subspace and that:

$$\mathrm{Im}\,\mathbf{G}_{\mathbf{x}}^{*} = (Ker\,\mathbf{G}_{\mathbf{x}})^{0}\,,$$

where $(Ker\,\mathbf{G}_{\mathbf{x}})^{0} := \{\mathbf{v}^{*} \in \mathbb{T}_{\mathbf{x}}^{*}\mathbb{C} \mid \langle\,\mathbf{v}^{*}, \mathbf{v}\,\rangle_{\mathbb{T}_{\mathbf{x}}^{*}\mathbb{C}\times\mathbb{T}_{\mathbf{x}}\mathbb{C}} = 0\,, \quad \forall\,\mathbf{v} \in Ker\,\mathbf{G}_{\mathbf{x}}\}$. The condition $\boldsymbol{\alpha}^{1} \in (Ker\,\mathbf{G}_{\mathbf{x}})^{0}$ is thus equivalent to $\boldsymbol{\alpha}^{1} \in \mathrm{Im}\,\mathbf{G}_{\mathbf{x}}^{*}$ and this means that there exists a $\boldsymbol{\lambda}_{\boldsymbol{\alpha}} \in E^{*}$ such that $\boldsymbol{\alpha}^{1} = \mathbf{G}_{\mathbf{x}}^{*} \cdot \boldsymbol{\lambda}_{\boldsymbol{\alpha}}$, or, equivalently, such that

$$\langle\,\boldsymbol{\alpha}^{1}, \mathbf{v}\,\rangle = \langle\,\mathbf{G}_{\mathbf{x}}^{*} \cdot \boldsymbol{\lambda}_{\boldsymbol{\alpha}}, \mathbf{v}\,\rangle = \langle\,\boldsymbol{\lambda}_{\boldsymbol{\alpha}}, \mathbf{G}_{\mathbf{x}} \cdot \mathbf{v}\,\rangle\,, \quad \forall\,\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}\,,$$

and the statement is proved. ■

In applications to mechanics of continua, $\mathbb{C}$ is the configuration manifold and the vector subbundle $\mathbf{p}_{\mathbb{C}} \in \mathrm{C}^{1}(\mathcal{A}\,;\mathbb{C})$ of the tangent bundle $\mathbb{TC}$ is the disjoint union of the linear subspaces of infinitesimal isometries of the body, at each configuration.

The linear operator $\mathbf{G}_{\mathbf{x}} \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{C}\,;E)$ provides the tangent strain at the configuration $\mathbf{x} \in \mathbb{C}$, corresponding to a virtual velocity field $\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}$.

The one-form $\boldsymbol{\alpha}^{1} \in \mathbb{T}_{\mathbf{x}}^{*}\mathbb{C}$ is a force system acting on the body at the placement $\mathbf{x} \in \mathbb{C}$ corresponding to the actual configuration and the duality pairing $\langle\,\boldsymbol{\alpha}^{1}, \mathbf{v}\,\rangle$ provides the virtual work performed by the force system $\boldsymbol{\alpha}^{1}$ for the velocity field $\mathbf{v} \in \mathbb{T}_{\mathbf{x}}\mathbb{C}$.

The LAGRANGE's multipliers method may be given a naïve mechanical interpretation, based on the following idea: the condition that the virtual work $\langle\,\boldsymbol{\alpha}^{1}, \mathbf{v}\,\rangle$ vanishes for any infinitesimal isometry means that the virtual work depends in fact on the tangent strain associated with the velocity field. For non-isometric velocities the virtual work may thus be expressed by the duality pairing between the tangent strain field associated with the velocity field and a dual stress field.

This is exactly what has been proved in Theorem 3.5.1 whose mechanical version takes the name of *theorem of virtual work* (see section 3.5.3).

The requirement that the linear map $\mathbf{G}_{\mathbf{x}} \in BL\,(\mathbb{T}_{\mathbf{x}}\mathbb{C}\,;E)$ has a closed range, is a technical one which is required since the space of kinematic fields is not finite dimensional. The issue will be discussed in the next section by endowing the linear kinematical space of virtual velocities with a suitable HILBERT's topology.

It is noteworthy that most linear differential operators governing classical problems of Mathematical Physics fulfill suitable closed range requirements of their restrictions to any closed subspace of the BANACH space of definition. This is at the basis of most existence results.

Although the title of the next section 3.5.2 could induce to think that only mathematical minded people should feel themselves interested in reading it, the ideas there proposed are of a genuine operative nature and correspond to what structural engineers put into the computational machinery to solve structural problems formulated in terms of a continuous model.

In this respect it is intriguing to highlight the natural resemblance between our general idea of a finite patchwork of regularity and computational methods of the finite element type.

3.5.2 Mathematical subtleties

Let us consider a virtual spatial flow $\mathbf{Fl}^{\mathbf{v}}_{\lambda} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$ dragging the body $\mathbb{B}$ in the space. The corresponding virtual velocity field $\mathbf{v} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{TS})$ of the body at the placement $\boldsymbol{\varphi}(\mathbb{B})$ is given by $\mathbf{v} = \partial_{\lambda=0}\, \mathbf{Fl}^{\mathbf{v}}_{\lambda}$.

The kinematical space is defined by requiring that the kinematic field $\mathbf{v}$ be $\mathbf{g}$-square integrable on the current placement $\boldsymbol{\varphi}(\mathbb{B})$ and that the tangent deformation $\mathrm{sym}\,\nabla\mathbf{v}$ be a piecewice $\mathbf{g}$-square integrable operator-valued distribution on a *patchwork* $\mathrm{P\scriptsize{AT}}_{\mathbf{v}}(\boldsymbol{\varphi}(\mathbb{B}))$ of nonoverlapping submanifolds whose union is a covering for $\boldsymbol{\varphi}(\mathbb{B})$. This means that any two submanifolds of the patchwork intersect only at boundary points and that their union contains the whole current placement $\boldsymbol{\varphi}(\mathbb{B})$.

The patchwork may vary from one kinematic field to another one and is named the *regularity patchwork* of the kinematic field. We will refer to these kinematic field as GREEN *regular kinematic fields*.

The pre-Hilbert kinematical space endowed with the topology induced by the mean square norm of the kinematic fields and of the regular part of the corresponding tangent deformation, is denoted by $\mathrm{K\scriptsize{IN}}(\boldsymbol{\varphi}(\mathbb{B}))$, or simply by $\mathrm{K\scriptsize{IN}}$.

The subspace $\mathrm{R\scriptsize{IG}} \subset \mathrm{K\scriptsize{IN}}$ of rigid kinematic fields is characterized by the property that $\mathrm{sym}\,\nabla\mathbf{v}$ vanishes on every element of the regularity patchwork $\mathrm{P\scriptsize{AT}}_{\mathbf{v}}(\boldsymbol{\varphi}(\mathbb{B}))$.

Force systems acting on the body at the placement $\boldsymbol{\varphi}(\mathbb{B})$ belong to the linear space $\mathrm{F\scriptsize{OR}} := \mathrm{K\scriptsize{IN}}^*$ which is the topological dual of $\mathrm{K\scriptsize{IN}}$.

The kinematic fields $\mathbf{v} \in \mathrm{K\scriptsize{IN}}$ which share a common regularity patchwork $\mathrm{P\scriptsize{AT}}(\boldsymbol{\varphi}(\mathbb{B}))$ form a linear closed subspace $\mathrm{K\scriptsize{IN}}(\mathrm{P\scriptsize{AT}}(\boldsymbol{\varphi}(\mathbb{B})))$, the PAT-regular kinematic space, which is a HILBERT space, i.e. a linear inner product space which is complete as a metric space, for the topology inherited by $\mathrm{K\scriptsize{IN}}$.

The boundary $\partial\mathrm{P\scriptsize{AT}}(\boldsymbol{\varphi}(\mathbb{B}))$ of the regularity patchwork is the collection of the boundaries of all the elements of the patchwork.

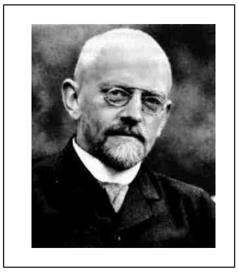

Figure 3.13: David Hilbert (1862 - 1943)

We say that *affine kinematical constraints* act on the body if boundary conditions are imposed on the fields of the PAT-regular kinematical space and define a closed flat manifold of *admissible kinematical fields* $\mathrm{ADM} \subset \mathrm{KIN}(\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B})))$. The closed linear space tangent to the manifold of admissible kinematical fields is denoted by $\mathrm{CONF} \subset \mathrm{KIN}(\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B})))$ and its elements are called *conforming kinematical fields*.

Due to linearity and closedness, the conformity space CONF is a HILBERT space for the topology inherited by KIN.

When the conformity space CONF is endowed with this hilbertian topology, it can be proven that the differential operator $\mathrm{sym}\,\nabla$ fulfills KORN's inequality:

$$\|\mathbf{v}\| + \|\mathrm{sym}\,\nabla\mathbf{v}\| \geq \alpha\,(\|\mathbf{v}\| + \|\nabla\mathbf{v}\|)\,, \quad \forall\, \mathbf{v} \in \mathrm{CONF}\,,$$

where $\|\|$ is the mean square norm on $\boldsymbol{\varphi}(\mathbb{B})$.

KORN's inequality states that the hilbertian topology of any conformity kinematic space is equivalent to the inner product topology of the SOBOLEV space $\mathcal{H}^1(\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B})))$, induced by the norm on the r.h.s. of the inequality.

Indeed symmetric and skew-symmetric components split the space of bounded linear operators into the sum of closed linear subspaces which are orthogonal supplements according to the usual inner product.

The inequality

$$\|\nabla\mathbf{v}\| \geq \|\mathrm{sym}\,\nabla\mathbf{v}\|\,, \quad \forall\, \mathbf{v} \in \mathrm{KIN}\,,$$

is a simple consequence of the pointwise inequality

$$\|\mathbf{L}(\mathbf{x})\|_{\mathbf{g}} \geq \|\mathrm{sym}\,\mathbf{L}(\mathbf{x})\|_{\mathbf{g}}\,,$$

which follows from PYTHAGORAS theorem.

The validity of KORN's inequality implies that the image thru $\operatorname{sym}\nabla$ of any closed linear subspace $\text{CONF} \subset \text{KIN}$ is closed and that the null-space of $\operatorname{sym}\nabla$ is finite dimensional. These two properties imply in turn the validity of KORN's inequality, [161], [165]. The mathematical construction described above opens the way to rely upon the LAGRANGE's multiplier method, introduced in theorem 3.5.1, to get the proof of the existence of a square integrable stress field equivalent to the force system acting in dynamical equilibrium on a body under arbitrary linear constraints defining a closed linear subspace of conforming kinematical fields. This representation result is discussed in detail in the next section.

3.5.3 Virtual work theorem

A load system $\ell \in \text{LOAD}$ acting on the body placed at $\boldsymbol{\varphi}(\mathbb{B})$, is an element of the HILBERT space $\text{LOAD} := \text{CONF}^*$ topological dual of CONF. Let a load $\ell \in \text{LOAD}$ meet the variational equilibrium condition:

$$\langle \ell, \mathbf{v} \rangle = 0\,, \quad \forall\, \mathbf{v} \in \text{CONF} \cap \text{RIG}\,.$$

Then theorem 3.5.1 ensures that there exists a (not necessarily unique) $\mathbf{g}$-square integrable field $\mathbf{T} : \boldsymbol{\varphi}(\mathbb{B}) \mapsto BL\,(\mathbb{TS}\,; \mathbb{TS})$ on the placement $\boldsymbol{\varphi}(\mathbb{B})$, whose point-values are $\mathbf{g}$-symmetric operators, fulfilling the following *virtual work identity*:

$$\langle \ell, \mathbf{v} \rangle = \int_{\text{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \operatorname{sym} \nabla \mathbf{v} \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,, \quad \forall\, \mathbf{v} \in \text{CONF}\,.$$

Here $\langle \cdot, \cdot \rangle_{\mathbf{g}}$ is the inner product between linear operators induced by the metric $\mathbf{g}$ (see section 1.1.4).

A LAGRANGE multiplier $\mathbf{T}(\boldsymbol{\varphi}(\mathbf{x})) \in BL\,(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,; \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ for the rigidity constraint is called a CAUCHY stress state.

The Cauchy stress tensor

$$\boldsymbol{\sigma}^*(\boldsymbol{\varphi}(\mathbf{x})) \in BL\,(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}^2\,; \Re) = BL\,(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,; \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})\,,$$

is the twice contravariant tensor defined by the relation

$$\mathbf{T} = \boldsymbol{\sigma}^* \circ \mathbf{g}\,, \qquad \boldsymbol{\sigma}^* = \mathbf{T} \circ \mathbf{g}^{-1}\,,$$

where the linear operator $\mathbf{g} = \mathbf{g}^\flat \in BL\,(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,; \mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ is the metric tensor.

Then, being $\frac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g} = \mathbf{g}\circ\operatorname{sym}\nabla\mathbf{v}$, according to the duality between the space $BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}^2;\Re)$ of twice covariant tensors and the space $BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}^2;\Re)$ of twice contravariant tensors, we have that

$$\langle \mathbf{T}, \operatorname{sym}\nabla\mathbf{v}\rangle_{\mathbf{g}} = \langle \boldsymbol{\sigma}^* \circ \mathbf{g}, \mathbf{g}^{-1} \circ \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}\rangle_{\mathbf{g}} = \langle \boldsymbol{\sigma}^*, \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}\rangle \,.$$

Here the second term in the equality is independent of the choice of the metric tensor (see section 1.1.4 proposition **??** on page **??**). The virtual work identity may then be written, in terms of tensor fields in the current placement, as

$$\langle \ell, \mathbf{v}\rangle = \int_{\textsc{Pat}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \boldsymbol{\sigma}^*, \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}\rangle \; \boldsymbol{\mu}\,, \quad \forall\, \mathbf{v} \in \textsc{Conf}\,.$$

Remark 3.5.1 *The assumption of a symmetric* CAUCHY *stress tensor* $\boldsymbol{\sigma}^* \in BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}^2;\Re)$ *and hence of a* $\mathbf{g}$*-symmetric* CAUCHY *stress operator* $\mathbf{T} \in BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S};\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ *is a natural choice due to the* $\mathbf{g}$*-symmetry of the* EULER*'s operator* $\operatorname{sym}\nabla\mathbf{v} \in BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S};\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ *and not a provable theorem, in spite of the common claim in textbooks on continuum mechanics. The choice of a non* $\mathbf{g}$*-symmetric Cauchy stress field* $\mathbf{T} \in BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S};\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ *is a permissible one but not a convenient one. Indeed the ineffective skew-symmetric part of it would perform virtual work for the symmetric* EULER*'s operator and would thus lead, through integration by parts, to a representation of the force system which includes body couples per unit volume [163]. But still worse thing would come, since all the nice and useful properties of the spectrum of a symmetrizable operator would be lost. Thus,* CAUCHY*'s choice of a symmetric stress tensor provides the most convenient representation of a system of forces in equilibrium.*

3.5.4 Boundary value problems

Boundary value problems are characterized by the following property.

- The closed linear subspace $\textsc{Conf} \subset \textsc{Kin}$ of conforming kinematical fields includes the whole linear subspace of kinematical fields with vanishing boundary values.

The basic tool in boundary value problems governed by a linear differential operator $\textsc{Diff}$ of order n, is GREEN's formula of integration by parts:

$$\int_{\textsc{Pat}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \textsc{Diff}\,\mathbf{v}\rangle_{\mathbf{g}}\, \boldsymbol{\mu} = \int_{\textsc{Pat}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\textsc{AdjDiff}\,\mathbf{T}, \mathbf{v})\, \boldsymbol{\mu}$$

$$+ \int_{\partial\textsc{Pat}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\textsc{Flux}\,\mathbf{T}, \textsc{Val}\,\mathbf{v})\, \partial\boldsymbol{\mu}\,, \quad \begin{cases} \forall\, \mathbf{v} \in \textsc{Kin}\,, \\ \forall\, \mathbf{T} \in \textsc{Stress}\,. \end{cases}$$

where $\textsc{AdjDiff}$ is a differential operator of order n said to be the *formal adjoint* of $\textsc{Diff}$. The boundary integral is over the duality pairing between two fields of the type $\textsc{Flux}\,\mathbf{T}$ and $\textsc{Val}\,\mathbf{v}$ where the differential operators $\textsc{Flux}$ and $\textsc{Val}$ are n-tuples of normal derivatives of order from 0 to $n-1$ in inverse sequence, to that the duality pairing is the sum of n terms such that in the k-th term the normal derivatives of thee two fields appear respectively to the order k and $n-1-k$.

In boundary value problems of continuum mechanics, it is assumed that a loading $\ell_{\{\mathbf{b},\mathbf{t}\}} \in \textsc{Load}$ is associated with a patchwork $\textsc{Pat}_{\{\mathbf{b},\mathbf{t}\}}(\boldsymbol{\varphi}(\mathbb{B}))$ and is composed by a vector field $\mathbf{b} \in \mathcal{L}^2(\boldsymbol{\varphi}(\mathbb{B})\,;V)$ of body force densities, i.e. forces per unit volume, and a vector field $\mathbf{t} \in \mathcal{L}^2(\partial\textsc{Pat}(\boldsymbol{\varphi}(\mathbb{B}))\,;V)$ of boundary tractions, i.e. forces per unit area, according to the definition:

$$\langle \ell_{\{\mathbf{b},\mathbf{t}\}}, \mathbf{v} \rangle := \int_{\boldsymbol{\varphi}(\mathbb{B})} \mathbf{g}(\mathbf{b},\mathbf{v})\,\boldsymbol{\mu} + \int_{\partial\textsc{Pat}_{\{\mathbf{b},\mathbf{t}\}}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{t},\mathbf{v})\,\partial\boldsymbol{\mu}\,, \quad \forall\,\mathbf{v} \in \textsc{Kin}\,.$$

Then we have the following result.

Theorem 3.5.2 (Cauchy's differential law of equilibrium) *In a boundary value problem, a stress field* $\mathbf{T}$ *in equilibrium with a load* $\ell_{\{\mathbf{b},\mathbf{t}\}}$*, i.e. fulfilling the virtual work identity*

$$\langle \ell_{\{\mathbf{b},\mathbf{t}\}}, \mathbf{v} \rangle = \int_{\textsc{Pat}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \operatorname{sym}\nabla\mathbf{v} \rangle_{\mathbf{g}}\,\boldsymbol{\mu}\,, \quad \forall\,\mathbf{v} \in \textsc{Conf}\,,$$

has a distributional divergence $\textsc{Div}\,\mathbf{T}$ *whose restriction to each element* $\textsc{Elem} \in \textsc{Pat}(\boldsymbol{\varphi}(\mathbb{B}))$ *of the patchwork is* $\mathbf{g}$*-square integrable with* $-\textsc{Div}\,\mathbf{T} = \mathbf{b}$*.*

Proof. In boundary value problems the test fields in the principle of virtual work may be taken to be kinematical fields with vanishing boundary values in each element $\textsc{Elem} \in \textsc{Pat}_{\{\mathbf{b},\mathbf{t}\}}(\boldsymbol{\varphi}(\mathbb{B}))$, so that

$$\langle \ell, \mathbf{v} \rangle = \int_{\textsc{Elem}} \mathbf{g}(\mathbf{b},\mathbf{v})\,\boldsymbol{\mu} = \int_{\textsc{Elem}} \langle \mathbf{T}, \operatorname{sym}\nabla\mathbf{v} \rangle_{\mathbf{g}}\,\boldsymbol{\mu}\,, \quad \forall\,\mathbf{v} \in \mathit{Ker}\,\textsc{Val}\,(\textsc{Elem})\,.$$

Hence, by the definition of distributional divergence $\textsc{Div}\,\mathbf{T}$:

$$\int_{\partial\textsc{Elem}} \langle \textsc{Div}\,\mathbf{T}, \mathbf{v} \rangle_{\mathbf{g}}\,\boldsymbol{\mu} := -\int_{\partial\textsc{Elem}} \langle \mathbf{T}, \operatorname{sym}\nabla\mathbf{v} \rangle_{\mathbf{g}}\,\boldsymbol{\mu}\,, \quad \forall\,\mathbf{v} \in \mathit{Ker}\,\textsc{Val}\,(\textsc{Elem})\,,$$

we infer that

$$\int_{\textsc{Elem}} \langle \textsc{Div}\,\mathbf{T}, \mathbf{v} \rangle_{\mathbf{g}}\,\boldsymbol{\mu} := \int_{\textsc{Elem}} \mathbf{g}(\mathbf{b},\mathbf{v})\,\boldsymbol{\mu}\,, \quad \forall\,\mathbf{v} \in \mathit{Ker}\,\textsc{Val}\,(\textsc{Elem})\,,$$

that is the meaning of the piecewise equality $-\textsc{Div}\,\mathbf{T} = \mathbf{b}$. ■

Stress fields whose distributional divergence $\mathrm{DIV}\,\mathbf{T}$ is piecewise representable by a square integrable field are said to be GREEN *regular stress fields* and we will write $\mathbf{T} \in \mathrm{STRESS}$ denoting by $\mathrm{PAT}_{\mathbf{T}}$ the regularity patchwork.

The GREEN regularity of stress and kinematic fields ensures that all the terms in the relevant GREEN's formula are well defined, so that:

$$\int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \mathrm{sym}\,\nabla\mathbf{v} \rangle_{\mathbf{g}}\, \boldsymbol{\mu} = \int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(-\mathrm{DIV}\,\mathbf{T}, \mathbf{v})\, \boldsymbol{\mu}$$

$$+ \int_{\partial\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathrm{FLUX}\,\mathbf{T}, \mathrm{VAL}\,\mathbf{v})\, \partial\boldsymbol{\mu}\,, \quad \begin{cases} \forall\, \mathbf{v} \in \mathrm{KIN}\,, \\ \forall\, \mathbf{T} \in \mathrm{STRESS}\,. \end{cases}$$

where $\mathrm{FLUX}\,\mathbf{T} = \mathbf{Tn}$ with $\mathbf{n}$ outward unit normal to the boundary $\partial\mathrm{ELEM} \in \partial\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))$, of a patchwork PAT finer than $\mathrm{PAT}_{\mathbf{v}}$ and $\mathrm{PAT}_{\mathbf{T}}$, and $\mathrm{VAL}\,\mathbf{v} = \mathbf{v}|_{\partial\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))}$ is the boundary value of the field $\mathbf{v} \in \mathrm{KIN}$, i.e. its restriction to $\partial\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))$.

Remark 3.5.2 *The partial order relation* $\mathrm{PAT}_1(\boldsymbol{\varphi}(\mathbb{B})) \prec \mathrm{PAT}_2(\boldsymbol{\varphi}(\mathbb{B}))$*, to be read:* $\mathrm{PAT}_2(\boldsymbol{\varphi}(\mathbb{B}))$ ***finer than*** $\mathrm{PAT}_1(\boldsymbol{\varphi}(\mathbb{B}))$*, means that every element of the patchwork* $\mathrm{PAT}_2(\boldsymbol{\varphi}(\mathbb{B}))$ *is included in an element of the patchwork* $\mathrm{PAT}_1(\boldsymbol{\varphi}(\mathbb{B}))$*. A patchwork finer then a given pair of patchworks is provided by the* ***grid*** $\mathrm{PAT}_2(\boldsymbol{\varphi}(\mathbb{B})) \wedge \mathrm{PAT}_1(\boldsymbol{\varphi}(\mathbb{B}))$ *of the two patchwork defined as the one whose elements are intersections of two elements of the given pair of patchworks. The set of all patchworks is then a direct set under the order relation* ***finer than****.*

Theorem 3.5.3 (Cauchy's boundary law of equilibrium) *In a boundary value problem, let* $\mathbf{T}$ *be a stress field in equilibrium with a load* $\ell_{\{\mathbf{b},\mathbf{t}\}}$*, i.e. fulfilling the virtual work identity*

$$\langle \ell_{\{\mathbf{b},\mathbf{t}\}}, \mathbf{v} \rangle = \int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \mathrm{sym}\,\nabla\mathbf{v} \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,, \quad \forall\, \mathbf{v} \in \mathrm{CONF}\,,$$

and $\mathrm{PAT}^*(\boldsymbol{\varphi}(\mathbb{B}))$ *a patchwork finer than the grid* $\mathrm{PAT}_{\{\mathbf{b},\mathbf{t}\}}(\boldsymbol{\varphi}(\mathbb{B})) \wedge \mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))$*. Then the jump*

$$[[\mathbf{Tn}]] := \mathbf{T}^+\mathbf{n}^+ + \mathbf{T}^-\mathbf{n}^- = \mathbf{T}^+\mathbf{n}^+ - \mathbf{T}^-\mathbf{n}^+\,,$$

of the flux $\mathbf{Tn}$ *across the interfaces* $+$ *and* $-$ *between the element of the patchwork* $\mathrm{PAT}^*(\boldsymbol{\varphi}(\mathbb{B}))$ *is such that*

$$[[\mathbf{Tn}]] = \mathbf{t}^+ + \mathbf{t}^- + \mathrm{CONF}^{\perp}\,,$$

where the field $\mathbf{t}$ *is extended to zero outside its domain of definition.*

Proof. From the virtual work identity and GREEN's formula we get

$$\begin{aligned}\langle \ell_{\{\mathbf{b},\mathbf{t}\}}, \mathbf{v} \rangle &:= \int_{\boldsymbol{\varphi}(\mathbb{B})} \mathbf{g}(\mathbf{b}, \mathbf{v})\, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}^*(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{t}, \mathbf{v})\, \partial\boldsymbol{\mu} \\ &= \int_{\mathrm{PAT}^*(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(-\mathrm{DIV}\, \mathbf{T}, \mathbf{v})\, \boldsymbol{\mu} \\ &+ \int_{\partial \mathrm{PAT}^*(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{T}\mathbf{n}, \mathbf{v})\, \partial\boldsymbol{\mu}\,, \quad \forall\, \mathbf{v} \in \mathrm{CONF}\,,\end{aligned}$$

and by Cauchy's differential law of equilibrium (Theorem 3.5.2) we infer that

$$\int_{\partial \mathrm{PAT}^*(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{t}, \mathbf{v})\, \partial\boldsymbol{\mu} = \int_{\partial \mathrm{PAT}^*(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{T}\mathbf{n}, \mathbf{v})\, \partial\boldsymbol{\mu}\,, \quad \forall\, \mathbf{v} \in \mathrm{CONF}\,,$$

and hence the result. ■

Let us now observe that the virtual work

$$\int_{\mathrm{PAT}_{\mathbf{v}}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \mathrm{sym}\, \nabla\mathbf{v} \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,, \quad \mathbf{v} \in \mathrm{KIN}\,,$$

is well-defined for any (even nonconforming) kinematic field $\mathbf{v} \in \mathrm{KIN}$.

Then, by making recourse to GREEN's formula, we may define the reactive force system $\mathbf{r}(\mathbf{t}, \mathbf{b}, \mathbf{T})$, associated with a body force field $\mathbf{b}$, a boundary traction field $\mathbf{t}$ and a stress field $\mathbf{T} \in \mathrm{STRESS}$, by the relation

$$\begin{aligned}\langle \mathbf{r}, \mathbf{v} \rangle &:= \int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \mathrm{sym}\, \nabla\mathbf{v} \rangle_{\mathbf{g}}\, \boldsymbol{\mu} - \int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{b}, \mathbf{v})\, \boldsymbol{\mu} - \int_{\partial \mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{t}, \mathbf{v})\, \boldsymbol{\mu} \\ &= \int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(-\mathrm{div}\, \mathbf{T} - \mathbf{b}, \mathbf{v})\, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{T}\mathbf{n} - \mathbf{t}, \mathbf{v})\, \boldsymbol{\mu}\,, \quad \mathbf{v} \in \mathrm{KIN}\,,\end{aligned}$$

where PAT is a patchwork finer than $\mathrm{PAT}_{\mathbf{v}}$, $\mathrm{PAT}_{\mathbf{T}}$ and $\mathrm{PAT}_{\{\mathbf{b},\mathbf{t}\}}(\boldsymbol{\varphi}(\mathbb{B}))$.

Due to the density of the linear space $\mathrm{C}_0^\infty(\mathrm{ELEM})$ of infinitely differentiable field with compact support in the space $\mathcal{L}^2(\mathrm{ELEM})$ of square integrable vector fields on each element ELEM, and being $\mathrm{C}_0^\infty(\mathrm{ELEM}) \subset \mathit{Ker}\, \mathrm{VAL}\,(\mathrm{ELEM})$, choosing $\mathbf{v} \in \mathit{Ker}\, \mathrm{VAL}\,(\mathrm{ELEM})$, we infer that $\mathrm{div}\, \mathbf{T} = -\mathbf{b}$ and hence that

$$\langle \mathbf{r}, \mathbf{v} \rangle = \int_{\partial \mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{g}(\mathbf{T}\mathbf{n} - \mathbf{t}, \mathbf{v})\, \boldsymbol{\mu} = 0\,, \quad \forall\, \mathbf{v} \in \mathrm{KIN}\,.$$

We underline the well-known characteristic property of linear constraints:

$$\langle \mathbf{r}, \mathbf{v} \rangle = 0\,, \quad \forall\, \mathbf{v} \in \text{Conf} \iff \mathbf{r} \in \text{Conf}^{\perp}\,,$$

stating that reactive force systems perform no virtual work for conforming virtual displacements.

3.5.5 Referential dynamical equilibrium

In finite deformation analyses of a dynamical equilibrium problem of a continuous body, the current placement of the body is an unknown of the problem. It is then be convenient to refer the state variables to a reference placement $\mathbb{B}$.

Moreover the concept of elastic behavior requires to assume an elastic potential which is a function of the configuration change from a reference natural placement of the body. These motivations require to express the equilibrium condition in terms of fields defined in a reference placement $\mathbb{B}$.

To this end, from section 1.2.6 on page 36, we recall that

- the pull-back of a twice covariant tensor $\beta \in BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,; \mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ from $\boldsymbol{\varphi}(\mathbf{x}) \in \boldsymbol{\varphi}(\mathbb{B})$ to $\mathbf{x} \in \mathbb{B}$ is given by:

$$\mathbf{g}_{\mathbf{x}}^{-1}(\boldsymbol{\varphi}{\downarrow}\boldsymbol{\beta}) = d\boldsymbol{\varphi}^T(\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{x})}^{-1}\boldsymbol{\beta})\, d\boldsymbol{\varphi}\,,$$

- the pull-back of a twice contravariant tensor $\alpha^* \in BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,; \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ from $\boldsymbol{\varphi}(\mathbf{x}) \in \boldsymbol{\varphi}(\mathbb{B})$ to $\mathbf{x} \in \mathbb{B}$ is given by:

$$(\boldsymbol{\varphi}{\downarrow}\boldsymbol{\alpha}^*)\,\mathbf{g}_{\mathbf{x}} = d\boldsymbol{\varphi}^{-1}(\boldsymbol{\alpha}^*\mathbf{g}_{\boldsymbol{\varphi}(\mathbf{x})})\, d\boldsymbol{\varphi}^{-T}\,.$$

Accordingly, the pull-back of the tangent deformation tensor from the actual placement $\boldsymbol{\varphi}(\mathbb{B})$ to the reference placement $\mathbb{B}$ is given by:

$$\boldsymbol{\varphi}{\downarrow}(\tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}) = \boldsymbol{\varphi}{\downarrow}(\mathbf{g}(\text{sym}\,\nabla\mathbf{v})) = \mathbf{g}(d\boldsymbol{\varphi}^T\text{sym}\,(\nabla\mathbf{v})d\boldsymbol{\varphi})\,.$$

The pull-back of the symmetric stress tensor $\boldsymbol{\sigma}^*(\boldsymbol{\varphi}(\mathbf{x})) \in BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,; \mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ from the actual placement $\boldsymbol{\varphi}(\mathbb{B})$ to the reference placement $\mathbb{B}$ is given by:

$$\boldsymbol{\varphi}{\downarrow}\boldsymbol{\sigma}^* = \boldsymbol{\varphi}{\downarrow}(\mathbf{T}\mathbf{g}^{-1}) = d\boldsymbol{\varphi}^{-1}\mathbf{T}d\boldsymbol{\varphi}^{-T}\mathbf{g}^{-1}\,.$$

To provide an expression of the virtual work in terms of operators defined in the reference placement, we introduce:

- The Kirchhoff stress tensor $\mathbf{k}^*(\boldsymbol{\varphi}(\mathbf{x})) \in BL(\mathbb{T}^*_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,;\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ which is the symmetric twice contravariant tensor defined by $\mathbf{k}^* := \mathbf{J}_{\boldsymbol{\varphi}}\,\mathbf{s}^*$. The mixed form $\mathbf{K}(\boldsymbol{\varphi}(\mathbf{x})) \in BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,;\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$ is then given by $\mathbf{K} := \mathbf{J}_{\boldsymbol{\varphi}}\,\mathbf{T}$.

- The Piola-Kirchhoff stress which is the symmetric twice contravariant tensor $\mathbf{s}^*(\mathbf{x}) \in BL(\mathbb{T}^*_{\mathbf{x}}\mathbb{S}\,;\mathbb{T}_{\mathbf{x}}\mathbb{S})$ related to the Kirchhoff stress by the pull-back correspondence:
$$\mathbf{s}^* := \boldsymbol{\varphi}{\downarrow}\mathbf{k}^*\,.$$

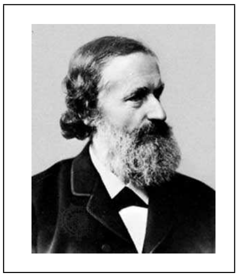

Figure 3.14: Gustav Robert Kirchhoff (1824 - 1887)

Their mixed forms $\mathbf{S}(\mathbf{x}) \in BL(\mathbb{T}_{\mathbf{x}}\mathbb{S}\,;\mathbb{T}_{\mathbf{x}}\mathbb{S})$ and $\mathbf{K}(\boldsymbol{\varphi}(\mathbf{x})) \in BL(\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S}\,;\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{x})}\mathbb{S})$, given by $\mathbf{S} = \mathbf{s}^*\mathbf{g}$ and $\mathbf{K} = \mathbf{k}^*\mathbf{g}$, are related by

$$\mathbf{S} = \boldsymbol{\varphi}{\downarrow}(\mathbf{K}\mathbf{g}^{-1})\mathbf{g} = d\boldsymbol{\varphi}^{-1}\mathbf{K}d\boldsymbol{\varphi}^{-T}\,,$$

and the following invariance property holds:

$$\begin{aligned}\langle \mathbf{K}, \operatorname{sym}\nabla\mathbf{v}\rangle_{\mathbf{g}} \circ \boldsymbol{\varphi} &= \langle \mathbf{S}, d\boldsymbol{\varphi}^T \operatorname{sym}(\nabla\mathbf{v})\, d\boldsymbol{\varphi}\rangle_{\mathbf{g}} \\ &= \langle \mathbf{S}, \operatorname{sym}(d\boldsymbol{\varphi}^T\nabla(\mathbf{v}\circ\boldsymbol{\varphi}))\rangle_{\mathbf{g}}\,.\end{aligned}$$

The virtual work identity may thus be written as

$$\begin{aligned}\langle \mathbf{f}, \mathbf{v} \rangle &= \int_{\mathrm{PAT}_{\mathbf{v}}(\boldsymbol{\varphi}(\mathbb{B}))} \langle \mathbf{T}, \operatorname{sym} \nabla \mathbf{v} \rangle_{\mathbf{g}}\, \boldsymbol{\mu} \\ &= \int_{\mathrm{PAT}_{\mathbf{v}}(\boldsymbol{\varphi}(\mathbb{B}))} \mathbf{J}_{\boldsymbol{\varphi}}^{-1} \langle \mathbf{K}, \operatorname{sym} \nabla \mathbf{v} \rangle_{\mathbf{g}}\, \boldsymbol{\mu} \\ &= \int_{\mathrm{PAT}_{(\mathbf{v}\circ\boldsymbol{\varphi})}(\mathbb{B})} (\langle \mathbf{K}, \operatorname{sym} \nabla \mathbf{v} \rangle_{\mathbf{g}} \circ \boldsymbol{\varphi})\, \boldsymbol{\mu} \\ &= \int_{\mathrm{PAT}_{(\mathbf{v}\circ\boldsymbol{\varphi})}(\mathbb{B})} \langle \mathbf{S}, d\boldsymbol{\varphi}^T \operatorname{sym} (\nabla \mathbf{v})\, d\boldsymbol{\varphi} \rangle\, \boldsymbol{\mu} \\ &= \int_{\mathrm{PAT}_{(\mathbf{v}\circ\boldsymbol{\varphi})}(\mathbb{B})} \langle d\boldsymbol{\varphi}\, \mathbf{S}, \nabla(\mathbf{v} \circ \boldsymbol{\varphi}) \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,,\end{aligned}$$

where $\mathbf{J}_{\boldsymbol{\varphi}} = \det(d\boldsymbol{\varphi})$ is the Jacobian deteminant of the configuration map.

Let us now assume that the divergence field $\operatorname{div}(d\boldsymbol{\varphi}\,\mathbf{S})$ is piecewise square integrable according to a regularity patchwork $\mathrm{PAT}_{\mathbf{S}}(\mathbb{B})$.

Then GREEN's formula yields

$$\begin{aligned}\langle \mathbf{f}, \mathbf{v} \rangle &= \int_{\mathrm{PAT}(\mathbb{B})} \langle d\boldsymbol{\varphi}\, \mathbf{S}, \nabla(\mathbf{v} \circ \boldsymbol{\varphi}) \rangle_{\mathbf{g}}\, \boldsymbol{\mu} \\ &= -\int_{\mathrm{PAT}(\mathbb{B})} \langle \operatorname{div}(d\boldsymbol{\varphi}\, \mathbf{S}), \mathbf{v} \circ \boldsymbol{\varphi} \rangle_{\mathbf{g}}\, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}(\mathbb{B})} \langle d\boldsymbol{\varphi}\, \mathbf{S}\, \mathbf{n}_{\mathbb{B}}, \mathbf{v} \circ \boldsymbol{\varphi} \rangle_{\mathbf{g}}\, (\partial \boldsymbol{\mu})\end{aligned}$$

where $\mathrm{PAT}(\mathbb{B})$ is a patchwork finer than $\mathrm{PAT}_{\mathbf{S}}(\mathbb{B}) \wedge \mathrm{PAT}_{(\mathbf{v}\circ\boldsymbol{\varphi})}(\mathbb{B})$ and $\mathbf{n}_{\mathbb{B}}$ is the outward unit normal to the elements of $\mathrm{PAT}(\mathbb{B})$.

GREEN's formula states that the system of referential forces may be represented by

- a field of body forces $-\operatorname{div}(d\boldsymbol{\varphi}\,\mathbf{S})$ and
- a field of surface tractions $(d\boldsymbol{\varphi}\,\mathbf{S})\,\mathbf{n}_{\mathbb{B}}$.

Introducing the PIOLA stress field $\mathbf{P} := d\boldsymbol{\varphi}\,\mathbf{S}$, we may state that the system of referential forces are composed by

- a field of body forces $-\operatorname{div}\mathbf{P}$ and
- a field of surface tractions $\mathbf{P}\,\mathbf{n}_{\mathbb{B}}$.

3.6 Continuum dynamics

The peculiar geometric feature of continuous dynamical systems is that two differentiable structures are playmates: the *ambient* finite dimensional riemannian manifold $(\mathcal{S}, \mathbf{g})$ (usually the flat euclidean 3D space) in which motions take place, and the *configuration* infinite dimensional manifold $\mathbb{C}$, describing the states of the system. The corresponding tangent bundles are denoted by $\boldsymbol{\tau}_{\mathcal{S}} \in \mathrm{C}^1(\mathbb{T}\mathcal{S} ; \mathcal{S})$ and $\boldsymbol{\tau}_{\mathbb{C}} \in \mathrm{C}^1(\mathbb{T}\mathbb{C} ; \mathbb{C})$. In discrete systems both manifolds are finite dimensional, sometimes taken to be coincident. In continuous systems, points of the *configuration* manifold are diffeomorphic maps with a fixed domain, a reference submanifold of the ambient manifold, and with codomains which are *placements*, submanifolds of the ambient manifold. To a vector tangent to the configuration manifold at a configuration, there corresponds a field of vectors tangent to the ambient manifold on the corresponding placement submanifold. The theory of continuous dynamical systems is then a *field* theory and it is essential to express differential properties of the *configuration* manifold in terms of the ones of the *ambient* manifold. Since morphisms, flows and tensor fields in the *configuration* and the *ambient* manifold must be carefully distinguished, in this sections and in subsequent ones, a superscript $(\cdot)^{\mathbb{C}}$ will be used to denote quantities pertaining to the former, when there are analogous quantities pertaining to the latter. Moreover geometrical objects in the two manifolds will be labeled by the prefixes $\mathbb{C}$- and $\mathcal{S}$- respectively.

3.6.1 The evaluation map

We denote by $\mathrm{EVAL}_{\mathbf{x}}$ the evaluator at $\mathbf{x} \in \mathcal{S}$ of fields on $\mathcal{S}$. A trajectory of a dynamical system through a configuration $\gamma_t \in \mathrm{C}^1(\mathcal{B} ; \mathcal{S})$ is described by a time-parametrized $\mathbb{C}$-curve $\gamma^{\mathbb{C}} \in \mathrm{C}^1(I ; \mathbb{C})$ with $\gamma^{\mathbb{C}}_t(\gamma_t) = \gamma_t$. The images of the trajectory are placements $\boldsymbol{\Omega}_\tau := \gamma_\tau(\mathcal{B})$ with $\tau \in I$.

The *displacement* from the placement $\boldsymbol{\Omega}_t$ to the placement $\boldsymbol{\Omega}_\tau$ is the diffeomorphism: $\gamma_{\tau,t} := \gamma_\tau \circ {\gamma_t}^{-1} \in \mathrm{C}^1(\boldsymbol{\Omega}_t ; \boldsymbol{\Omega}_\tau)$. To a trajectory in the configuration manifold, there corresponds a sheaf of trajectories $\mathrm{EVAL}_{\mathbf{x}}(\gamma_{\tau,t}) \in \mathrm{C}^1(I ; \mathcal{S})$, also denoted by $\mathrm{EVAL}_{\mathbf{x}}(\gamma^{\mathbb{C}}(\gamma_t)) \in \mathrm{C}^1(I ; \mathcal{S})$, through the points $\mathbf{x} \in \boldsymbol{\Omega}_t = \gamma_t(\mathcal{B})$, so that $\gamma_{\tau,t}(\mathbf{x}) = \gamma_\tau({\gamma_t}^{-1}(\mathbf{x})) \in \boldsymbol{\Omega}_\tau$.

The *velocity* of a particle $\mathbf{p} \in \mathcal{B}$ is the time-derivative $\dot{\gamma_t}(\mathbf{p}) = \partial_{\tau=t}\, \gamma_\tau(\mathbf{p}) \in \mathbb{T}_{\gamma_t(\mathbf{p})}\mathcal{S}$, and the velocity field at $\gamma_t \in \mathrm{C}^1(\mathcal{B} ; \mathcal{S})$, given by $\mathbf{v}^{\mathbb{C}}(\gamma_t) = \dot{\gamma_t} \in \mathrm{C}^1(\mathcal{B} ; \mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S})$, is a section of the pull-back bundle $\mathrm{C}^1(\gamma_t {\downarrow} \mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S} ; \mathcal{B})$. With a little abuse the same notation is also adopted for the corresponding velocity field on the position $\boldsymbol{\Omega}_t := \gamma_t(\mathcal{B})$, given by $\mathbf{v}^{\mathbb{C}}(\gamma_t) := \partial_{\tau=t}\, \gamma_{\tau,t} \in \mathrm{C}^1(\boldsymbol{\Omega}_t ; \mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S})$ which is a section of the vector bundle $\mathrm{C}^1(\mathbb{T}_{\mathbf{\Omega}_t}\mathcal{S}\,;\mathbf{\Omega}_t)$. Along the trajectory $\gamma^{\mathbb{C}}(\gamma_t) \in \mathrm{C}^1(I\,;\mathbb{C})$ in the configuration manifold, a virtual flow $\boldsymbol{\varphi}^{\mathbb{C}}(\gamma_t) \in \mathrm{C}^1(I\,;\mathbb{C})$ defines a virtual velocity field given by $\delta\mathbf{v}^{\mathbb{C}}(\gamma_t) := \partial_{\lambda=0}\,\boldsymbol{\varphi}^{\mathbb{C}}_\lambda(\gamma_t) \in \mathbb{T}_{\gamma_t}\mathbb{C}$. A $\mathbb{C}$-curve $\boldsymbol{\varphi}^{\mathbb{C}}(\gamma) \in \mathrm{C}^1(I\,;\mathbb{C})$ with $\boldsymbol{\varphi}^{\mathbb{C}}_0(\gamma) = \gamma \in \mathrm{C}^1(\mathcal{B}\,;\mathcal{S})$ is associated with a sheaf of $\mathcal{S}$-curves $\boldsymbol{\varphi}(\mathbf{x}) = \mathrm{EVAL}_{\mathbf{x}}(\boldsymbol{\varphi}^{\mathbb{C}}(\gamma)) \in \mathrm{C}^1(I\,;\mathcal{S})$ with $\boldsymbol{\varphi}_0(\mathbf{x}) = \mathbf{x} \in \mathbf{\Omega} = \gamma(\mathcal{B})$. Then, setting $\delta\mathbf{v}^{\mathbb{C}}(\gamma) := \partial_{\lambda=0}\,\boldsymbol{\varphi}^{\mathbb{C}}_\lambda(\gamma) \in \mathbb{T}_\gamma\mathbb{C}$ and $\delta\mathbf{v}(\mathbf{x}) := \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda(\mathbf{x}) \in \mathbb{T}_{\mathbf{x}}\mathcal{S}$, with $\mathbf{x} \in \mathbf{\Omega}$, we have that

$$\delta\mathbf{v}(\mathbf{x}) := \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda(\mathbf{x}) = \partial_{\lambda=0}\,\mathrm{EVAL}_{\mathbf{x}}(\boldsymbol{\varphi}^{\mathbb{C}}_\lambda(\gamma)) = \mathrm{EVAL}_{\mathbf{x}}(\delta\mathbf{v}^{\mathbb{C}}(\gamma))\,,$$

with $\delta\mathbf{v} \in \mathrm{C}^1(\mathbf{\Omega}\,;\mathbb{T}_{\mathbf{\Omega}}\mathcal{S})$.

Definition 3.6.1 (Induced connection) *The special geometric feature of the configuration manifold $\mathbb{C}$ permits to define a connection induced by a given connection in the finite dimensional ambient manifold $\mathcal{S}$. The procedure is best described in terms of parallel transport and consists in performing the parallel transport of a vector field, along a $\mathbb{C}$-curve from one configuration to another one, by transporting pointwise the vectors along the sheaf of $\mathcal{S}$-curves in the ambient manifold corresponding to the $\mathbb{C}$-curve. Setting $\mathbf{v}^{\mathbb{C}} = \partial_{\lambda=0}\,\boldsymbol{\varphi}^{\mathbb{C}}_\lambda$, the covariant derivatives are related by the formula:*

$$\begin{aligned}
\mathrm{EVAL}_{\mathbf{x}}(\nabla^{\mathbb{C}}_{\mathbf{v}^{\mathbb{C}}}\,\mathbf{u}^{\mathbb{C}}(\gamma)) &= \mathrm{EVAL}_{\mathbf{x}}(\partial_{\lambda=0}\,\boldsymbol{\varphi}^{\mathbb{C}}_\lambda\!\Downarrow \mathbf{u}^{\mathbb{C}}(\boldsymbol{\varphi}^{\mathbb{C}}_\lambda(\gamma)))\\
&= \partial_{\lambda=0}\,\boldsymbol{\varphi}_\lambda\!\Downarrow \mathbf{u}(\boldsymbol{\varphi}_\lambda(\mathbf{x}))\\
&= \nabla_{\mathbf{v}}\mathbf{u}(\mathbf{x})\,, \qquad \forall\,\mathbf{x} \in \mathbf{\Omega} = \gamma(\mathcal{B})\,,
\end{aligned}$$

where $\boldsymbol{\varphi}_\lambda(\mathbf{x}) = \mathrm{EVAL}_{\mathbf{x}}(\boldsymbol{\varphi}^{\mathbb{C}}_\lambda(\gamma))$ *and* $\mathbf{u}(\mathbf{x}) = \mathrm{EVAL}_{\mathbf{x}}(\mathbf{u}^{\mathbb{C}}(\gamma))$.

The following result is at the core of the theory of continuous dynamical systems developed in this paper. Its proof is based on the tensoriality property of the torsion of a connection and on a simple but tricky geometrical construction of vector fields associated with a given pair of vectors in the configuration manifold.

The naturality result provided by Lemma 3.6.1 will be resorted to as an essential ingredient in the proof of Theorem 3.6.2.

Lemma 3.6.1 (Evaluation of the torsion) *Let ∇ be a connection in the ambient manifold $\mathcal{S}$ with torsion* TORS *and $\nabla^{\mathbb{C}}$ be the induced connection in the configuration manifold $\mathbb{C}$ with torsion $\mathrm{TORS}^{\mathbb{C}}$. Then, the torsion $\mathrm{TORS}^{\mathbb{C}}$ evaluated at a pair of $\mathbb{C}$-vectors $\mathbf{v}^{\mathbb{C}}_\gamma, \mathbf{u}^{\mathbb{C}}_\gamma \in \mathbb{T}_\gamma\mathbb{C}$ is a $\mathcal{S}$-vector field on $\mathbf{\Omega} = \gamma(\mathcal{B})$*

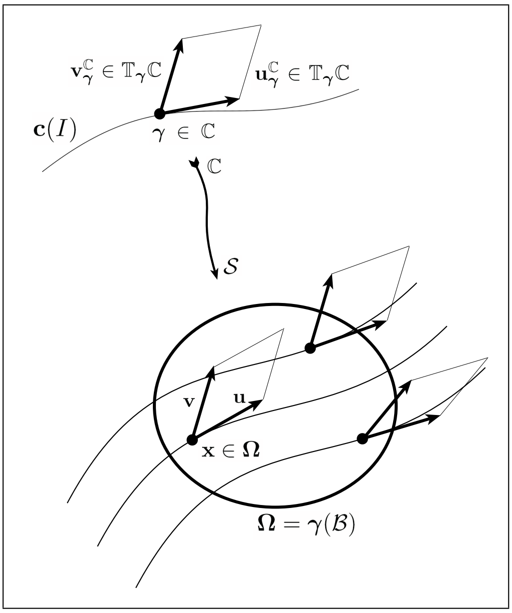

Figure 3.15: Sheaf of trajectories

whose value at a point $\mathbf{x} \in \boldsymbol{\Omega}$ *is equal to the torsion* TORS *evaluated at the pair of vectors* $\mathbf{v}(\mathbf{x}) = \text{EVAL}_{\mathbf{x}}(\mathbf{v}_\gamma^{\mathbb{C}}), \mathbf{u}(\mathbf{x}) = \text{EVAL}_{\mathbf{x}}(\mathbf{u}_\gamma^{\mathbb{C}}) \in \mathbb{T}_{\mathbf{x}}\mathcal{S}$ *:*

$$\text{EVAL}_{\mathbf{x}}(\text{TORS}^{\mathbb{C}}(\mathbf{v}_\gamma^{\mathbb{C}}, \mathbf{u}_\gamma^{\mathbb{C}})) = \text{TORS}(\mathbf{v}(\mathbf{x}), \mathbf{u}(\mathbf{x}))\,,$$

i.e. the torsion is natural with respect to the evaluation map.

Proof. Let us consider a pair $\mathbf{v}_\gamma^{\mathbb{C}}, \mathbf{u}_\gamma^{\mathbb{C}} \in \mathbb{T}_\gamma \mathbb{C}$ of $\mathbb{C}$-vectors, and the plane spanned by them. A 2-D submanifold of $\mathbb{C}$ passing through $\gamma \in \mathbb{C}$ and tangent there to this plane, is generated as follows. First we draw a curve $c \in \mathrm{C}^1(I\,;\mathbb{C})$ having the vector $\mathbf{u}_\gamma^{\mathbb{C}} \in \mathbb{T}_\gamma \mathbb{C}$ as tangent at $c(0) = \gamma(t) \in \mathbb{C}$ and denote the field of tangent vectors by $\mathbf{u}^{\mathbb{C}}(c(t)) := \partial_{\tau=t}\, c(\tau) \in \mathrm{C}^1(c(I)\,;\mathbb{T}\mathbb{C})$. Then the vector $\mathbf{v}_\gamma^{\mathbb{C}} \in \mathbb{T}_\gamma \mathbb{C}$ is extended to a vector field $\mathbf{v}^{\mathbb{C}} \in \mathrm{C}^1(c(I)\,;\mathbb{T}\mathbb{C})$ along this curve. Hence an extrusion of the curve $c \in \mathrm{C}^1(I\,;\mathbb{C})$ is performed by a flow $\boldsymbol{\varphi}_\lambda^{\mathbb{C}} \in \mathrm{C}^1(c(I)\,;\mathbb{C})$ with $\lambda \in J$ and velocity $\mathbf{v}^{\mathbb{C}} = \partial_{\mu=\lambda}\, \boldsymbol{\varphi}_\mu^{\mathbb{C}} \in \mathrm{C}^1(\boldsymbol{\varphi}_\lambda^{\mathbb{C}}(c(I))\,;\mathbb{T}\mathbb{C})$ such that $\mathbf{v}^{\mathbb{C}}(\gamma) = \mathbf{v}_\gamma^{\mathbb{C}}$. This generates a 2-D submanifold $\boldsymbol{\Sigma} \subset \mathbb{C}$ around $\gamma \in \mathbb{C}$. At last the tangent vector field $\mathbf{u}^{\mathbb{C}} \in \mathrm{C}^1(c(I)\,;\mathbb{T}c(I))$ is extended to a vector field $\mathbf{u}^{\mathbb{C}} \in \mathrm{C}^1(\boldsymbol{\Sigma}\,;\mathbb{T}\boldsymbol{\Sigma})$.

The extrusion of the curve $c \in C^1(I\,;\mathbb{C})$ defines a chart on $\boldsymbol{\Sigma}$ with origin at $\gamma \in \mathbb{C}$ and coordinates $(t\,,\lambda) \in I \times J \subset \Re^2$. The pair of vector fields $\mathbf{v}^{\mathbb{C}}, \mathbf{u}^{\mathbb{C}} \in C^1(\boldsymbol{\Sigma}\,;\mathbb{T}\boldsymbol{\Sigma})$, which at $\gamma \in \mathbb{C}$ take the values $\mathbf{v}^{\mathbb{C}}_\gamma, \mathbf{u}^{\mathbb{C}}_\gamma \in \mathbb{T}_\gamma\boldsymbol{\Sigma}$, provide a mobile frame associated with this coordinate system. If the extension of the vector field $\mathbf{u}^{\mathbb{C}} \in C^1(c(I)\,;\mathbb{T}c(I))$ to a vector field $\mathbf{u}^{\mathbb{C}} \in C^1(\boldsymbol{\Sigma}\,;\mathbb{T}\boldsymbol{\Sigma})$ is performed by pushing it along the flow $\boldsymbol{\varphi}^{\mathbb{C}}_\lambda \in C^1(c(I)\,;\mathbb{C})$, the frame is natural and the Lie bracket of the pair $\mathbf{v}^{\mathbb{C}}, \mathbf{u}^{\mathbb{C}} \in C^1(\boldsymbol{\Sigma}\,;\mathbb{T}\boldsymbol{\Sigma})$ vanishes identically on $\boldsymbol{\Sigma} \subset \mathbb{C}$. This is the choice which leads to the proof of the naturality property in the statement. The construction illustrated above reproduces itself at any point of the manifold $\boldsymbol{\Omega} \subset \mathcal{S}$ thus generating around each $\mathbf{x} \in \boldsymbol{\Omega}$ a 2-D submanifold $\boldsymbol{\Sigma}_{\mathbf{x}} \subset \mathcal{S}$ spanned by the coordinate system $(t\,,\lambda) \in I \times J \subset \Re^2$ and by the frame $(\mathbf{v}\,,\mathbf{u})$ with $\mathbf{v} \in C^1(\boldsymbol{\Sigma}_{\mathbf{x}}\,;\mathbb{T}\boldsymbol{\Sigma}_{\mathbf{x}})$ given by $\mathbf{v}(\mathbf{y}) = \text{EVAL}_{\mathbf{y}}(\mathbf{v}^{\mathbb{C}}(\boldsymbol{\xi}))$ where $\mathbf{y} \in \boldsymbol{\xi}(\mathcal{B})$ with $\boldsymbol{\xi} \in \boldsymbol{\Sigma}$ and similarly for $\mathbf{u} \in C^1(\boldsymbol{\Sigma}_{\mathbf{x}}\,;\mathbb{T}\boldsymbol{\Sigma}_{\mathbf{x}})$. Then, in particular, we have that

$$\text{EVAL}_{\mathbf{x}}([\mathbf{v}^{\mathbb{C}}\,,\mathbf{u}^{\mathbb{C}}](\gamma)) = [\mathbf{v}\,,\mathbf{u}](\mathbf{x}) = 0\,, \quad \forall\, \mathbf{x} \in \boldsymbol{\Omega} = \gamma(\mathcal{B})\,.$$

By tensoriality, to evaluate the torsion of the connection $\nabla^{\mathbb{C}}$ at a pair of $\mathbb{C}$-vectors $\mathbf{v}^{\mathbb{C}}_\gamma, \mathbf{u}^{\mathbb{C}}_\gamma \in \mathbb{T}_\gamma\mathbb{C}$, we may extend them according to the previously illustrated procedure.

Then, applying the formula for the torsion of a pair of vector fields, we get:

$$\begin{aligned}
\text{EVAL}_{\mathbf{x}}(\text{TORS}^{\mathbb{C}}(\mathbf{v}^{\mathbb{C}}_\gamma, \mathbf{u}^{\mathbb{C}}_\gamma)) &= \text{EVAL}_{\mathbf{x}}(\text{TORS}^{\mathbb{C}}_\gamma(\mathbf{v}^{\mathbb{C}}, \mathbf{u}^{\mathbb{C}})) \\
&= \text{EVAL}_{\mathbf{x}}((\nabla^{\mathbb{C}}_{\mathbf{v}^{\mathbb{C}}}\mathbf{u}^{\mathbb{C}} - \nabla^{\mathbb{C}}_{\mathbf{u}^{\mathbb{C}}}\mathbf{v}^{\mathbb{C}} - [\mathbf{v}^{\mathbb{C}}\,,\mathbf{u}^{\mathbb{C}}])(\gamma)) \\
&= \text{EVAL}_{\mathbf{x}}((\nabla^{\mathbb{C}}_{\mathbf{v}^{\mathbb{C}}}\mathbf{u}^{\mathbb{C}} - \nabla^{\mathbb{C}}_{\mathbf{u}^{\mathbb{C}}}\mathbf{v}^{\mathbb{C}})(\gamma)) \\
&= (\nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v})(\mathbf{x}) \\
&= (\nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v} - [\mathbf{v}\,,\mathbf{u}])(\mathbf{x}) \\
&= \text{TORS}(\mathbf{v}(\mathbf{x}), \mathbf{u}(\mathbf{x}))\,,
\end{aligned}$$

the last equality being again due to the tensoriality of the torsion. ■

3.6.2 Law of motion

A proper formulation of the law of motion for a continuous body, in an *ambient* finite dimensional riemannian manifold $(\mathcal{S}\,,\mathbf{g})$, needs a sufficiently general definition of the linear space of spatial virtual velocity fields on the position

$\mathbf{\Omega}_t := \gamma_t(\mathcal{B})$ at time $t \in I$ along the trajectory in the ambient manifold. To this end, let us give the following definitions. A *patchwork* $\text{PAT}(\mathbf{\Omega}_t)$ is a finite family of open connected, non-overlapping subsets of $\mathbf{\Omega}_t$, called elements, such that the union of their closures is a covering for $\mathbf{\Omega}_t$. The set of all patchworks of $\mathbf{\Omega}_t$ is a directed set for the relation *finer than* and the coarsest patchwork finer than two given ones $\text{PAT}_1(\mathbf{\Omega}_t)$ and $\text{PAT}_2(\mathbf{\Omega}_t)$ is the *grid* $\text{PAT}_1(\mathbf{\Omega}_t) \wedge \text{PAT}_2(\mathbf{\Omega}_t)$. The kinematic space $\text{KIN}(\mathbf{\Omega}_t)$ is made up of vector fields $\mathbf{v}_t \in \mathrm{C}^1(\mathbf{\Omega}_t \,; \mathbb{T}_{\mathbf{\Omega}_t}\mathcal{S})$ which are square integrable with a distributional gradient which is square integrable in the elements of a patchwork $\text{PAT}_{\mathbf{v}_t}(\mathbf{\Omega}_t)$. This space is pre-HILBERT with the positive definite symmetric bilinear form:

$$\int_{\text{PAT}_{(\mathbf{v}_t,\mathbf{w}_t)}(\mathbf{\Omega}_t)} (\mathbf{g}\,(\mathbf{v}_t, \mathbf{w}_t) + \langle \nabla \mathbf{v}_t, \nabla \mathbf{w}_t \rangle_{\mathbf{g}})\, \boldsymbol{\mu}\,,$$

where $\text{PAT}_{(\mathbf{v}_t,\mathbf{w}_t)}(\mathbf{\Omega}_t) = \text{PAT}_{\mathbf{v}_t}(\mathbf{\Omega}_t) \wedge \text{PAT}_{\mathbf{w}_t}(\mathbf{\Omega}_t)$ and $\langle \cdot, \cdot \rangle_{\mathbf{g}}$ is the inner product between tensors induced by the metric $\mathbf{g}$. A continuous body at $\mathbf{\Omega}_t$ is defined by a fixed patchwork $\text{PAT}(\mathbf{\Omega}_t)$ and by a closed linear subspace of conforming virtual displacements $\text{CONF}(\mathbf{\Omega}_t) \subset \text{KIN}(\mathbf{\Omega}_t)$ such that all of its vector fields have $\text{PAT}(\mathbf{\Omega}_t)$ as a regularity patchwork. Then $\text{CONF}(\mathbf{\Omega}_t)$ is a HILBERT space for the topology induced by $\text{KIN}(\mathbf{\Omega}_t)$. Since $\text{CONF}(\mathbf{\Omega}_t)$ is a linear space, this definition includes any linear or affine kinematical constraint. Non-linear constraints must rather be modeled by suitable constitutive laws described by fiberwise monotone maximal graphs in the WHITNEY bundle whose fiber is the product of tangent vector and covector spaces [173]. In the tangent bundle $\boldsymbol{\tau}_{\mathcal{S}} \in \mathrm{C}^1(\mathbb{T}\mathcal{S}\,; \mathcal{S})$, the subbundle of infinitesimal isometries (or rigid body velocities) at the position $\mathbf{\Omega}_t$ is denoted by $\text{RIG}(\mathbf{\Omega}_t)$. These are vector fields $\delta\mathbf{v}_t \in \mathrm{C}^1(\mathbf{\Omega}_t \,; \mathbb{T}_{\mathbf{\Omega}_t}\mathcal{S})$ characterized by the condition $\mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g} = 0$. The property of the LIE derivative: $\mathcal{L}_{[\mathbf{u}\,,\mathbf{v}]} = [\mathcal{L}_{\mathbf{u}}\,, \mathcal{L}_{\mathbf{v}}]$ for any pair of tangent vector fields $\mathbf{u}, \mathbf{v} \in \mathrm{C}^1(\mathcal{S}\,; \mathbb{T}\mathcal{S})$, ensures that the subbundle $\text{RIG}(\mathbf{\Omega}_t)$ is involutive, i.e. that $\mathcal{L}_{\mathbf{u}}\mathbf{g} = \mathcal{L}_{\mathbf{v}}\mathbf{g} = 0 \implies \mathcal{L}_{[\mathbf{u}\,,\mathbf{v}]}\mathbf{g} = 0$, and hence integrable by FROBENIUS theorem, see e.g. [3], [80]. This property is at the basis of the classical analytical dynamics which considers dynamical trajectories evolving in a leaf of the foliation induced by the rigidity condition on the velocity fields. Let ∇ be a connection in the ambient manifold $\{\mathcal{S}\,, \mathbf{g}\}$ and $\text{TORS} \in \Lambda^2(\mathcal{S}\,; \mathbb{T}\mathcal{S})$ be the tangent-valued torsion 2-form: $\text{TORS}(\mathbf{v}, \mathbf{u}) := (\nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v}) - [\mathbf{v}\,, \mathbf{u}]$, [3]. We quote hereafter a generalized version of EULER's classical formula for the stretching $\frac{1}{2}(\mathcal{L}_{\mathbf{v}}\,\mathbf{g})$, valid in an ambient riemannian manifold with an arbitrary connection.

Lemma 3.6.2 *Let* $\{\mathcal{S}\,, \mathbf{g}\}$ *be a riemannian manifold,* ∇ *a connection in* $\mathcal{S}$ *with torsion* $\text{TORS} \in \Lambda^2(\mathcal{S}\,; \mathbb{T}\mathcal{S})$ *and* $\text{TORS}(\mathbf{v})$ *the field of linear operators* *defined by:*

$$\mathrm{T\textsc{ors}}(\mathbf{v}) \cdot \mathbf{u} = \textsc{tors}(\mathbf{v}, \mathbf{u}) , \quad \forall\, \mathbf{v}, \mathbf{u} \in \mathrm{C}^1(\mathcal{S}\,; \mathbb{T}\mathcal{S}) .$$

Then, for any vector field $\mathbf{v} \in \mathrm{C}^1(\mathcal{S}\,; \mathbb{T}\mathcal{S})$ *:*

$$\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}}\, \mathbf{g}) = \mathbf{g} \circ (\mathrm{sym}\, \nabla\mathbf{v}) + \tfrac{1}{2}(\nabla_{\mathbf{v}}\, \mathbf{g}) + \mathbf{g} \circ (\mathrm{sym}\, \mathrm{T\textsc{ors}}(\mathbf{v})) .$$

If ∇ *is* LEVI-CIVITA*, i.e. metric,* $\nabla_{\mathbf{v}}\, \mathbf{g} = 0$*, and torsion-free,* $\mathrm{T\textsc{ors}}(\mathbf{v}) = 0$*,* EULER*'s formula for the stretching is recovered:*

$$\tfrac{1}{2}(\mathcal{L}_{\mathbf{v}}\, \mathbf{g}) = \mathbf{g} \circ (\mathrm{sym}\, \nabla\mathbf{v}) .$$

Proof. Applying the LEIBNIZ rule to the LIE derivative and to the covariant derivative, we have that, for any vector fields $\mathbf{v}, \mathbf{u}, \mathbf{w} \in \mathrm{C}^1(\mathcal{S}\,; \mathbb{T}\mathcal{S})$:

$$\begin{aligned}
(\mathcal{L}_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) &= \mathcal{L}_{\mathbf{v}}\, (\mathbf{g}\, (\mathbf{u}, \mathbf{w})) - \mathbf{g}\, (\mathcal{L}_{\mathbf{v}}\mathbf{u}, \mathbf{w}) - \mathbf{g}\, (\mathbf{u}, \mathcal{L}_{\mathbf{v}}\mathbf{w}) , \\
(\nabla_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) &= \nabla_{\mathbf{v}}\, (\mathbf{g}\, (\mathbf{u}, \mathbf{w})) - \mathbf{g}\, (\nabla_{\mathbf{v}}\mathbf{u}, \mathbf{w}) - \mathbf{g}\, (\mathbf{u}, \nabla_{\mathbf{v}}\mathbf{w}) .
\end{aligned}$$

Since the LIE derivative and the covariant derivative of a scalar field coincide, we also have that $\mathcal{L}_{\mathbf{v}}\, (\mathbf{g}\, (\mathbf{u}, \mathbf{w})) = \nabla_{\mathbf{v}}\, (\mathbf{g}\, (\mathbf{u}, \mathbf{w}))$ and hence:

$$\begin{aligned}
(\mathcal{L}_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) = (\nabla_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) &+ \mathbf{g}\, (\nabla_{\mathbf{v}}\mathbf{u}, \mathbf{w}) + \mathbf{g}\, (\mathbf{u}, \nabla_{\mathbf{v}}\mathbf{w}) \\
&- \mathbf{g}\, (\mathcal{L}_{\mathbf{v}}\mathbf{u}, \mathbf{w}) - \mathbf{g}\, (\mathbf{u}, \mathcal{L}_{\mathbf{v}}\mathbf{w}) .
\end{aligned}$$

Moreover, since $\textsc{tors}(\mathbf{v}, \mathbf{u}) := (\nabla_{\mathbf{v}}\mathbf{u} - \nabla_{\mathbf{u}}\mathbf{v}) - [\mathbf{v}, \mathbf{u}]$ we may write

$$\begin{aligned}
(\mathcal{L}_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) = (\nabla_{\mathbf{v}}\, \mathbf{g})(\mathbf{u}, \mathbf{w}) &+ \mathbf{g}\, (\textsc{tors}(\mathbf{v}, \mathbf{u}), \mathbf{w}) + \mathbf{g}\, (\nabla_{\mathbf{u}}\mathbf{v}, \mathbf{w}) \\
&+ \mathbf{g}\, (\textsc{tors}(\mathbf{v}, \mathbf{w}), \mathbf{u}) + \mathbf{g}\, (\nabla_{\mathbf{w}}\mathbf{v}, \mathbf{u}) ,
\end{aligned}$$

which gives the result. ■

Let $L_t \in \mathrm{C}^1(\mathbb{T}_{\mathbf{\Omega}_t}\mathcal{S}\,; \Re)$ be the lagrangian per unit mass at the position $\mathbf{\Omega}_t := \gamma_t(\mathcal{B})$, $\boldsymbol{\mu}$ be the volume form in $\mathcal{S}$ and $\mathbf{m}_t = \rho_t\, \boldsymbol{\mu}$ be the mass form related to the scalar density $\rho_t \in \mathrm{C}^1(\mathbf{\Omega}_t\,; \Re)$.

In continuum dynamics the lagrangian functional on the tangent bundle to the configuration manifold: $L^{\mathbb{C}}_t \in \mathrm{C}^1(\mathbb{T}\mathbb{C}\,; \Re)$, is defined by the integral:

$$(L^{\mathbb{C}}_t \circ \mathbf{v}^{\mathbb{C}})(\gamma_t) := \int_{\mathbf{\Omega}_t} (L_t \circ \mathbf{v}_t)\, \mathbf{m}_t ,$$

where $\mathbf{v}_t(\mathbf{x}) = \textsc{eval}_{\mathbf{x}}(\mathbf{v}^{\mathbb{C}}(\gamma_t))$ for $\mathbf{x} \in \mathbf{\Omega}_t = \gamma_t(\mathcal{B})$.

The next theorem provides the expression of the law of dynamics in an ambient riemannian manifold $\{\mathcal{S}\, , \mathbf{g}\}$, independent of a connection. The volume form $\boldsymbol{\mu}$ is the one induced by the metric tensor $\mathbf{g}$.

Theorem 3.6.1 (Law of motion in the ambient manifold) *The law of motion of a continuous dynamical system in the ambient riemannian manifold* $\{\mathcal{S}\,,\mathbf{g}\}$*, is expressed by the variational condition:*

$$\partial_{\tau=t} \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \langle\, d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau \,\rangle\, \mathbf{m}_\tau - \partial_{\lambda=0} \int_{\boldsymbol{\varphi}_{\lambda,t}(\boldsymbol{\Omega}_t)} (L_t \circ \boldsymbol{\varphi}_{\lambda,t}{\uparrow}\mathbf{v}_t)\, \mathbf{m}_t \\ = \langle\, \mathbf{F}^{\mathbb{C}}_t(\gamma_t), \delta\mathbf{v}^{\mathbb{C}}_t \,\rangle\,,$$

for any virtual flow $\boldsymbol{\varphi}_{\lambda,t} \in \mathrm{C}^1(\boldsymbol{\Omega}_t\,;\mathcal{S})$ *at time* $t \in I$ *such that the virtual velocity field* $\delta\mathbf{v}_t := \partial_{\lambda=0}\, \boldsymbol{\varphi}_{\lambda,t} \in \mathrm{C}^1(\boldsymbol{\Omega}_t\,;\mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S})$ *is conforming and isometric, i.e.* $\delta\mathbf{v}_t \in \mathrm{C{\small ONF}}(\boldsymbol{\Omega}_t) \cap \mathrm{RIG}(\boldsymbol{\Omega}_t)$.

Proof. According to Theorem 2.2.1, the law of motion in the configuration manifold is expressed by LAGRANGE's variational condition:

$$\partial_{\tau=t}\, \langle\, d_{\mathrm{F}} L^{\mathbb{C}}_\tau(\mathbf{v}^{\mathbb{C}}_\tau), \delta\mathbf{v}^{\mathbb{C}}_\tau \,\rangle - \partial_{\lambda=0}\, L^{\mathbb{C}}_t(T\boldsymbol{\varphi}^{\mathbb{C}}_\lambda \cdot \mathbf{v}^{\mathbb{C}}_t) = \langle\, \mathbf{F}^{\mathbb{C}}_t(\gamma_t), \delta\mathbf{v}^{\mathbb{C}}_t \,\rangle\,,$$

where, being $\mathbf{v}_t(\mathbf{x}) = \mathrm{EVAL}_{\mathbf{x}}(\mathbf{v}^{\mathbb{C}}(\gamma_t))$ and $\delta\mathbf{v}_t(\mathbf{x}) = \mathrm{EVAL}_{\mathbf{x}}(\delta\mathbf{v}^{\mathbb{C}}(\gamma_t))$ we have that

$$\langle\, d_{\mathrm{F}} L^{\mathbb{C}}_\tau(\mathbf{v}^{\mathbb{C}}_\tau), \delta\mathbf{v}^{\mathbb{C}}_\tau \,\rangle = \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \langle\, d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau \,\rangle\, \mathbf{m}_\tau\,,$$

where $\delta\mathbf{v}_\tau := \partial_{\lambda=0}\, \boldsymbol{\varphi}_{\lambda,\tau} \in \mathrm{C}^1(\boldsymbol{\Omega}_\tau\,;\mathbb{T}_{\boldsymbol{\Omega}_\tau}\mathcal{S})$ and $\boldsymbol{\varphi}_{\lambda,t}(\mathbf{x}) = \mathrm{EVAL}_{\mathbf{x}}(\boldsymbol{\varphi}^{\mathbb{C}}_\lambda(\gamma_t))$. On the other hand:

$$L^{\mathbb{C}}_t(T\boldsymbol{\varphi}^{\mathbb{C}}_\lambda \cdot \mathbf{v}^{\mathbb{C}}_t) = (L^{\mathbb{C}}_t \circ \boldsymbol{\varphi}^{\mathbb{C}}_\lambda{\uparrow}\mathbf{v}^{\mathbb{C}}_t \circ \boldsymbol{\varphi}^{\mathbb{C}}_\lambda)(\gamma_t) = \int_{\boldsymbol{\varphi}_{\lambda,t}(\boldsymbol{\Omega}_t)} L_t(\boldsymbol{\varphi}_{\lambda,t}{\uparrow}\mathbf{v}_t)\, \mathbf{m}_t\,.$$

Substituting, we get the result. ■

Each term at the l.h.s. of the law of motion in Theorem 3.6.1 depends on the choice of the family of virtual flows $\boldsymbol{\varphi}_{\lambda,\tau} \in \mathrm{C}^1(\boldsymbol{\Omega}_\tau\,;\mathbb{T}\mathcal{S})$ with $\tau \in I$. However, in Theorem 3.6.2 it will be proved that the expression at the l.h.s. of the law of motion defines a bounded linear functional $\mathcal{F} \in \mathrm{C{\small ONF}}^*(\boldsymbol{\Omega}_t)$.

This basic result, which is a generalized version of Euler's law of motion makes an essential recourse to the notion of a connection in the ambient manifold and of the induced connection in the infinite dimensional configuration manifold. The proof is based on a subtle argument whose key points are the vanishing of the LIE derivative leading to the expression of covariant derivatives in terms of the torsion and the tensoriality property of the torsion of a connection. Moreover to get the result we need to make an assumption of mass conservation along virtual flows.

Precisely in the sequel we will assume that, in performing the variations, the following condition is fulfilled.

Ansatz 3.6.1 (Virtual mass-conservation) *The virtual flows drag the mass-form or equivalently along the virtual flows the mass of any sub-body is preserved, that is:*

$$\mathcal{L}_{\delta \mathbf{v}_t} \mathbf{m}_t = 0 \iff \partial_{\lambda=0} \int_{\boldsymbol{\varphi}_{\lambda,t}(\mathcal{P})} \mathbf{m}_t = 0 \,, \quad \forall\, \mathcal{P} \subseteq \boldsymbol{\Omega}_t \,.$$

This assumption amounts in defining a proper way of extending the mass-form to positions of the body outside the trajectory and mimics the one tacitly made in analytical mechanics in assuming that the material particles retain their mass-measure along the variations.

Theorem 3.6.2 (Generalized Euler's law of motion) *Let* ∇ *be a connection in the ambient manifold* $\mathcal{S}$ *with parallel transport* $\Uparrow$ *and torsion* TORS. *The law of motion is then expressed by the variational condition:*

$$\partial_{\tau=t} \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \langle\, d_{\text{F}} L_\tau(\mathbf{v}_\tau), \gamma_{\tau,t}\!\Uparrow \delta \mathbf{v}_t \,\rangle\, \mathbf{m}_\tau - \int_{\boldsymbol{\Omega}_t} \langle\, d_{\text{B}} L_t(\mathbf{v}_t), \delta \mathbf{v}_t \,\rangle\, \mathbf{m}_t$$

$$+ \int_{\boldsymbol{\Omega}_t} \langle\, d_{\text{F}} L_t(\mathbf{v}_t), \text{TORS}(\mathbf{v}_t, \delta \mathbf{v}_t) \,\rangle\, \mathbf{m}_t = \langle\, \mathbf{F}^{\mathbb{C}}_t(\gamma_t), \delta \mathbf{v}^{\mathbb{C}}_t \,\rangle \,,$$

for any virtual velocity field $\delta \mathbf{v}_t \in \text{CONF}(\boldsymbol{\Omega}_t) \cap \text{RIG}(\boldsymbol{\Omega}_t)$.

Proof. By Theorem **??**, the l.h.s. of the law of motion in the configuration manifold, according to the connection $\nabla^{\mathbb{C}}$ there induced by the connection ∇ in the ambient manifold, writes:

$$\partial_{\tau=t} \langle\, d_{\text{F}} L^{\mathbb{C}}_\tau(\mathbf{v}^{\mathbb{C}}_\tau), \gamma_{\tau,t}\!\Uparrow \delta \mathbf{v}^{\mathbb{C}}_t \,\rangle - \langle\, d_{\text{B}} L^{\mathbb{C}}_t(\mathbf{v}^{\mathbb{C}}_t), \delta \mathbf{v}^{\mathbb{C}}_t \,\rangle$$
$$+ \langle\, d_{\text{F}} L^{\mathbb{C}}_t(\mathbf{v}^{\mathbb{C}}_t), \text{TORS}^{\mathbb{C}}(\mathbf{v}^{\mathbb{C}}_t, \delta \mathbf{v}^{\mathbb{C}}_t) \,\rangle \,.$$

Translating in terms of fields in the ambient manifold, by Lemma 3.6.1 we have:

$$\langle\, d_{\text{F}} L^{\mathbb{C}}_\tau(\mathbf{v}^{\mathbb{C}}_\tau), \gamma_{\tau,t}\!\Uparrow \delta \mathbf{v}^{\mathbb{C}}_t \,\rangle = \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \langle\, d_{\text{F}} L_\tau(\mathbf{v}_\tau), \gamma_{\tau,t}\!\Uparrow \delta \mathbf{v}_t \,\rangle\, \mathbf{m}_\tau \,,$$

$$\langle\, d_{\text{F}} L^{\mathbb{C}}_t(\mathbf{v}^{\mathbb{C}}_t), \text{TORS}^{\mathbb{C}}(\mathbf{v}^{\mathbb{C}}_t, \delta \mathbf{v}^{\mathbb{C}}_t) \,\rangle = \int_{\boldsymbol{\Omega}_t} \langle\, d_{\text{F}} L_t(\mathbf{v}_t), \text{TORS}(\mathbf{v}_t, \delta \mathbf{v}_t) \,\rangle\, \mathbf{m}_t \,,$$

$$
\begin{aligned}
\langle d_{\mathrm{B}} L^{\mathbb{C}}_t(\mathbf{v}^{\mathbb{C}}_t), \delta\mathbf{v}^{\mathbb{C}}_t \rangle &= \partial_{\lambda=0}\, L^{\mathbb{C}}_t(\boldsymbol{\varphi}^{\mathbb{C}}_{\lambda,t}\Uparrow \mathbf{v}^{\mathbb{C}}_t) \\
&= \partial_{\lambda=0} \int_{\boldsymbol{\varphi}_{\lambda,t}(\boldsymbol{\Omega}_t)} L_t(\boldsymbol{\varphi}_{\lambda,t}\Uparrow \mathbf{v}_t)\, \mathbf{m}_t \\
&= \int_{\boldsymbol{\Omega}_t} \partial_{\lambda=0}\, \boldsymbol{\varphi}_{\lambda,t}\!\downarrow\![L_t(\boldsymbol{\varphi}_{\lambda,t}\Uparrow \mathbf{v}_t)\, \mathbf{m}_t] \\
&= \int_{\boldsymbol{\Omega}_t} \langle d_{\mathrm{B}} L_t(\mathbf{v}_t), \delta\mathbf{v}_t \rangle\, \mathbf{m}_t + \int_{\boldsymbol{\Omega}_t} L_t(\mathbf{v}_t)\, \mathcal{L}_{\delta\mathbf{v}_t}\mathbf{m}_t\,.
\end{aligned}
$$

Setting $\mathcal{L}_{\delta\mathbf{v}_t}\mathbf{m}_t = 0$ we get the result. ■

The law of motion provided by Theorem 3.6.2 defines a bounded linear functional $\mathcal{F} \in \mathrm{C\scriptsize ONF}^*(\boldsymbol{\Omega}_t)$. Then the next theorem shows that the rigidity condition on virtual velocities may be eliminated by introducing a LAGRANGE's multiplier dual to the stretching.

The proof is based on the property that the image by the differential operator $\mathrm{sym}\,\nabla$ of any closed subspace of the HILBERT space $\mathrm{C\scriptsize ONF}(\boldsymbol{\Omega}_t)$ is a closed subspace of $\mathrm{S\scriptsize QIT}(\boldsymbol{\Omega}_t)$, the HILBERT space of square integrable tensor fields on $\boldsymbol{\Omega}_t$. In turn this property is inferred from KORN's second inequality [56], [39], [161], [165].

Theorem 3.6.3 (Law of motion in terms of a stress field) *There exists at least a square integrable twice contravariant stress tensor field* $\boldsymbol{\sigma}_t \in \mathrm{S\scriptsize QIT}(\boldsymbol{\Omega}_t)$ *such that the law of motion of a continuous dynamical system in the ambient riemannian manifold* $\{\mathcal{S}\,, \mathbf{g}\}$ *is equivalent to the variational condition:*

$$
\begin{aligned}
\partial_{\tau=t} \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \langle d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \delta\mathbf{v}_\tau \rangle\, \mathbf{m}_\tau - \partial_{\lambda=0} \int_{\boldsymbol{\varphi}_{\lambda,t}(\boldsymbol{\Omega}_t)} (L_t \circ \boldsymbol{\varphi}_{\lambda,t}\!\uparrow\!\mathbf{v}_t)\, \mathbf{m}_t \\
= \langle \mathbf{F}^{\mathbb{C}}_t(\gamma_t), \delta\mathbf{v}^{\mathbb{C}}_t \rangle - \int_{\boldsymbol{\Omega}_t} \langle \boldsymbol{\sigma}_t, \mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g} \rangle\, \boldsymbol{\mu}\,,
\end{aligned}
$$

for any virtual flow $\boldsymbol{\varphi}_{\lambda,t} \in \mathrm{C}^1(\boldsymbol{\Omega}_t\,; \mathcal{S})$ *at time* $t \in I$ *whose virtual velocity field* $\delta\mathbf{v}_t := \partial_{\lambda=0}\, \boldsymbol{\varphi}_{\lambda,t} \in \mathrm{C}^1(\boldsymbol{\Omega}_t\,; \mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S})$ *is conforming, i.e.* $\delta\mathbf{v}_t \in \mathrm{C\scriptsize ONF}(\boldsymbol{\Omega}_t)$.

Proof. The duality between the twice covariant stretching tensor $\mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g}(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathcal{S}\,; \mathbb{T}^*_{\mathbf{x}}\mathcal{S}) = BL\,(\mathbb{T}_{\mathbf{x}}\mathcal{S}, \mathbb{T}_{\mathbf{x}}\mathcal{S}\,; \Re)$ and the twice contravariant stress tensor $\boldsymbol{\sigma}_t(\mathbf{x}) \in BL\,(\mathbb{T}^*_{\mathbf{x}}\mathcal{S}\,; \mathbb{T}_{\mathbf{x}}\mathcal{S}) = BL\,(\mathbb{T}^*_{\mathbf{x}}\mathcal{S}, \mathbb{T}^*_{\mathbf{x}}\mathcal{S}\,; \Re)$ is defined by the linear invariant of their composition $(\boldsymbol{\sigma}_t \circ \mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g})(\mathbf{x}) \in BL\,(\mathbb{T}_{\mathbf{x}}\mathcal{S}\,; \mathbb{T}_{\mathbf{x}}\mathcal{S})$, that is:

$$
\langle \boldsymbol{\sigma}_t, \mathcal{L}_{\delta\mathbf{v}_\tau}\mathbf{g} \rangle := I_1(\boldsymbol{\sigma}_t \circ \mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g})\,.
$$

By the isomorphisms $\mathbf{g}^\flat(\mathbf{x}) \in BL(\mathbb{T}_\mathbf{x}\mathcal{S}\,;\mathbb{T}^*_\mathbf{x}\mathcal{S})$ and $\mathbf{g}^\sharp(\mathbf{x}) \in BL(\mathbb{T}^*_\mathbf{x}\mathcal{S}\,;\mathbb{T}_\mathbf{x}\mathcal{S})$ with $\mathbf{g}^\sharp(\mathbf{x}) = (\mathbf{g}^\flat)^{-1}(\mathbf{x})$, induced by the metric $\mathbf{g}(\mathbf{x}) \in BL(\mathbb{T}_\mathbf{x}\mathcal{S}, \mathbb{T}_\mathbf{x}\mathcal{S}\,;\Re)$ and assuming the LEVI-CIVITA connection in $\{\mathcal{S}\,,\mathbf{g}\}$, we may set

$$\begin{cases} \boldsymbol{\sigma}_t = \mathbf{T}_t \circ \mathbf{g}^\sharp \\ \mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g} = \mathbf{g}^\flat \circ (\operatorname{sym}\nabla\delta\mathbf{v}_t)\,, \end{cases}$$

with $\mathbf{T}_t(\mathbf{x}), \operatorname{sym}\nabla\delta\mathbf{v}_t(\mathbf{x}) \in BL(\mathbb{T}_\mathbf{x}\mathcal{S}\,;\mathbb{T}_\mathbf{x}\mathcal{S})$, and the inner product given by $\langle\mathbf{T}_t, \operatorname{sym}\nabla\delta\mathbf{v}_t\rangle_\mathbf{g} := I_1(\boldsymbol{\sigma}_t \circ \mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g})$. The HILBERT space $\mathrm{SQIT}(\boldsymbol{\Omega}_t)$ is identified with its dual by the RIESZ-FRÉCHET theorem (see e.g. [192], [159]). The dual operator $(\operatorname{sym}\nabla)^* \in BL(\mathrm{SQIT}(\boldsymbol{\Omega}_t)\,;\mathrm{CONF}^*(\boldsymbol{\Omega}_t))$ of the kinematic operator $\operatorname{sym}\nabla \in BL(\mathrm{CONF}(\boldsymbol{\Omega}_t)\,;\mathrm{SQIT}(\boldsymbol{\Omega}_t))$ is then defined by the identity:

$$\langle(\operatorname{sym}\nabla)^*\mathbf{T}_t, \delta\mathbf{v}_t\rangle := \int_{\mathrm{PAT}(\boldsymbol{\Omega}_t)} \langle\mathbf{T}_t, \operatorname{sym}\nabla\delta\mathbf{v}_t\rangle_\mathbf{g}\,\boldsymbol{\mu}\,,$$

for all $\delta\mathbf{v}_t \in \mathrm{CONF}(\boldsymbol{\Omega}_t)$. Now the difference between the r.h.s. and the l.h.s. of the equation of motion in Theorem 3.6.1 defines a bounded linear functional $\mathcal{F} \in \mathrm{CONF}^*(\boldsymbol{\Omega}_t)$, as was proven in Theorem 3.6.2. Moreover KORN's inequality implies that the linear subspace $\operatorname{sym}\nabla(\mathrm{CONF}(\boldsymbol{\Omega}_t))$ is closed in $\mathrm{SQIT}(\boldsymbol{\Omega}_t)$, see e.g. [159], and BANACH's closed range theorem assures that $(\operatorname{sym}\nabla)^*(\mathrm{SQIT}(\boldsymbol{\Omega}_t))$ is closed in $\mathrm{CONF}^*(\boldsymbol{\Omega}_t)$, [192]. The law of motion expressed by the variational condition in Theorem 3.6.1 may then be written as:

$$\mathcal{F} \in (\ker\operatorname{sym}\nabla)^\circ \subset (\ker\operatorname{sym}\nabla \cap \mathrm{CONF}(\boldsymbol{\Omega}_t))^\circ = (\operatorname{sym}\nabla)^*(\mathrm{SQIT}(\boldsymbol{\Omega}_t))\,,$$

where $(\bullet)^\circ$ denotes the annihilator, i.e. the closed subspace of bounded linear functionals vanishing on $\bullet$.

This means that there exists a stress tensor field $\mathbf{T}_t \in \mathrm{SQIT}(\boldsymbol{\Omega}_t)$ such that $\mathcal{F} = (\operatorname{sym}\nabla)^*\mathbf{T}_t$, that is, for all $\delta\mathbf{v}_t \in \mathrm{CONF}(\boldsymbol{\Omega}_t)$:

$$\begin{aligned} \langle\mathcal{F}, \delta\mathbf{v}_t\rangle = \langle(\operatorname{sym}\nabla)^*\mathbf{T}_t\,, \delta\mathbf{v}_t\rangle &= \int_{\mathrm{PAT}(\boldsymbol{\Omega}_t)} \langle\mathbf{T}_t, \operatorname{sym}\nabla\delta\mathbf{v}_t\rangle_\mathbf{g}\,\boldsymbol{\mu} \\ &= \int_{\mathrm{PAT}(\boldsymbol{\Omega}_t)} \langle\boldsymbol{\sigma}_t, \mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g}\rangle\,\boldsymbol{\mu}\,. \end{aligned}$$

The proof of the converse result is trivial since for rigid virtual velocity fields $\delta\mathbf{v}_t \in \mathrm{RIG}(\boldsymbol{\Omega}_t)$ the variational condition above, being $\mathcal{L}_{\delta\mathbf{v}_t}\mathbf{g} = 0$, gives: $\langle\mathcal{F}, \delta\mathbf{v}_t\rangle = 0$ which is the condition in Theorem 3.6.1. ■

It is straightforward to see that the law of dynamics of Theorem 3.6.3 implies as a simple corollary a generalized statement of E. NOETHER's theorem for continuous dynamical systems, [172]. The energy $E_t \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S}\,;\Re)$ per unit mass is defined by LEGENDRE transform: $E_t(\mathbf{v}_t) := \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_t \,\rangle - L_t(\mathbf{v}_t)\,.$

3.6.3 Special forms of the law of motion

From the general law of motion provided in Theorems 3.6.1 and 3.6.2 other expressions valid under special assumptions may be derived. The following one is the extension to continuous systems of the law of dynamics formulated by POINCARÉ in the context of analytical dynamics for systems described in terms of vector components in a mobile reference frame [7], [174].

Theorem 3.6.4 (Euler-Poincaré law of motion) *Let ∇ be a connection in the ambient manifold $\mathcal{S}$ with a distant parallel transport $\Uparrow$ and torsion* TORS *. The law of motion is then expressed by the variational condition:*

$$\partial_{\tau=t} \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \langle\, d_{\mathrm{F}} L_\tau(\mathbf{v}_\tau), \gamma_{\tau,t}\!\Uparrow \delta\mathbf{v}_t \,\rangle\, \mathbf{m}_\tau - \int_{\boldsymbol{\Omega}_t} \langle\, d_{\mathrm{B}} L_t(\mathbf{v}_t), \delta\mathbf{v}_t \,\rangle\, \mathbf{m}_t$$

$$- \int_{\boldsymbol{\Omega}_t} \langle\, d_{\mathrm{F}} L_t(\mathbf{v}_t), [\mathbf{S}(\mathbf{v}_t)\,, \mathbf{S}(\delta\mathbf{v}_t)] \,\rangle\, \mathbf{m}_t = \langle\, \mathbf{F}_t^{\mathbb{C}}(\gamma_t), \delta\mathbf{v}_t^{\mathbb{C}} \,\rangle\,,$$

for any virtual velocity field $\delta\mathbf{v}_t \in \mathrm{C}\textsc{onf}(\boldsymbol{\Omega}_t) \cap \mathrm{RIG}(\boldsymbol{\Omega}_t)$.

Proof. To evaluate the torsion on a given pair of vectors $\mathbf{u}_\mathbf{x}, \mathbf{v}_\mathbf{x} \in \mathbb{T}_\mathbf{x}\mathcal{S}$ we may extend them by distant parallel transport to a pair of vector fields $\mathbf{S}(\mathbf{u}_\mathbf{x}), \mathbf{S}(\mathbf{v}_\mathbf{x}) \in \mathrm{C}^1(\mathcal{S}\,;\mathbb{T}\mathcal{S})$ so that:

$$\mathrm{TORS}(\mathbf{u}_\mathbf{x}, \mathbf{v}_\mathbf{x}) := \nabla_{\mathbf{u}_\mathbf{x}} \mathbf{S}(\mathbf{v}_\mathbf{x}) - \nabla_{\mathbf{v}_\mathbf{x}} \mathbf{S}(\mathbf{u}_\mathbf{x}) - [\mathbf{S}(\mathbf{u}_\mathbf{x})\,, \mathbf{S}(\mathbf{v}_\mathbf{x})] = -[\mathbf{S}(\mathbf{u}_\mathbf{x})\,, \mathbf{S}(\mathbf{v}_\mathbf{x})]\,,$$

and the result follows from Theorem 3.6.2. ■

The *standard* bulk lagrangian per unit mass is: $L_t = K_t + P_t \circ \boldsymbol{\tau}_\mathcal{S} \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S}\,;\Re)$, where $K_t = \frac{1}{2}\,\mathbf{g} \circ \mathrm{DIAG} \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S}\,;\Re)$ is the positive definite quadratic form of the bulk *kinetic energy* per unit mass, with $\mathrm{DIAG}(\mathbf{v}) := (\mathbf{v}\,, \mathbf{v})$ so that $K_t(\mathbf{v}_t) = \frac{1}{2}\,\mathbf{g}\,(\mathbf{v}_t, \mathbf{v}_t)$, and of $P_t \in \mathrm{C}^1(\boldsymbol{\Omega}_t\,;\Re)$ is the bulk *load potential* per unit mass.

Lemma 3.6.3 *Let the ambient manifold $\{\mathcal{S}\,, \mathbf{g}\}$ be a riemannian manifold with the* LEVI-CIVITA *connection ∇ . Then the scalar fields $K_t \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S}\,;\Re)$ and*

$P_t \in \mathrm{C}^1(\mathbf{\Omega}_t\,;\Re)$ *fulfil the relations:*

$$\begin{cases} d_{\mathrm{F}} K_t = \mathbf{g}^\flat\,, \\ d_{\mathrm{B}} K_t = \tfrac{1}{2}\, d_{\mathrm{B}}(\mathbf{g} \circ \mathrm{DIAG}) = 0\,, \end{cases}$$

$$\begin{cases} d_{\mathrm{F}}(P_t \circ \boldsymbol{\tau}_{\mathcal{S}}) = 0\,, \\ d_{\mathrm{B}}(P_t \circ \boldsymbol{\tau}_{\mathcal{S}}) = TP_t \circ \boldsymbol{\tau}_{\mathcal{S}}\,. \end{cases}$$

Then, being $L_t := K_t + P_t \circ \boldsymbol{\tau}_{\mathcal{S}}$ *with*

$$K_t(\mathbf{v}_t) := \tfrac{1}{2}\,\langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_t \rangle\,,$$
$$E_t(\mathbf{v}_t) := \langle d_{\mathrm{F}} L_t(\mathbf{v}_t), \mathbf{v}_t \rangle - L_t(\mathbf{v}_t)\,,$$

we have the relation: $E_t = 2\,K_t - L_t = K_t - P_t \circ \boldsymbol{\tau}_{\mathcal{S}}\,.$

Proof. By definition of fiber and base derivative, for any $\mathbf{u}_t, \mathbf{v}_t, \delta\mathbf{v}_t \in \mathbb{T}_{\mathbf{\Omega}_t}\mathcal{S}$ with $\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{u}_t) = \boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t) = \boldsymbol{\tau}_{\mathcal{S}}(\delta\mathbf{v}_t)$ we have that:

$$\begin{aligned}
\langle d_{\mathrm{F}} K_t(\mathbf{u}_t), \mathbf{v}_t \rangle &= \partial_{\varepsilon=0}\, K_t(\mathbf{u}_t + \varepsilon\mathbf{v}_t) = \partial_{\varepsilon=0}\, \tfrac{1}{2}\mathbf{g}(\mathbf{u}_t + \varepsilon\mathbf{v}_t, \mathbf{u}_t + \varepsilon\mathbf{v}_t) \\
&= \mathbf{g}(\mathbf{u}_t, \mathbf{v}_t)\,, \\
\langle d_{\mathrm{B}} K_t(\mathbf{v}_t), \delta\mathbf{v}_t \rangle &= \partial_{\lambda=0}\, \boldsymbol{\varphi}_{\lambda,t}{\downarrow} K_t(\boldsymbol{\varphi}_{\lambda,t}{\Uparrow}\, \mathbf{v}_t) = \partial_{\lambda=0}\, K_t(\boldsymbol{\varphi}_{\lambda,t}{\Uparrow}\, \mathbf{v}_t) \circ \boldsymbol{\varphi}_{\lambda,t} = 0\,, \\
d_{\mathrm{F}}(P_t \circ \boldsymbol{\tau}_{\mathcal{S}})(\mathbf{v}_t) \cdot \delta\mathbf{v}_t &= TP_t(\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t)) \cdot T\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t) \cdot \nabla\mathbf{v}_t \cdot \delta\mathbf{v}_t = 0\,, \\
d_{\mathrm{B}}(P_t \circ \boldsymbol{\tau}_{\mathcal{S}})(\mathbf{v}_t) \cdot \delta\mathbf{v}_t &= TP_t(\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t)) \cdot T\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t) \cdot \mathbf{H}\mathbf{v}_t \cdot \delta\mathbf{v}_t \\
&= TP_t(\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t)) \cdot \delta\mathbf{v}_t\,,
\end{aligned}$$

where $\delta\mathbf{v}_t := \partial_{\lambda=0}\, \boldsymbol{\varphi}_{\lambda,t}\,.$

The second equality in the above list holds since the Levi-Civita parallel transport in $\{\mathcal{S}\,, \mathbf{g}\}$ preserves the metric, that is:

$$\mathbf{g}\,(\boldsymbol{\varphi}_{\lambda,t}{\Uparrow}\, \mathbf{v}_t, \boldsymbol{\varphi}_{\lambda,t}{\Uparrow}\, \mathbf{v}_t) \circ \boldsymbol{\varphi}_{\lambda,t} = \mathbf{g}\,(\mathbf{v}_t, \mathbf{v}_t)\,.$$

The last two equalities follow from the verticality of the covariant derivative and the fact that the horizontal lift is a right inverse to $T\boldsymbol{\tau}_{\mathcal{S}}$, the tangent map to the projection, so that $T\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t) \cdot \mathbf{H}\mathbf{v}_t = \mathbf{id}_{\,\mathbb{T}_{\mathbf{\Omega}_t}\mathcal{S}}\,.$ ∎

Corollary 3.6.1 (Euler's law of motion: special form) *Let the lagrangian per unit mass have the standard form:* $L_t = K_t + P_t \circ \boldsymbol{\tau}_{\mathcal{S}} \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\Omega}_t}\mathcal{S}\,;\Re)$ *and* ∇ *be the* LEVI-CIVITA *connection in the riemannian ambient manifold* $\{\mathcal{S}\,,\mathbf{g}\}$ *. Then the law of motion writes:*

$$\partial_{\tau=t} \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \mathbf{g}(\mathbf{v}_\tau, \gamma_{\tau,t}\Uparrow\delta\mathbf{v}_t)\,\mathbf{m}_\tau = \int_{\boldsymbol{\Omega}_t} \langle TP_t(\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t)), \delta\mathbf{v}_t\,\rangle\,\mathbf{m}_t + \langle\,\mathbf{F}^{\mathbb{C}}_t(\gamma_t), \delta\mathbf{v}^{\mathbb{C}}_t\,\rangle\,,$$

for any virtual velocity field $\delta\mathbf{v}_t \in \mathrm{CONF}(\boldsymbol{\Omega}_t)\cap\mathrm{RIG}(\boldsymbol{\Omega}_t)$ *.*

Proof. By the metric property of the LEVI-CIVITA connection: $\nabla\mathbf{g} = 0$ and the mass-preserving ansatz on the virtual velocities: $\mathcal{L}_{\delta\mathbf{v}_t}\mathbf{m}_t = 0$, we have:

$$\partial_{\lambda=0} \int_{\boldsymbol{\varphi}_{\lambda,t}(\boldsymbol{\Omega}_t)} K_t(\boldsymbol{\varphi}_{\lambda,t}\Uparrow\mathbf{v}_t)\,\mathbf{m}_t = \int_{\boldsymbol{\Omega}_t} \partial_{\lambda=0}\,\boldsymbol{\varphi}_{\lambda,t}\!\downarrow\![K_t(\boldsymbol{\varphi}_{\lambda,t}\Uparrow\mathbf{v}_t)\,\mathbf{m}_t]$$

$$= \int_{\boldsymbol{\Omega}_t} \partial_{\lambda=0}\,\boldsymbol{\varphi}_{\lambda,t}\!\downarrow\! K_t(\boldsymbol{\varphi}_{\lambda,t}\Uparrow\mathbf{v}_t)\,\mathbf{m}_t + \int_{\boldsymbol{\Omega}_t} K_t(\mathbf{v}_t)\,\mathcal{L}_{\delta\mathbf{v}_t}\mathbf{m}_t = 0\,,$$

$$\partial_{\lambda=0} \int_{\boldsymbol{\varphi}_{\lambda,t}(\boldsymbol{\Omega}_t)} P_t(\boldsymbol{\tau}_{\mathcal{S}}(\boldsymbol{\varphi}_{\lambda,t}\Uparrow\mathbf{v}_t))\,\mathbf{m}_t = \int_{\boldsymbol{\Omega}_t} \partial_{\lambda=0}\,\boldsymbol{\varphi}_{\lambda,t}\!\downarrow\![P_t(\boldsymbol{\tau}_{\mathcal{S}}(\boldsymbol{\varphi}_{\lambda,t}\Uparrow\mathbf{v}_t))\,\mathbf{m}_t]$$

$$= \int_{\boldsymbol{\Omega}_t} \langle\,d_{\mathrm{B}}P_t(\mathbf{v}_t), \delta\mathbf{v}_t\,\rangle\,\mathbf{m}_t + \int_{\boldsymbol{\Omega}_t} P_t(\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t))\,\mathcal{L}_{\delta\mathbf{v}_t}\mathbf{m}_t\,,$$

with the last integral vanishing. By Lemma 3.6.3 we have that $\langle\,d_{\mathrm{B}}P_t(\mathbf{v}_t), \delta\mathbf{v}_t\,\rangle = \langle TP_t(\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t)), \delta\mathbf{v}_t\rangle$ and $\langle\,d_{\mathrm{F}}L_\tau(\mathbf{v}_\tau), \gamma_{\tau,t}\Uparrow\delta\mathbf{v}_t\,\rangle = \mathbf{g}(\mathbf{v}_\tau, \gamma_{\tau,t}\Uparrow\delta\mathbf{v}_t)$. Moreover the LEVI-CIVITA connection is torsion-free, that is $\mathrm{TORS}(\mathbf{v}_t, \delta\mathbf{v}_t) = 0$, and the result follows from Theorem 3.6.2. ■

In the euclidean ambient space, a simple body is defined by the property that conforming isometric virtual displacement fields are simple infinitesimal isometries, expressible as the sum of a *speed of translation* and of an *angular velocity* around a pole. Then we recover the classical EULER's laws for the time-rate of variation of momentum and of moment of momentum.

Corollary 3.6.2 (d'Alembert's law of motion) *By conservation of mass the special* EULER*'s law of motion translates into* D'ALEMBERT*'s law:*

$$\int_{\boldsymbol{\Omega}_t} \mathbf{g}\,(\nabla_{\mathbf{v}_t}\mathbf{v}_t, \delta\mathbf{v}_t)\,\mathbf{m}_t = \int_{\boldsymbol{\Omega}_t} \langle TP_t(\boldsymbol{\tau}_{\mathcal{S}}(\mathbf{v}_t)), \delta\mathbf{v}_t\,\rangle\,\mathbf{m}_t + \langle\,\mathbf{F}^{\mathbb{C}}_t(\gamma_t), \delta\mathbf{v}^{\mathbb{C}}_t\,\rangle\,,$$

for any virtual velocity field $\delta\mathbf{v}_t \in \mathrm{CONF}(\boldsymbol{\Omega}_t)\cap\mathrm{RIG}(\boldsymbol{\Omega}_t)$ *.*

Proof. Applying the transport formula and LEIBNIZ rule we get the identity:

$$\begin{aligned}
\partial_{\tau=t} \int_{\gamma_{\tau,t}(\boldsymbol{\Omega}_t)} \mathbf{g}(\mathbf{v}_\tau, \gamma_{\tau,t}\Uparrow \delta\mathbf{v}_t)\, \mathbf{m}_\tau &= \int_{\boldsymbol{\Omega}_t} \partial_{\tau=t}\, \gamma_{\tau,t}\!\downarrow\![\mathbf{g}(\mathbf{v}_\tau, \gamma_{\tau,t}\Uparrow \delta\mathbf{v}_t)\, \mathbf{m}_\tau] \\
&= \int_{\boldsymbol{\Omega}_t} [\partial_{\tau=t}\, \mathbf{g}(\mathbf{v}_\tau, \gamma_{\tau,t}\Uparrow \delta\mathbf{v}_t) \circ \gamma_{\tau,t}]\, \mathbf{m}_t + \int_{\boldsymbol{\Omega}_t} (\mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \partial_{\tau=t}\, \gamma_{\tau,t}\!\downarrow\!\mathbf{m}_\tau \\
&= \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\partial_{\tau=t}\, \gamma_{\tau,t}\Downarrow \mathbf{v}_\tau, \delta\mathbf{v}_t)\, \mathbf{m}_t + \int_{\boldsymbol{\Omega}_t} (\mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathcal{L}_{t,\mathbf{v}_t}\, \mathbf{m} \\
&= \int_{\boldsymbol{\Omega}_t} \mathbf{g}(\nabla_{\mathbf{v}_t}\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathbf{m}_t + \int_{\boldsymbol{\Omega}_t} (\mathbf{g}(\mathbf{v}_t, \delta\mathbf{v}_t)\, \mathcal{L}_{t,\mathbf{v}_t}\, \mathbf{m}\,,
\end{aligned}$$

where $\mathbf{g}(\mathbf{v}_\tau, \gamma_{\tau,t}\Uparrow \delta\mathbf{v}_t) \circ \gamma_{\tau,t} = \mathbf{g}(\gamma_{\tau,t}\Downarrow \mathbf{v}_\tau, \delta\mathbf{v}_t)$ since LEVI-CIVITA connection is metric. Imposing conservation of mass: $\mathcal{L}_{t,\mathbf{v}_t}\mathbf{m} := \partial_{\tau=t}\, \mathbf{m}_\tau + \mathcal{L}_{\mathbf{v}_t}\mathbf{m}_t = 0$, the result follows from Corollary 3.6.1. ∎

3.6.4 Boundary value problems

The basic tool in boundary value problems governed by a linear partial differential operator DIFF of order n, is GREEN's formula of integration by parts, which formally may be written as:

$$\begin{aligned}
\int_{\mathrm{PAT}(\boldsymbol{\Omega}_t)} \langle \bullet, \mathrm{DIFF} \circ \rangle \boldsymbol{\mu} &= \int_{\mathrm{PAT}(\boldsymbol{\Omega}_t)} \langle \mathrm{ADJDIFF} \bullet, \circ \rangle \boldsymbol{\mu} \\
&\quad + \oint_{\partial\mathrm{PAT}(\boldsymbol{\Omega}_t)} \langle \mathrm{FLUX} \bullet, \mathrm{VAL} \circ \rangle \partial\boldsymbol{\mu}\,,
\end{aligned}$$

where $\boldsymbol{\Omega}_t$ is a submanifold of a finite dimensional riemannian space $\{\mathcal{S}, \mathbf{g}\}$, $\mathrm{PAT}(\boldsymbol{\Omega}_t)$ is a fixed patchwork, $\partial\mathrm{PAT}(\boldsymbol{\Omega}_t)$ is its boundary, $\partial\boldsymbol{\mu}$ is the volume form induced on the surfaces $\partial\mathrm{PAT}(\boldsymbol{\Omega}_t)$ by the volume form in $\mathcal{S}$ and all the integrals are assumed to take a finite value. The differential operator ADJDIFF of order n is the *formal adjoint* of DIFF . The boundary integral acts on the duality pairing between the two fields FLUX $\bullet$ and VAL $\circ$ with the differential operators FLUX and VAL being n-tuples of normal derivatives of order from 0 to $n-1$ in inverse sequence, so that the duality pairing is the sum of n terms, whose k-th term is the pairing of normal derivatives of two fields respectively of order k and $n-1-k$.

Boundary value problems are characterized by the property that the closed linear subspace $\textsc{Conf}(\mathbf{\Omega}_t)$ of conforming test fields includes the whole linear subspace $\ker(\textsc{Val})$ of test fields in $\textsc{Kin}(\mathbf{\Omega}_t)$ with vanishing boundary values on $\partial\textsc{Pat}(\mathbf{\Omega}_t)$, i.e.

$$\ker(\textsc{Val}) \subseteq \textsc{Conf}(\mathbf{\Omega}_t)\,.$$

Let us assume that the force virtual power $\langle \mathbf{F}_t, \delta\mathbf{v}_t \rangle$ is expressed in terms of forces per unit volume $\mathbf{b} \in \textsc{Sqiv}(\mathbf{\Omega}_t)$ ($\textsc{Sqiv}$:= square integrable vector fields) and of forces per unit area (tractions) $\mathbf{t} \in \textsc{Sqiv}(\partial\textsc{Pat}(\mathbf{\Omega}_t))$, so that the force virtual power is given by:

$$\langle \mathbf{F}^{\mathbb{C}}_t(\gamma_t), \delta\mathbf{v}^{\mathbb{C}}_t \rangle := \int_{\mathbf{\Omega}_t} \mathbf{g}\,(\mathbf{b}_t, \delta\mathbf{v}_t)\,\boldsymbol{\mu} + \int_{\partial\textsc{Pat}(\mathbf{\Omega}_t)} \mathbf{g}\,(\mathbf{t}_t, \delta\mathbf{v}_t)\,\partial\boldsymbol{\mu}\,.$$

d'Alembert's law, may then be rewritten as

$$\int_{\mathbf{\Omega}_t} \mathbf{g}\,(\nabla_{\mathbf{v}_t}\mathbf{v}_t, \delta\mathbf{v}_t)\,\mathbf{m}_t + \int_{\textsc{Pat}(\mathbf{\Omega}_t)} \langle \mathbf{T}_t, \mathrm{sym}\,\nabla\delta\mathbf{v}_t \rangle_{\mathbf{g}}\,\boldsymbol{\mu}$$

$$= \int_{\mathbf{\Omega}_t} \mathbf{g}\,(\mathbf{b}_t, \delta\mathbf{v}_t)\,\boldsymbol{\mu} + \int_{\partial\textsc{Pat}(\mathbf{\Omega}_t)} \mathbf{g}\,(\mathbf{t}_t, \delta\mathbf{v}_t)\,\partial\boldsymbol{\mu}\,,$$

and a standard localization procedure, leads to the differential equation:

$$-\textsc{Div}\,\mathbf{T}_t = \mathbf{b}_t - \rho_t \cdot \mathbf{g}^\flat \circ \nabla_{\mathbf{v}_t}\mathbf{v}_t\,, \quad \text{in} \quad \textsc{Pat}_\infty(\mathbf{\Omega}_t)\,,$$

and the boundary conditions on the jump $[[\mathbf{T}_t\mathbf{n}]]$ across the boundary of the domain $\mathbf{\Omega}_t$ and across the interfaces of the patchwork $\textsc{Pat}_\infty(\mathbf{\Omega}_t)$ fulfills the conditions:

$$\begin{aligned} \mathbf{T}_t\mathbf{n} &\in \mathbf{t} + \textsc{Conf}^\circ\,, && \text{on} \quad \mathbf{\Omega}_t \\ [[\mathbf{T}_t\mathbf{n}]] &\in \mathbf{t}^+ + \mathbf{t}^- + \textsc{Conf}^\circ\,, && \text{on} \quad \textsc{Sing}(\textsc{Pat}_\infty(\mathbf{\Omega}_t)) \end{aligned}$$

where the fields $\mathbf{t}$ of surfacial forces are taken to be zero outside their domain of definition and $\textsc{Pat}_\infty$ denotes a patchwork sufficiently fine for the statement at hand.

Chapter 4

Elasticity

The theory of elasticity is a fundamental chapter of Mathematical Physics which leads to results that, under suitable generalizations, can be applied to the analysis of other constitutive models, describing different physical phenomena, but sharing in the meanwhile the same formal properties.

This chapter is devoted to an abstract presentation of the characteristic properties of an elastic behaviour, with a generalized formulation encompassing constitutive models governed by monotone conservative multivalued relations which cannot be dealt with by the classical theory.

Constrained elasticity, such as for incompressible materials, is dealt with by assuming that at each point admissible strains belong to a differentiable manifold. The elastic constitutive law is defined in the general case and specialized to linear strain spaces and linear elasticity. The issue of linearization of general nonlinear laws is also briefly investigated.

The treatment of a general monotone and conservative elastic behaviour is based on the presentation of the theory provided in [164].

The specialization of this general model to the classical one-to-one and possibily linearly elastic behaviour shows that well-known results can be recovered as special cases of the ones established in the new more comprehensive framework.

The theory is applied to the modelling of several widely adopted constitutive laws which can be framed into the general scheme of monotone laws governed by convex potentials.

At the end of the chapter the theory of associated plasticity and viscoplasticity is revisited in the unitary framework provided by the generalized elastic model.

4.1 Elastic behaviour

A fundamental assumption for the development of the theory of elasticity is the existence of a relation between dual vector quantities, representing the kinematic and static state variables, which depend only upon their actual values and not on their past history, so that an elastic material is a *material without memory*.

Moreover an elastic relation enjoys the properties of being invertible and conservative, and hence both the direct and the inverse constitutive laws admit a potential.

Starting from the classical scheme of a one-to-one linear relation between stress and strain, it is possible to develop a general scheme which includes a much wider class of constitutive relations involving either values and rates of the kinematic and static state variables encompassing most of the engineering models of material behaviour.

This general model is called *generalized elasticity* to recall that its genesis consists in a suitable extension of the classical linear elastic relation.

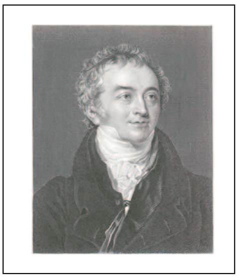

Figure 4.1: Thomas Young (1773 - 1829)

4.2 Constrained elastic law

Let us consider a reference placement $\mathbb{B}$ and the actual placement $\mathbf{\Omega} = \boldsymbol{\chi}(\mathbb{B})$ of the body in the euclidean space $\{\mathbb{S}, \mathbf{g}\}$.

In a general setting, admissible strain fields in $\mathbb{B}$ are described by symmetric tensor fields taking their point values at a particle $\mathbf{p} \in \mathbb{B}$ in a nonlinear finite dimensional manifold $\mathbf{D}$, called the *admissible strain manifold* at $\mathbf{p} \in \mathbb{B}$.

A standard example of a nonlinear manifold of admissible strains is provided by the assumption of the incompressibility constraint (isochoric replacements).

The admissible strain manifold is then the **unimodular group** of symmetric tensors $\boldsymbol{\varepsilon} \in BL(\mathbb{T}_{\mathbf{p}}\mathbb{B}^2 ; \Re)$ such that

$$\det \mathrm{GRAM}(\boldsymbol{\varepsilon}) = 1\,,$$

with $\mathrm{GRAM}_{ij}(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}(\mathbf{d}_i, \mathbf{d}_j)$ and $\{\mathbf{d}_i\,, i = 1, 2, 3\}$ orthonormal basis in $\mathbb{T}_{\mathbf{p}}\mathbb{B}$.

Definition 4.2.1 *An **elastic law** on a nonlinear **admissible strain manifold** $\boldsymbol{D}$ is a vector bundle homomorphism $\mathcal{E} \in \mathrm{C}^1(\mathbb{T}\boldsymbol{D}; \mathbb{T}\Re)$, that is a fiber preserving, fiber linear and differentiable map from the tangent bundle $\boldsymbol{\pi_D} \in \mathrm{C}^1(\mathbb{T}\boldsymbol{D}; \boldsymbol{D})$ to the tangent bundle $\boldsymbol{\pi}_{\Re} \in \mathrm{C}^1(\mathbb{T}\Re; \Re)$. An elastic law admits an **elastic potential** if there exists a map $\varphi_{\mathcal{E}} \in \mathrm{C}^1(\boldsymbol{D}; \Re)$ such that*

$$T\varphi_{\mathcal{E}} = \mathcal{E}\,.$$

Theorem 4.2.1 (Elastic law) *An elastic law is equivalently described by a cross section of the cotangent bundle $\mathbb{T}^*\boldsymbol{D}$, that is a differential one-form $\mathcal{E} \in \mathrm{C}^1(\boldsymbol{D}; \mathbb{T}^*\boldsymbol{D})$ with $\boldsymbol{\pi}^*_{\boldsymbol{D}} \circ \mathcal{E} = \mathbf{id}_{\boldsymbol{D}}$.*

Proof. The vector bundle homomorphism $\mathcal{E} \in \mathrm{C}^1(\mathbb{T}\mathbf{D}\,; \mathbb{T}\Re)$ is fiber preserving and hence defines a base morphism $\varphi_{\mathcal{E}} \in \mathrm{C}^1(\mathbf{D}\,; \Re)$ by the commutative diagram:

$$\begin{array}{ccc} \mathbb{T}\mathbf{D} & \xrightarrow{\ \mathcal{E}\ } & \mathbb{T}\Re \\ {\scriptstyle \boldsymbol{\pi}_{\mathbf{D}}}\big\downarrow & & \big\downarrow{\scriptstyle \boldsymbol{\pi}_{\Re}} \\ \mathbf{D} & \xrightarrow{\ \varphi_{\mathcal{E}}\ } & \Re \end{array} \iff \quad \varphi_{\mathcal{E}} \circ \boldsymbol{\pi}_{\mathbf{D}} = \boldsymbol{\pi}_{\Re} \circ \mathcal{E} \in \mathrm{C}^0(\mathbb{T}\mathbf{D}\,; \Re)\,.$$

The elastic law $\mathcal{E} \in \mathrm{C}^1(\mathbb{T}\mathbf{D}\,; \mathbb{T}\Re)$ may then be regarded as a field which associates with any strain $\boldsymbol{\varepsilon} \in \mathbf{D}$ a linear map $\mathcal{E}(\boldsymbol{\varepsilon}) \in BL(\mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}\,; \mathbb{T}_{\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon})}\Re)$. By the isomorphism between $\mathbb{T}_{\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon})}\Re$ and $\{\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon})\} \times \Re$ and the identification between $\{\alpha_0\} \times \Re$ and $\Re$ made by setting $\{\alpha_0\,, \alpha\} \simeq \{0\,, \alpha\} \simeq \alpha$ for all $\alpha \in \Re$, we may assume that $\mathcal{E}(\boldsymbol{\varepsilon}) \in BL(\mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}\,; \Re) = \mathbb{T}^*_{\boldsymbol{\varepsilon}}\mathbf{D}$ and the vector bundle homomorphism may be considered as a differential one-form on the admissible strain manifold $\mathbf{D}$, i.e. $\mathcal{E} \in \mathrm{C}^1(\mathbf{D}\,; \mathbb{T}^*\mathbf{D})$ with $\boldsymbol{\pi}^*_{\mathbf{D}} \circ \mathcal{E} = \mathbf{id}_{\mathbf{D}}$. ■

The covectors in $\mathbb{T}^*_{\boldsymbol{\varepsilon}}\mathbf{D}$ are the *effective stresses* at $\boldsymbol{\varepsilon} \in \mathbf{D}$ and the covector $\mathcal{E}(\boldsymbol{\varepsilon}) \in \mathbb{T}^*_{\boldsymbol{\varepsilon}}\mathbf{D}$ is the stress field elastically associated with the strain $\boldsymbol{\varepsilon} \in \mathbf{D}$. The elements of the linear tangent space $\mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}$ are called *admissible tangent strains* at $\boldsymbol{\varepsilon} \in \mathbf{D}$. If the manifold $\mathbf{D}$ of admissible strains is endowed with a metric field, there is an isomorphism $\mathbf{g}_{\mathbf{D}} \in BL(\mathbb{T}\mathbf{D}\,; \mathbb{T}^*\mathbf{D})$ and each effective stress $\boldsymbol{\sigma} \in \mathbb{T}^*\mathbf{D}$ can be represented by a tangent strain-like vector $\mathbf{g}^{-1}_{\mathbf{D}} \circ \boldsymbol{\sigma} \in \mathbb{T}\mathbf{D}$.

4.2.1 Elastic potential

The specific work performed at the particle $\mathbf{p} \in \mathbb{B}$ by the elastic stress field $\mathcal{E} \circ \gamma \in \mathrm{C}^1(I\,; \mathbb{T}^*\mathbf{D})$ along a loop $\gamma \in \mathrm{C}^1(I\,; \mathbf{D})$ in the *admissible strain* manifold, is provided by the *circuitation* integral

$$\oint_\gamma \mathcal{E} = \int_a^b \mathcal{E}(\gamma(\lambda)) \cdot \partial_{\mu=\lambda}\, \gamma(\mu)\, d\lambda\,,$$

with $\gamma(a) = \gamma(b)$ and $\partial_{\mu=\lambda}\, \gamma(\mu) \in \mathbb{T}_{\gamma(\lambda)}\gamma \subset \mathbb{T}_{\gamma(\lambda)}\mathbf{D}$.

Let us now assume that the 1-D Betti's number of the admissible strain manifold $\mathbf{D}$ vanishes, i.e. that any loop in $\mathbf{D}$ is the boundary of a 2-D submanifold $\Sigma \subset \mathbf{D}$.

Then we may put $\gamma = \partial\Sigma$, and the circuitation of the one-form $\mathcal{E}$ along any loop vanishes if and only if it is a closed form on $\mathbb{D}$, since this means that its exterior derivative vanishes $d\mathcal{E} = 0$ and then, by Stokes formula:

$$\oint_\gamma \mathcal{E} = \oint_{\partial\Sigma} \mathcal{E} = \int_\Sigma d\mathcal{E} = 0\,.$$

The exterior derivative $d\mathcal{E}(\boldsymbol{\varepsilon})$ at $\boldsymbol{\varepsilon} \in \mathbf{D}$ is evaluated by the formula

$$d\mathcal{E}(\boldsymbol{\varepsilon}) \cdot \mathbf{X}(\boldsymbol{\varepsilon}) \cdot \mathbf{Y}(\boldsymbol{\varepsilon}) = d_{\mathbf{X}(\boldsymbol{\varepsilon})}(\mathcal{E} \cdot \mathbf{Y}) - d_{\mathbf{Y}(\boldsymbol{\varepsilon})}(\mathcal{E} \cdot \mathbf{X}) - \mathcal{E}(\boldsymbol{\varepsilon}) \cdot [\mathbf{X}\,, \mathbf{Y}](\boldsymbol{\varepsilon})\,,$$

where $\mathbf{X}, \mathbf{Y} \in \mathrm{C}^1(\mathbf{D}\,; \mathbb{T}\mathbf{D})$ are vector fields of admissible tangent strains.

Being tensorial, the exterior derivative $d\mathcal{E}(\boldsymbol{\varepsilon})$ depends only on the point values $\mathbf{X}(\boldsymbol{\varepsilon}), \mathbf{Y}(\boldsymbol{\varepsilon}) \in \mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}$. However, none of the terms at the r.h.s. of the defining equality is tensorial.

We recall that the differential of the functional $\mathcal{E} \cdot \mathbf{Y} \in \mathrm{C}^1(\mathbf{D}\,; \Re)$ at the point $\boldsymbol{\varepsilon} \in \mathbb{D}$ is the linear map $T_{\boldsymbol{\varepsilon}}(\mathcal{E} \cdot \mathbf{Y}) \in BL(\mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}\,; \Re)$ such that for all vectors $\mathbf{X}(\boldsymbol{\varepsilon}) \in \mathbb{T}_{\boldsymbol{\varepsilon}}\mathbb{D}$:

$$T_{\boldsymbol{\varepsilon}}(\mathcal{E} \cdot \mathbf{Y}) \cdot \mathbf{X}(\boldsymbol{\varepsilon}) := T_{\mathbf{X}(\boldsymbol{\varepsilon})}(\mathcal{E} \cdot \mathbf{Y}) = d_{\mathbf{X}(\boldsymbol{\varepsilon})}(\mathcal{E} \cdot \mathbf{Y})\,.$$

In a local chart $\varphi \in \mathrm{C}^1(\mathbf{D}\,; E)$, in terms of partial derivatives:

$$d(\varphi\!\uparrow\!\mathcal{E})(\varepsilon) \cdot X(\varepsilon) \cdot Y(\varepsilon) = d_{X(\varepsilon)}(\varphi\!\uparrow\!\mathcal{E} \cdot Y)(\varepsilon) - d_{Y(\varepsilon)}(\varphi\!\uparrow\!\mathcal{E} \cdot X)(\varepsilon)\,,$$

where $\varepsilon \in E$ and $X, Y \in \mathrm{C}^1(E\,; E)$ are constant fields, that is they are parallel transported according to the distant connection by translation in the linear space E. Then $[X\,, Y] = 0$ since the corresponding flows commute.

Let us further assume that the admissible strain manifold is star shaped, i.e. that it can be homotopically contracted to a point. Then the closedness of the one-form $\mathcal{E} \in C^1(\mathbf{D}\,;\mathbb{T}^*\mathbf{D})$ implies, by Poincaré Lemma 1.6.1 on page 155, that there exists a scalar potential $\varphi_{\mathcal{E}} \in C^1(\mathbf{D}\,;\Re)$ such that

$$\mathcal{E} = d\varphi_{\mathcal{E}} = T\varphi_{\mathcal{E}}\,,$$

since the exterior differential reduced to the differential, for scalar potentials.

The *elastic potential* $\varphi_{\mathcal{E}} \in C^1(\mathbf{D}\,;\Re)$ is given by

$$\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon}) - \varphi_{\mathcal{E}}(\boldsymbol{\varepsilon}_0) = \oint_{\partial\gamma} \mathcal{E} = \int_{\gamma} d\mathcal{E}\,,$$

where γ is any 1-D chain with boundary $\partial\gamma = \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_0$.

4.2.2 Unconstrained elasticity

If the admissible strain manifold $\mathbf{D}$ is a linear space, denoting by $\mathbf{D}^*$ the dual space, the tangent bundle $\mathbb{T}\mathbf{D}$ is isomorphic to the cartesian product $\mathbf{D}\times\mathbf{D}$ and may be identified with the linear space $\mathbf{D}$ by setting $\{\boldsymbol{\varepsilon}_0\,,\boldsymbol{\varepsilon}\} \simeq \{0\,,\boldsymbol{\varepsilon}\} \simeq \boldsymbol{\varepsilon}$ for all $\boldsymbol{\varepsilon} \in \mathbf{D}$. The cotangent bundle $\mathbb{T}^*\mathbf{D}$ is isomorphic to the cartesian product $\mathbf{D}\times\mathbf{D}^*$ and may be identified with the dual space $\mathbf{D}^*$ by setting $\{\boldsymbol{\varepsilon}_0\,,\boldsymbol{\sigma}\} \simeq \{0\,,\boldsymbol{\sigma}\} \simeq \boldsymbol{\sigma}$ for all $\boldsymbol{\sigma} \in \mathbf{D}^*$.

An *elastic law* on a linear strain manifold $\mathbf{D}$ is then a map $\mathcal{E} \in C^1(\mathbf{D}\,;\mathbf{D}^*)$ and the exterior derivative $d\mathcal{E}(\boldsymbol{\varepsilon})$ may be expressed as

$$d\mathcal{E}(\boldsymbol{\varepsilon})\cdot\mathbf{X}(\boldsymbol{\varepsilon})\cdot\mathbf{Y}(\boldsymbol{\varepsilon}) = d_{\mathbf{X}(\boldsymbol{\varepsilon})}(\mathcal{E}\cdot\mathbf{Y})(\boldsymbol{\varepsilon}) - d_{\mathbf{Y}(\boldsymbol{\varepsilon})}(\mathcal{E}\cdot\mathbf{X})(\boldsymbol{\varepsilon})\,,$$

where $\boldsymbol{\varepsilon} \in \mathbf{D}$ and $\mathbf{X},\mathbf{Y} \in C^1(\mathbf{D}\,;\mathbf{D})$ are constant fields in the linear space $\mathbf{D}$. Then, by evaluating the differential $T_{\boldsymbol{\varepsilon}}(\mathcal{E}\cdot\mathbf{X}) \in BL(\mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}\,;\Re)$ of the functional $\mathcal{E}\cdot\mathbf{X} \in C^1(\mathbf{D}\,;\Re)$ at the point $\boldsymbol{\varepsilon} \in \mathbb{D}$, by the constancy of $\mathbf{X} \in C^1(\mathbf{D}\,;\mathbf{D})$, we get that

$$T_{\boldsymbol{\varepsilon}}(\mathcal{E}\cdot\mathbf{X}) = T_{\boldsymbol{\varepsilon}}\mathcal{E}\cdot\mathbf{X}(\boldsymbol{\varepsilon})\,.$$

and the integrability condition may be written

$$\begin{aligned} d\mathcal{E}(\boldsymbol{\varepsilon})\cdot\mathbf{X}(\boldsymbol{\varepsilon})\cdot\mathbf{Y}(\boldsymbol{\varepsilon}) &= d_{\mathbf{X}(\boldsymbol{\varepsilon})}(\mathcal{E}\cdot\mathbf{Y})(\boldsymbol{\varepsilon}) - d_{\mathbf{Y}(\boldsymbol{\varepsilon})}(\mathcal{E}\cdot\mathbf{X})(\boldsymbol{\varepsilon}) \\ &= T_{\boldsymbol{\varepsilon}}\mathcal{E}\cdot\mathbf{Y}(\boldsymbol{\varepsilon})\cdot\mathbf{X}(\boldsymbol{\varepsilon}) - T_{\boldsymbol{\varepsilon}}\mathcal{E}\cdot\mathbf{X}(\boldsymbol{\varepsilon})\cdot\mathbf{Y}(\boldsymbol{\varepsilon}) = 0\,, \end{aligned}$$

If moreover the elastic map is linear, that is $\mathcal{E} \in BL(\mathbf{D}\,;\mathbf{D}^*)$, we have that

$$T_{\boldsymbol{\varepsilon}}\mathcal{E}\cdot\mathbf{X}(\boldsymbol{\varepsilon}) = \lim_{\lambda\to 0}\frac{1}{\lambda}(\mathcal{E}(\boldsymbol{\varepsilon}+\lambda\mathbf{X}(\boldsymbol{\varepsilon})) - \mathcal{E}(\boldsymbol{\varepsilon})) = \mathcal{E}\cdot\mathbf{X}(\boldsymbol{\varepsilon})\,,$$

that is

$$T_{\boldsymbol{\varepsilon}}\mathcal{E} = \mathcal{E}\,,$$

and the integrability condition writes

$$(d\mathcal{E} \cdot \mathbf{X} \cdot \mathbf{Y})(\boldsymbol{\varepsilon}) = \mathcal{E} \cdot \mathbf{X}(\boldsymbol{\varepsilon}) \cdot \mathbf{Y}(\boldsymbol{\varepsilon}) - \mathcal{E} \cdot \mathbf{Y}(\boldsymbol{\varepsilon}) \cdot \mathbf{X}(\boldsymbol{\varepsilon}) = 0\,.$$

Then, by considering the elastic map $\mathcal{E} \in BL(\mathbf{D}\,;\mathbf{D}^*)$ as a twice covariant tensor $\mathcal{E} \in BL(\mathbf{D}^2\,;\Re)$, the integrability condition amounts to require that the elastic tensor be symmetric.

4.2.3 Linearized elasticity

The linearization of an elastic law $\mathcal{E} \in \mathrm{C}^1(\mathbf{D}\,;\mathbb{T}^*\mathbf{D})$ cannot be performed by the associated tangent map $T_{\boldsymbol{\varepsilon}}\mathcal{E} \in \mathrm{C}^1(\mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}\,;\mathbb{T}_{\mathcal{E}(\boldsymbol{\varepsilon})}\mathbb{T}^*\mathbf{D})$, Indeed it does not transform a tangent vector $\delta\boldsymbol{\varepsilon} \in \mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}$ at a strain point $\boldsymbol{\varepsilon} \in \mathbf{D}$ into a stress form but rather into a tangent vector $T_{\boldsymbol{\varepsilon}}\mathcal{E} \cdot \delta\boldsymbol{\varepsilon} \in \mathbb{T}_{\mathcal{E}(\boldsymbol{\varepsilon})}\mathbb{T}^*\mathbf{D}$ at the stress form $\mathcal{E}(\boldsymbol{\varepsilon}) \in \mathbb{T}^*_{\boldsymbol{\varepsilon}}\mathbf{D}$. Then the natural candidate for linearization is the following.

Definition 4.2.2 *The* ***incremental form*** *at $\boldsymbol{\varepsilon} \in \boldsymbol{D}$ of the nonlinear elastic law $\mathcal{E} \in \mathrm{C}^1(\boldsymbol{D};\mathbb{T}^*\boldsymbol{D})$ according to the connection ∇ on $\boldsymbol{D}$ is the map*

$$\nabla_{\boldsymbol{\varepsilon}}\mathcal{E} \in BL(\mathbb{T}_{\boldsymbol{\varepsilon}}\boldsymbol{D};\mathbb{T}^*_{\boldsymbol{\varepsilon}}\boldsymbol{D})\,,$$

which associates, with any tangent strain $\delta\boldsymbol{\varepsilon} \in \mathbb{T}_{\boldsymbol{\varepsilon}}\boldsymbol{D}$, the corresponding ***tangent stress*** *$\delta\boldsymbol{\sigma}(\boldsymbol{\varepsilon}) \in \mathbb{V}_{\mathcal{E}(\boldsymbol{\varepsilon})}\mathbb{T}^*\boldsymbol{D} \simeq \mathbb{T}^*_{\boldsymbol{\varepsilon}}\boldsymbol{D}$ which is the* ***covariant derivative*** *of the elastic law $\mathcal{E} \in \mathrm{C}^1(\boldsymbol{D};\mathbb{T}^*\boldsymbol{D})$ along the tangent strain $\delta\boldsymbol{\varepsilon} \in \mathbb{T}_{\boldsymbol{\varepsilon}}\boldsymbol{D}$:*

$$\delta\boldsymbol{\sigma}(\boldsymbol{\varepsilon}) = \nabla\mathcal{E} \cdot \delta\boldsymbol{\varepsilon}\,.$$

Let us recall that the covariant derivative of a field of one-forms $\mathcal{E} \in \mathrm{C}^1(\mathbf{D}\,;\mathbb{T}^*\mathbf{D})$ at $\boldsymbol{\varepsilon} \in \mathbf{D}$ is defined by a formal application of LEIBNIZ rule:

$$\langle \nabla_{\delta_1\boldsymbol{\varepsilon}}\mathcal{E}, \delta_2\boldsymbol{\varepsilon} \rangle = \nabla_{\delta_1\boldsymbol{\varepsilon}}\langle \mathcal{E}, \delta_2 \rangle + \langle \mathcal{E}(\boldsymbol{\varepsilon}), \nabla_{\delta_1\boldsymbol{\varepsilon}}\delta_2 \rangle\,, \quad \delta_1\boldsymbol{\varepsilon}, \delta_2\boldsymbol{\varepsilon} \in \mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}\,.$$

According to the given definition, the linearization of the nonlinear elastic law $\mathcal{E} \in \mathrm{C}^1(\mathbf{D}\,;\mathbb{T}^*\mathbf{D})$ at a strain $\boldsymbol{\varepsilon} \in \mathbf{D}$ depends on the chosen connection. If the manifold $\mathbf{D}$ is included by the map $\mathbf{i} \in \mathrm{C}^1(\mathbf{D}\,;\mathbb{M})$ into a larger manifold $\mathbb{M}$ endowed with a connection $\nabla_{\mathbb{M}}$, it is natural to assume in $\mathbf{D}$ the induced connection, given by

$$\nabla_{\mathbf{D}} := \mathbf{i}{\downarrow}\nabla_{\mathbb{M}}\,,$$

or explicitly: $\nabla_{\mathbf{D}}\delta_1 \cdot \delta_2\boldsymbol{\varepsilon} = \nabla_{\mathbb{M}}(\mathbf{i}{\uparrow}\delta_1) \cdot \mathbf{i}{\uparrow}\delta_2\boldsymbol{\varepsilon}\,.$

In a riemannian manifold $\{\mathbb{M}, \mathbf{g}_{\mathbb{M}}\}$ the induced metric on $\mathbf{D}$ is $\mathbf{g}_{\mathbf{D}} = \mathbf{i}\!\downarrow\!\mathbf{g}_{\mathbb{M}}$ and the LEVI-CIVITA connections in $\{\mathbb{M}, \mathbf{g}_{\mathbb{M}}\}$ and $\{\mathbf{D}, \mathbf{g}_{\mathbf{D}}\}$ are related by $\nabla_{\mathbf{D}} = \mathbf{P}_{\mathbb{M}} \circ \nabla_{\mathbb{M}}$ or, explicitly:

$$\nabla_{\mathbf{D}}\delta_1 \cdot \delta_2\boldsymbol{\varepsilon} = \mathbf{P}_{\mathbb{M}}(\nabla_{\mathbb{M}}(\mathbf{i}\!\uparrow\!\delta_1) \cdot \mathbf{i}\!\uparrow\!\delta_2\boldsymbol{\varepsilon}) .$$

where the morphism $\mathbf{P}_{\mathbb{M}} \in \mathrm{C}^1(\mathbb{TM}; \mathbb{TM})$ is the orthogonal projector on the fiberwise linear images of $\mathbb{TQ}$ by the tangent inclusion map $T\mathbf{i} \in \mathrm{C}^1(\mathbb{TQ}; \mathbb{TM})$. Then

$$\delta\boldsymbol{\sigma}(\boldsymbol{\varepsilon}) = \nabla_{\mathbf{D}}\mathcal{E} \cdot \delta\boldsymbol{\varepsilon} \in \mathbb{T}^*_{\boldsymbol{\varepsilon}}\mathbf{D} .$$

Let us denote by $\delta\boldsymbol{\sigma}_{\mathbf{D}}(\boldsymbol{\varepsilon}) \in \mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}$ the stress vector associated with the one form $\delta\boldsymbol{\sigma}(\boldsymbol{\varepsilon}) \in \mathbb{T}^*_{\boldsymbol{\varepsilon}}\mathbf{D}$ according to the relation $\mathbf{g}_{\mathbf{D}} \circ \delta\boldsymbol{\sigma}_{\mathbf{D}}(\boldsymbol{\varepsilon}) := \delta\boldsymbol{\sigma}(\boldsymbol{\varepsilon})$. Then, setting $\mathcal{E} = \mathbf{g}_{\mathbf{D}} \circ \mathcal{E}_{\mathbf{D}}$ and recalling the metric property of the LEVI-CIVITA connection, we have that

$$\mathbf{g}_{\mathbf{D}} \circ \delta\boldsymbol{\sigma}_{\mathbf{D}}(\boldsymbol{\varepsilon}) = \nabla_{\mathbf{D}}(\mathbf{g}_{\mathbf{D}} \circ \mathcal{E}_{\mathbf{D}}) \cdot \delta\boldsymbol{\varepsilon} = \mathbf{g}_{\mathbf{D}} \circ \nabla_{\mathbf{D}}\mathcal{E}_{\mathbf{D}} \cdot \delta\boldsymbol{\varepsilon} ,$$

that is

$$\delta\boldsymbol{\sigma}_{\mathbf{D}}(\boldsymbol{\varepsilon}) = \nabla_{\mathbf{D}}\mathcal{E}_{\mathbf{D}} \cdot \delta\boldsymbol{\varepsilon} = \mathbf{P}_{\mathbb{M}}(\nabla_{\mathbb{M}}(\mathbf{i}\!\uparrow\!\mathcal{E}_{\mathbf{D}}) \cdot \mathbf{i}\!\uparrow\!\delta\boldsymbol{\varepsilon}) .$$

The relation above may be rewritten in terms of the WEINGARTEN map, introduced in section 1.9.2 on page 198, as

$$\delta\boldsymbol{\sigma}_{\mathbf{D}}(\boldsymbol{\varepsilon}) = \nabla_{\mathbb{M}}(\mathbf{i}\!\uparrow\!\mathcal{E}_{\mathbf{D}}) \cdot \mathbf{i}\!\uparrow\!\delta\boldsymbol{\varepsilon} - \mathbf{W}(\mathbf{i}\!\uparrow\!\mathcal{E}_{\mathbf{D}}(\boldsymbol{\varepsilon}), \delta\boldsymbol{\varepsilon}) .$$

This is the expression of the linearized elastic law proposed in [139]. Since the WEINGARTEN map is tensorial, bilinear and symmetric, the second term on the r.h.s. vanishes if the stress $\mathcal{E}(\boldsymbol{\varepsilon}) = \mathbf{g}_{\mathbf{D}}(\mathcal{E}_{\mathbf{D}}(\boldsymbol{\varepsilon}))$ vanishes.

Let us now discuss of a special circumstance under which the dependence of linearization upon the chosen connection disappears.

If the elastic law admits a potential, that is if $\mathcal{E} = d\varphi_{\mathcal{E}} = \nabla\varphi_{\mathcal{E}}$, the linearized elastic law is expresed by the hessian of the potential. The hessian is defined, again by LEIBNIZ rule, as the second covariant derivative, through the identity:

$$\begin{aligned}\nabla_{\delta_1\boldsymbol{\varepsilon},\delta_2\boldsymbol{\varepsilon}}\varphi_{\mathcal{E}} &:= \nabla_{\delta_1\boldsymbol{\varepsilon}}\nabla_{\delta_2}\varphi_{\mathcal{E}} + \nabla_{\nabla_{\delta_1\boldsymbol{\varepsilon}}\delta_2}\varphi_{\mathcal{E}} \\ &= \nabla_{\boldsymbol{\varepsilon}}(\nabla_{\delta_2}\varphi_{\mathcal{E}}) \cdot \delta_1\boldsymbol{\varepsilon} + \nabla_{\boldsymbol{\varepsilon}}\varphi_{\mathcal{E}} \cdot \nabla_{\delta_1\boldsymbol{\varepsilon}}\delta_2 .\end{aligned}$$

Although the evaluation of the terms at the r.h.s. require that the vectors $\delta_1\boldsymbol{\varepsilon}, \delta_2\boldsymbol{\varepsilon} \in \mathbb{T}_{\boldsymbol{\varepsilon}}\mathbf{D}$ be extended to vector fields $\delta_1, \delta_2 \in \mathrm{C}^1(\mathbf{D}; \mathbb{T}\mathbf{D})$, their sum, and hence the l.h.s., is independent of the extension. Then the hessian is a twice covariant tensor field. Symmetry of the hessian is ensured if the connection is torsion-free.

Let $\boldsymbol{\varepsilon} \in \mathbf{D}$ be a critical point of the elastic potential $\varphi_{\mathcal{E}} \in \mathrm{C}^1(\mathbf{D}\,;\Re)$, that is a stress-free point: $\nabla_{\boldsymbol{\varepsilon}}\varphi_{\mathcal{E}} = \mathcal{E}(\boldsymbol{\varepsilon}) = 0$. Then, being $\nabla_{\boldsymbol{\varepsilon}}\varphi_{\mathcal{E}} \cdot \nabla_{\delta_1\boldsymbol{\varepsilon}}\delta_2 = 0$, the formula for the hessian gives

$$\nabla_{\delta_1\boldsymbol{\varepsilon},\delta_2\boldsymbol{\varepsilon}}\varphi_{\mathcal{E}} = \nabla_{\delta_1\boldsymbol{\varepsilon}}\nabla_{\delta_2}\varphi_{\mathcal{E}}\,.$$

Hence, at a critical point of the elastic potential, the hessian of the elastic potential is independent of the connection.

This result is a correction of the statement in [106], section 4.1.9, which claims, without proof, that if $\mathcal{E}(\boldsymbol{\varepsilon}) = 0$ then the linearized elastic law is independent of the connection.

4.3 Monotone laws and convex potentials

We shall denote by $\mathbf{x}$ a point of the domain $\boldsymbol{\Omega}$ occupied by the body. Let $\mathbf{D}$ and $\mathbf{S}$ be the dual finite dimensional vector spaces of local strains $\boldsymbol{\varepsilon}_{\mathbf{x}}$ and stresses $\boldsymbol{\sigma}_{\mathbf{x}}$. The subscript $\mathbf{x}$ recalls that we are dealing with local values, such as the variables appearing in the constitutive relations, of global fields pertaining to the whole structure. In this section we will deal only with local relations and hence the subscripts $\mathbf{x}$ will be dropped to simplify the notation.

A generalized elastic behaviour $\mathcal{E} \in BL(\mathbf{D}\,;\mathbf{S})$ is a relation between the local strain and stress spaces $\mathbf{D}$ and $\mathbf{S}$, such that the graph $\mathcal{G}(\mathcal{E}) \in \mathbf{D} \times \mathbf{S}$ fulfills the following properties:

i)	$\mathcal{G}(\mathcal{E})$	is maximal monotone,
ii)	$\mathcal{G}(\mathcal{E})$	is conservative,
iii)	$\mathrm{dom}\,\mathcal{E} \subset \mathbf{D}$ and $\mathrm{dom}\,\mathcal{E}^{-1} \subset \mathbf{S}$	are convex sets.

The definition of a *monotone graph* consists in extending to a general context the essential properties of a two-dimensional graph which is drawn giving increments of the same sign along two cartesian axes. A monotone graph can have horizontal, upward or vertical lines but no downward lines.

This means that the tangent stiffness of the material is nonnegative even if the tangent compliance may vanish. Hence, the material has a stable behaviour.

In fig.4.2 a generalized elastic behaviour which is multivalued in both directions is sketched. If a point $\{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}\} \in \mathbf{D} \times \mathbf{S}$ belongs to the graph of a

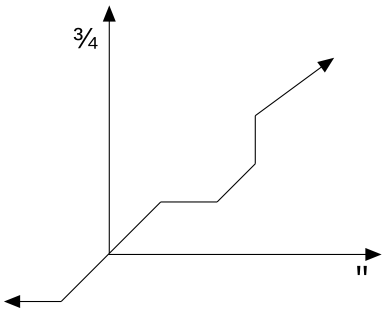

Figure 4.2: Typical generalized $\boldsymbol{\varepsilon} - \boldsymbol{\sigma}$ diagram

generalized elastic relation, we have that

$$\{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}\} \in \mathcal{G}(\mathcal{E}) \iff \boldsymbol{\sigma} \in \mathcal{E}(\boldsymbol{\varepsilon}) \iff \boldsymbol{\varepsilon} \in \mathcal{E}^{-1}(\boldsymbol{\sigma})\,.$$

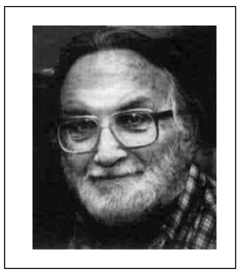

Figure 4.3: Max August Zorn (1906 - 1993)

Maximality of the graph requires that it can be drawn without lifting the pencil from the paper and extending (ideally) the graph in both directions so that it has no ends. From the conceptual point of view this last aspect is the most delicate to deal with. The proof of the existence of at least one maximal monotone graph is based on Zorn's Lemma (or equivalently on the Axiom of Choice) which is at the logical basis of modern mathematics [192].

- The formal statement of the *monotonicity property* requires that for any pair of points $\{\boldsymbol{\varepsilon}_1\,,\boldsymbol{\sigma}_1\}$ and $\{\boldsymbol{\varepsilon}_2\,,\boldsymbol{\sigma}_2\}$ belonging to $\mathbf{D}\times\mathbf{S}$ it is

$$\langle \boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1, \boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_1 \rangle \geq 0 \qquad \forall\ \{\boldsymbol{\varepsilon}_1\,,\boldsymbol{\sigma}_1\}\,,\{\boldsymbol{\varepsilon}_2\,,\boldsymbol{\sigma}_2\} \in \mathcal{G}(\mathcal{E})\,.$$

- The *maximality property* can be stated as follows: if a point $\{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}\}$ can be added to the graph without violating the property of monotonicity, then this point must belong to the graph. In formulae:

$$\langle \boldsymbol{\sigma} - \boldsymbol{\sigma}_g, \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_g \rangle \geq 0 \qquad \forall\, \{\boldsymbol{\varepsilon}_g\,,\boldsymbol{\sigma}_g\} \in \mathcal{G}(\mathcal{E}) \implies \{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}\} \in \mathcal{G}(\mathcal{E})\,.$$

The potential theory for monotone multivalued operators is developed herafter in its essential aspects; readers interested in a more detailed presentation are referred to [154]

- The *conservativity* of the map $\mathcal{E}$ requires the vanishing of the work associated with the map $\mathcal{E}$ along any *closed polyline* $\mathring{\Pi}_{\boldsymbol{\varepsilon}}$ included in dom $\mathcal{E} \subseteq \mathbf{D}$ (see fig.4.4):

$$\oint_{\mathring{\Pi}_{\boldsymbol{\varepsilon}}} \langle \mathcal{E}(\boldsymbol{\varepsilon}), d\boldsymbol{\varepsilon} \rangle = 0 \qquad \forall\, \mathring{\Pi}_{\boldsymbol{\varepsilon}} \subseteq \text{dom}\ \mathcal{E}.$$

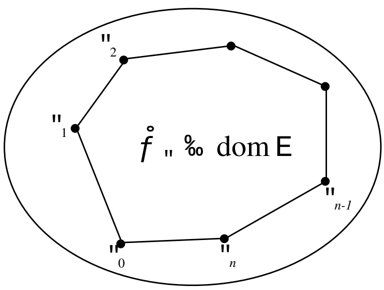

Figure 4.4: A closed polyline.

It is worth noting that the property of monotonicity of the map $\mathcal{E}$ ensures the existence of the integral along any segment belonging to dom $\mathcal{E}$; moreover, even if $\mathcal{E}(\boldsymbol{\varepsilon})$ is a set, the value of the integral does not depend on the particular choice of a point in the set $\mathcal{E}(\boldsymbol{\varepsilon})$.

Actually it can be proved [154] that, by virtue of the monotonicity of the graph $\mathcal{G}(\mathcal{E})$, the number of points in which the scalar product appearing in the integral above is multivalued is a set of null measure on any segment. Hence these points turn out to be inessential in the evaluation of the integral.

Let us now consider an arbitrary polyline $\Pi_{\boldsymbol{\varepsilon}}$ in $\mathbf{D}$ and let $i = 0 \ldots n$ be the number of its vertices. A *refinement* of $\Pi_{\boldsymbol{\varepsilon}}$ is any polyline included in $\Pi_{\boldsymbol{\varepsilon}}$.

By virtue of the monotonicity of the graph $\mathcal{G}(\mathcal{E})$ (see fig.4.5), the following formula holds for the integral along a polyline $\Pi_{\boldsymbol{\varepsilon}} \subseteq \mathrm{dom}\, \mathcal{E}$:

$$\sup\left\{\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle\right\} = \int_{\Pi_{\boldsymbol{\varepsilon}}} \langle \mathcal{E}(\boldsymbol{\varepsilon}), d\boldsymbol{\varepsilon} \rangle = \inf\left\{\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1}, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle\right\},$$

where the $\sum$ are referred to arbitrary refinements of $\Pi_{\boldsymbol{\varepsilon}}$ and the choice of $\boldsymbol{\sigma}_i \in \mathcal{E}(\boldsymbol{\varepsilon}_i)$ is inessential.

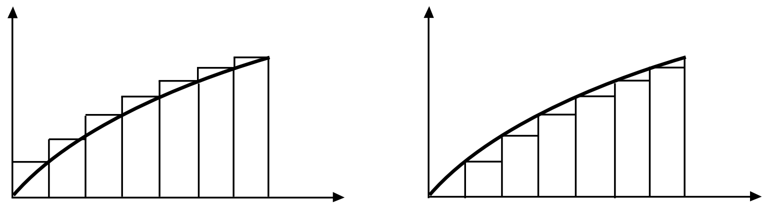

Figure 4.5:

If the map $\mathcal{E}$ is conservative, the vanishing of the integral along any closed polyline $\mathring{\Pi}_{\boldsymbol{\varepsilon}}$ implies the property of *cyclic monotonicity*:

- For any n-tuple $\{\boldsymbol{\varepsilon}_i\}$ with $i = 0 \ldots n$ and $\boldsymbol{\varepsilon}_n = \boldsymbol{\varepsilon}_0$, the following inequalities hold:

$$\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle \le 0\,, \qquad \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1}, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle \ge 0\,,$$

  where $\boldsymbol{\sigma}_i \in \mathcal{E}(\boldsymbol{\varepsilon}_i)$.

It is apparent that, vice versa, cyclic monotonicity of $\mathcal{E}$ implies conservativity.

Let us now remark that for any n-tuple of points $\{\boldsymbol{\varepsilon}_i\,, \boldsymbol{\sigma}_i\} \in \mathbf{D} \times \mathbf{S}$, with $i = 0 \ldots n$ and $\{\boldsymbol{\varepsilon}_n\,, \boldsymbol{\sigma}_n\} = \{\boldsymbol{\varepsilon}_0\,, \boldsymbol{\sigma}_0\}$, we have:

$$\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle = -\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} \rangle\,,$$

$$\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1}, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle = -\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_i \rangle\,.$$

It follows that the cyclic monotonicity of $\mathcal{E}$ implies the cyclic monotonicity of the inverse map $\mathcal{E}^{-1}$, i.e. for any n-tuple of vectors $\{\boldsymbol{\sigma}_i\}$ with $i = 0 \dots n$ and $\boldsymbol{\sigma}_n = \boldsymbol{\sigma}_0$, we have the inequalities:

$$\sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_i \rangle \leq 0 \,, \qquad \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} \rangle \geq 0 \,,$$

where $\boldsymbol{\varepsilon}_i \in \mathcal{E}^{-1}(\boldsymbol{\sigma}_i)$.

The cyclic monotonicity is then a characteristic property of the graph $\mathcal{G}(\mathcal{E})$.

On the basis of this result we can prove that, if $\mathcal{E}$ is conservative, its multivalued inverse map $\mathcal{E}^{-1}$ is conservative as well:

$$\oint_{\mathring{\Pi}_{\boldsymbol{\sigma}}} \langle \mathcal{E}^{-1}(\boldsymbol{\sigma}), d\boldsymbol{\sigma} \rangle = 0 \qquad \forall\, \mathring{\Pi}_{\boldsymbol{\sigma}} \subseteq \mathbf{S} \,,$$

where $\mathring{\Pi}_{\boldsymbol{\sigma}}$ is any closed polyline belonging to $\mathbf{S}$.

Therefore the conservativity property is an attribute of the graph $\mathcal{G}(\mathcal{E})$.

To prove the formula above we preliminarily note that the integral along a polyline $\Pi_{\boldsymbol{\sigma}} \subseteq \operatorname{dom} \mathcal{E}^{-1}$ fulfills the fundamental formula:

$$\sup \left\{ \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_i \rangle \right\} = \int_{\Pi_{\boldsymbol{\sigma}}} \langle \mathcal{E}^{-1}(\boldsymbol{\sigma}), d\boldsymbol{\sigma} \rangle = \inf \left\{ \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} \rangle \right\} ,$$

where the $\sum$ is referred to an arbitrary refinement of $\Pi_{\boldsymbol{\sigma}}$ and the choice of $\boldsymbol{\varepsilon}_i \in \mathcal{E}^{-1}(\boldsymbol{\sigma}_i)$ is inessential.

By virtue of the cyclic monotonicity of the map $\mathcal{E}^{-1}$, in the formula above the sup turns out to be nonpositive and the inf is nonnegative; hence the integral vanishes and the proof is complete.

For any conservative graph two complementary potentials $\phi : \mathbf{D} \mapsto \Re$ and $\psi : \mathbf{S} \mapsto \Re$ are associated with the multivalued monotone maps $\mathcal{E} : \mathbf{D} \mapsto \mathbf{S}$ and $\mathcal{E}^{-1} : \mathbf{S} \mapsto \mathbf{D}$.

Given a finite set of points $\{\boldsymbol{\varepsilon}_i \,, \boldsymbol{\sigma}_i\}$ with $i = 0, \dots, n+1$ belonging to $\mathcal{G}(\mathcal{E})$ such that $\{\boldsymbol{\varepsilon}_{n+1} \,, \boldsymbol{\sigma}_{n+1}\} = \{\boldsymbol{\varepsilon} \,, \boldsymbol{\sigma}\}$, the two complementary potentials $\phi : \mathbf{D} \mapsto \Re$ and $\psi : \mathbf{S} \mapsto \Re$ are recovered:

- $\phi : \mathbf{D} \mapsto \Re$ by integrating along the polyline $\Pi_{\boldsymbol{\varepsilon}} \subset \mathbf{D}$ having vertices $\{\boldsymbol{\varepsilon}_i\}$ with $i = 0, \dots, n+1$

- and $\psi : \mathbf{S} \mapsto \Re$ by integrating along the polyline $\Pi_{\boldsymbol{\sigma}} \subset \mathbf{S}$ having vertices $\{\boldsymbol{\sigma}_i\}$ with $i = 0, \dots, n+1$.

The properties of the graph $\mathcal{G}(\mathcal{E})$ ensure that the polylines $\Pi_{\boldsymbol{\varepsilon}}$ and $\Pi_{\boldsymbol{\sigma}}$ belong to the domains of $\mathcal{E}$ and of $\mathcal{E}^{-1}$ (see fig.4.6).

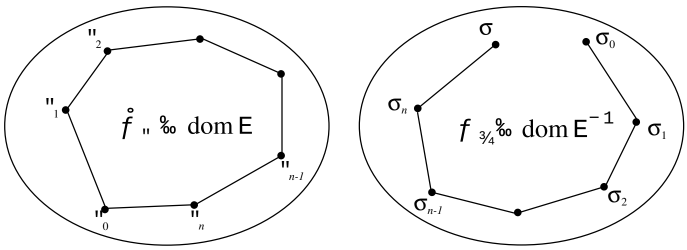

Figure 4.6: Polylines

The complementary potentials $\phi : \mathbf{D} \mapsto \Re$ and $\psi : \mathbf{S} \mapsto \Re$ can thus be defined on the domains of $\mathcal{E} : \mathbf{D} \mapsto \mathbf{S}$ and $\mathcal{E}^{-1} : \mathbf{S} \mapsto \mathbf{D}$ by integrating along the arbitrary polylines $\Pi_{\boldsymbol{\varepsilon}}$ and $\Pi_{\boldsymbol{\sigma}}$, according to the relations:

$$\phi(\boldsymbol{\varepsilon}) - \phi(\boldsymbol{\varepsilon}_o) = \int_{\Pi_{\boldsymbol{\varepsilon}}} \mathcal{E}\,, \qquad \psi(\boldsymbol{\sigma}) - \psi(\boldsymbol{\sigma}_o) = \int_{\Pi_{\boldsymbol{\sigma}}} \mathcal{E}^{-1}\,,$$

It is convenient to extend the two complementary potentials to extended real valued functions $\phi : \mathbf{D} \mapsto \mathcal{R} \cup \{+\infty\}$ and $\psi : \mathbf{S} \mapsto \mathcal{R} \cup \{+\infty\}$, by setting them equal to $+\infty$ respectively outside the domains of $\mathcal{E}$ and $\mathcal{E}^{-1}$.

To get the expressions of the potentials in a closed form, the integration can be conveniently performed along straight segments which join the initial and end points:

$$\phi(\boldsymbol{\varepsilon}) - \phi(\boldsymbol{\varepsilon}_o) = \int_0^1 \langle \mathcal{E}[\boldsymbol{\varepsilon}_o + t\,(\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_o)], \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_o \rangle dt\,,$$

$$\psi(\boldsymbol{\sigma}) - \psi(\boldsymbol{\sigma}_o) = \int_0^1 \langle \mathcal{E}^{-1}[\boldsymbol{\sigma}_o + t\,(\boldsymbol{\sigma} - \boldsymbol{\sigma}_o)], \boldsymbol{\sigma} - \boldsymbol{\sigma}_o \rangle dt.$$

The analysis of the properties of the potentials ϕ and ψ can be carried out by resorting to the basic properties of the integration of monotone multivalued maps along polylines. The following equalities can thus be inferred:

$$\phi(\boldsymbol{\varepsilon}) - \phi(\boldsymbol{\varepsilon}_o) = \sup\left\{ \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle + \langle \boldsymbol{\sigma}_n, \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_n \rangle \right\}$$

$$= \inf \left\{ \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1}, \boldsymbol{\varepsilon}_{i+1} - \boldsymbol{\varepsilon}_i \rangle + \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_n \rangle \right\},$$

where $\boldsymbol{\varepsilon} \in \mathrm{dom}\, \mathcal{E}$, $\boldsymbol{\sigma} \in \mathcal{E}(\boldsymbol{\varepsilon})$ and

$$\psi(\boldsymbol{\sigma}) - \psi(\boldsymbol{\sigma}_o) = \sup \left\{ \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_i \rangle + \langle \boldsymbol{\sigma} - \boldsymbol{\sigma}_n, \boldsymbol{\varepsilon}_n \rangle \right\}$$

$$= \inf \left\{ \sum_{i=0}^{n-1} \langle \boldsymbol{\sigma}_{i+1} - \boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_{i+1} \rangle + \langle \boldsymbol{\sigma} - \boldsymbol{\sigma}_n, \boldsymbol{\varepsilon} \rangle \right\},$$

where $\boldsymbol{\sigma} \in \mathrm{dom}\, \mathcal{E}^{-1}$ and $\boldsymbol{\varepsilon} \in \mathcal{E}^{-1}(\boldsymbol{\sigma})$. Note that the $\sum$ are referred to arbitrary refinements of $\Pi_{\boldsymbol{\varepsilon}}$ and $\Pi_{\boldsymbol{\sigma}}$.

We recall that:

- the *epigraph* of a function $\phi : \mathbf{D} \mapsto \mathcal{R} \cup \{+\infty\}$ is the subset $\mathrm{epi}\, \phi \subset \mathcal{E} \times \Re$ defined by:

$$\{\boldsymbol{\varepsilon}, \alpha\} \in \mathrm{epi}\, \phi \iff \alpha \geq \phi(\boldsymbol{\varepsilon}).$$

- a function $\phi : \mathbf{D} \mapsto \mathcal{R} \cup \{+\infty\}$ is *convex* if its epigraph is convex,
- a function $\phi : \mathbf{D} \mapsto \mathcal{R} \cup \{+\infty\}$ is *lower semicontinuous* if its epigraph is closed.

From the supremum formula above, we infer that the epigraph of the potential $\phi : \mathrm{dom}\, \mathcal{E} \mapsto \Re$ is convex, being the the intersection of a family of closed convex sets (closed half-spaces) with the convex set $(\mathrm{dom}\, \mathcal{E}) \times \Re$.

However, we cannot infer that the epigraph of $\phi : \mathrm{dom}\, \mathcal{E} \mapsto \Re$ is closed, unless $(\mathrm{dom}\, \mathcal{E}) \times \Re$ is closed. An analogous observation holds for the potential $\psi : \mathbf{S} \mapsto \Re$.

Let us now introduce a basic *invariance property* which links the potentials $\phi : \mathbf{D} \mapsto \mathcal{R} \cup \{+\infty\}$ and $\psi : \mathbf{S} \mapsto \mathcal{R} \cup \{+\infty\}$.

In fact we have:

$$
\begin{aligned}
\phi(\boldsymbol{\varepsilon})-\phi(\boldsymbol{\varepsilon}_o) &= \sup\left\{\sum_{i=0}^{n-1}\langle\,\boldsymbol{\sigma}_i,\boldsymbol{\varepsilon}_{i+1}-\boldsymbol{\varepsilon}_i\,\rangle+\langle\,\boldsymbol{\sigma}_n,\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}_n\,\rangle\right\}\\
&= \sup\left\{\sum_{i=0}^{n-1}\langle\,\boldsymbol{\sigma}_i-\boldsymbol{\sigma}_{i+1},\boldsymbol{\varepsilon}_{i+1}\,\rangle+\langle\,\boldsymbol{\sigma}_n-\boldsymbol{\sigma},\boldsymbol{\varepsilon}\,\rangle\right\}+\langle\,\boldsymbol{\sigma},\boldsymbol{\varepsilon}\,\rangle-\langle\,\boldsymbol{\sigma}_o,\boldsymbol{\varepsilon}_o\,\rangle\\
&= -\inf\left\{\sum_{i=0}^{n-1}\langle\,\boldsymbol{\sigma}_{i+1}-\boldsymbol{\sigma}_i,\boldsymbol{\varepsilon}_{i+1}\,\rangle+\langle\,\boldsymbol{\sigma}-\boldsymbol{\sigma}_n,\boldsymbol{\varepsilon}\,\rangle\right\}+\langle\,\boldsymbol{\sigma},\boldsymbol{\varepsilon}\,\rangle-\langle\,\boldsymbol{\sigma}_o,\boldsymbol{\varepsilon}_o\,\rangle\\
&= \psi(\boldsymbol{\sigma}_o)-\psi(\boldsymbol{\sigma})+\langle\,\boldsymbol{\sigma},\boldsymbol{\varepsilon}\,\rangle-\langle\,\boldsymbol{\sigma}_o,\boldsymbol{\varepsilon}_o\,\rangle\,.
\end{aligned}
$$

so that we have the:

- *invariance property*:

$$
\begin{aligned}
\phi(\boldsymbol{\varepsilon}_1)+\psi(\boldsymbol{\sigma}_1)-\langle\,\boldsymbol{\sigma}_1,\boldsymbol{\varepsilon}_1\,\rangle &= \phi(\boldsymbol{\varepsilon}_2)+\psi(\boldsymbol{\sigma}_2)-\langle\,\boldsymbol{\sigma}_2,\boldsymbol{\varepsilon}_2\,\rangle\,,\\
&\forall\,\{\boldsymbol{\varepsilon}_1\,,\boldsymbol{\sigma}_1\},\{\boldsymbol{\varepsilon}_2\,,\boldsymbol{\sigma}_2\}\in\mathcal{G}(\mathcal{E})\,.
\end{aligned}
$$

The *trinomial* $I(\boldsymbol{\varepsilon},\boldsymbol{\sigma}) = \phi(\boldsymbol{\varepsilon})+\psi(\boldsymbol{\sigma})-\langle\,\boldsymbol{\sigma},\boldsymbol{\varepsilon}\,\rangle$ is then a convex functional, defined on the product space $\mathbf{D}\times\mathbf{S}$, which is constant on the graph $\mathcal{G}(\mathcal{E})$.

Assuming $\{\boldsymbol{\varepsilon}_g\,,\boldsymbol{\sigma}_g\}\in\mathcal{G}(\mathcal{E})$, the monotonicity of the graph implies the following inequalities, equivalent to the property of subdifferentiability [72],[92],[149]:

$$
\begin{aligned}
\phi(\boldsymbol{\varepsilon})-\phi(\boldsymbol{\varepsilon}_g) &\geq \langle\,\boldsymbol{\sigma}_g,\boldsymbol{\varepsilon}-\boldsymbol{\varepsilon}_g\,\rangle \quad \forall\,\boldsymbol{\varepsilon}\in\mathbf{D} \iff \boldsymbol{\sigma}_g\in\partial\phi(\boldsymbol{\varepsilon}_g)\,,\\
\psi(\boldsymbol{\sigma})-\psi(\boldsymbol{\sigma}_g) &\geq \langle\,\boldsymbol{\sigma}-\boldsymbol{\sigma}_g,\boldsymbol{\varepsilon}_g\,\rangle \quad \forall\,\boldsymbol{\sigma}\in\mathbf{S} \iff \boldsymbol{\varepsilon}_g\in\partial\psi(\boldsymbol{\sigma}_g)\,.
\end{aligned}
$$

In the above relations, equality holds respectively if and only if we have $\{\boldsymbol{\varepsilon}_g\,,\boldsymbol{\sigma}_g\}\in\mathcal{G}(\partial\phi)$ and $\{\boldsymbol{\varepsilon}_g\,,\boldsymbol{\sigma}_g\}\in\mathcal{G}(\partial\psi)$.

The property of subdifferentiability is equivalent to the following inclusions:

$$
\mathcal{G}(\mathcal{E})\subseteq\mathcal{G}(\partial\phi)\,,\qquad \mathcal{G}(\mathcal{E})\subseteq\mathcal{G}(\partial\psi)\,.
$$

It is easy to prove that the graphs $\mathcal{G}(\partial\phi)$ and $\mathcal{G}(\partial\psi)$ are monotone (in particular cyclically monotone) and hence the maximality of the graph $\mathcal{G}(\mathcal{E})$ yields the equalities:

$$
\mathcal{G}(\mathcal{E})=\mathcal{G}(\partial\phi)=\mathcal{G}(\partial\psi)\,.
$$

On the basis of these properties, we infer that the trinomial invariant $I(\boldsymbol{\varepsilon},\boldsymbol{\sigma}) = \phi(\boldsymbol{\varepsilon}) + \psi(\boldsymbol{\sigma}) - \langle\boldsymbol{\sigma},\boldsymbol{\varepsilon}\rangle$ assumes on the graph $\mathcal{G}(\mathcal{E})$ an absolute minimum.

Actually the property of subdifferentiability is equivalent to require the minimum property:

$$I(\boldsymbol{\varepsilon},\boldsymbol{\sigma}_g) \geq I(\boldsymbol{\varepsilon}_g,\boldsymbol{\sigma}_g)\,, \quad \forall\,\boldsymbol{\varepsilon} \in \mathbf{D}\,, \qquad I(\boldsymbol{\varepsilon}_g,\boldsymbol{\sigma}) \geq I(\boldsymbol{\varepsilon}_g,\boldsymbol{\sigma}_g)\,, \quad \forall\,\boldsymbol{\sigma} \in \mathbf{S}\,,$$

with equality if and only if:

$$\{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}_g\} \in \mathcal{G}(\mathcal{E}) \qquad \text{and} \qquad \{\boldsymbol{\varepsilon}_g\,,\boldsymbol{\sigma}\} \in \mathcal{G}(\mathcal{E})\,.$$

This property implies that the graph $\mathcal{G}(\mathcal{E})$ is the minimal set for $I(\boldsymbol{\varepsilon},\boldsymbol{\sigma})$.

- The two complementary potentials ϕ and ψ are said to be *conjugate* one another if the corresponding integration constants are fixed so that $I(\boldsymbol{\varepsilon},\boldsymbol{\sigma})$ vanishes on the graph $\mathcal{G}(\mathcal{E})$.

The following FENCHEL's relations then hold [72], [92], [149]:

$$\phi(\boldsymbol{\varepsilon}) + \psi(\boldsymbol{\sigma}) \geq \langle\boldsymbol{\sigma},\boldsymbol{\varepsilon}\rangle\,, \quad \forall\,\{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}\} \in \mathbf{D}\times\mathbf{S}\,,$$

and

$$\phi(\boldsymbol{\varepsilon}) + \psi(\boldsymbol{\sigma}) = \langle\boldsymbol{\sigma},\boldsymbol{\varepsilon}\rangle \iff \{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}\} \in \mathcal{G}(\mathcal{E})\,.$$

FENCHEL's relations may be rewritten in the form:

$$\begin{aligned} \phi(\boldsymbol{\varepsilon}) &= \max\{\langle\boldsymbol{\sigma},\boldsymbol{\varepsilon}\rangle - \psi(\boldsymbol{\sigma}) \mid \boldsymbol{\sigma} \in \mathcal{S}\}\,, \quad \forall\,\boldsymbol{\varepsilon} \in \operatorname{dom}\,\mathcal{E}\,, \\ \psi(\boldsymbol{\sigma}) &= \max\{\langle\boldsymbol{\sigma},\boldsymbol{\varepsilon}\rangle - \phi(\boldsymbol{\varepsilon}) \mid \boldsymbol{\varepsilon} \in \mathcal{D}\}\,, \quad \forall\,\boldsymbol{\sigma} \in \operatorname{dom}\,\mathcal{E}^{-1}\,. \end{aligned}$$

As observed above, the conjugate potentials are convex but not necessarily lower semicontinuous, unless their domains are closed.

However, it is possible to carry out a regularization procedure by suitably modifying the values the potentials at the relative boundary points.

The modification consists in substituting, the value $+\infty$ with the smallest finite value compatible with the property of convexity. Hence the regularization can be performed by identifying, at the boundary points of their domains, the potentials with the lower semicontinuos convex functionals of which they are the restriction.

The expression of the regularized potentials are obtained by substituting the *max* with the *sup* and allowing the argument to range in the whole linear space.

In the interior of the domains, the two expressions coincide; at the relative boundary points equality holds after the regularization operation (closure of the epigraphs) has been performed:

$$\begin{aligned} \mathrm{cl}\,\phi(\boldsymbol{\varepsilon}) &= \sup\{\langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle - \psi(\boldsymbol{\sigma}) \mid \boldsymbol{\sigma} \in \mathcal{S}\} \quad \forall\, \boldsymbol{\varepsilon} \in \mathcal{E}\,, \\ \mathrm{cl}\,\psi(\boldsymbol{\sigma}) &= \sup\{\langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle - \phi(\boldsymbol{\varepsilon}) \mid \boldsymbol{\varepsilon} \in \mathcal{D}\} \quad \forall\, \boldsymbol{\sigma} \in \mathcal{S}\,. \end{aligned}$$

where the symbol $\mathrm{cl}\,\phi$ denotes the closed convex functional whose epigraph is the closure of the convex functional ϕ.

- The regularized potentials are known, in convex analysis, as the Fenchel's conjugate convex potentials, and are denoted by $\psi^*(\boldsymbol{\varepsilon})$ and $\phi^*(\boldsymbol{\sigma})$ [72], [92], [149], and we have:

$$\psi^*(\boldsymbol{\varepsilon}) = \mathrm{cl}\ \phi(\boldsymbol{\varepsilon})\,, \qquad \phi^*(\boldsymbol{\sigma}) = \mathrm{cl}\ \psi(\boldsymbol{\sigma})\,.$$

- The conjugate potentials ψ and ϕ are said to be *regular* if the following equalities hold:

$$\psi^*(\boldsymbol{\varepsilon}) = \phi(\boldsymbol{\varepsilon})\,, \qquad \phi^*(\boldsymbol{\sigma}) = \psi(\boldsymbol{\sigma})\,.$$

Regular potentials are lower semicontinuous and subdifferentiable at any point of their domains.

- A *regular graph* is a graph which is conservative, monotone and maximal and the such that two convex conjugate potentials associated with it are regular.

On the basis of the relations

$$\mathcal{G}(\partial\phi) \subseteq \mathcal{G}(\partial\phi^*) \ \text{ and } \ \mathcal{G}(\partial\psi) \subseteq \mathcal{G}(\partial\psi^*)\,,$$

the maximality property ensures that:

$$\mathcal{G}(\mathcal{E}) = \mathcal{G}(\partial\phi) = \mathcal{G}(\partial\psi) = \mathcal{G}(\partial\phi^*) = \mathcal{G}(\partial\psi^*)\,.$$

The convex functionals ϕ^* and ψ^* are subdifferentiable only at the points in which they coincide respectively with the potentials ψ and ϕ.

The possible differences between the potentials and their closure is exemplified in figs. 4.7 and 4.8. In fig.4.7 we consider a graph of an elastic behaviour in which the admissible strains $\boldsymbol{\varepsilon}$ must belong to the open set $\mathrm{dom}\,\mathcal{E}$. If $\boldsymbol{\varepsilon}$

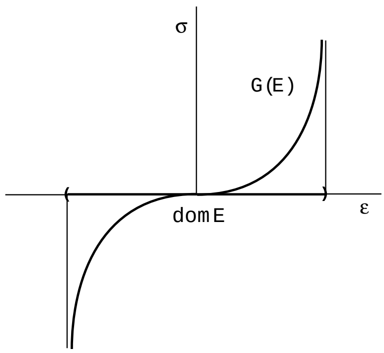

Figure 4.7: Open domain

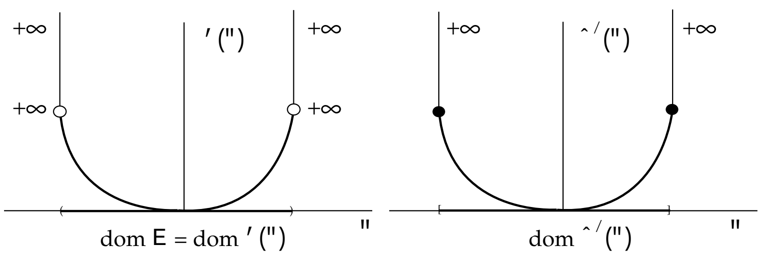

Figure 4.8: A potential and its regularization.

approaches to a boundary value, the stress $\boldsymbol{\sigma}$ goes to the infinity and in the boundary points the elastic potential ϕ is, by definition, $+\infty$.

Whenever the area under the graph has a finite value, the closure of the potential, or equivalently the functional ψ^*, is finite even at the boundary points but it is not subdifferentiable there (see fig.4.8).

In the sequel we will assume that the graphs of the constraint relations are regular. Accordingly, the tools of the subdifferential calculus can be applied.

In the mechanics of elastic structures, the convex potentials ϕ and ψ are respectively denoted the elastic energy and the complementary elastic energy.

In the following section we will show how the results of the theory outlined before, can be specialized to an elastic strictly monotone behaviour and, in particular, to a linear elastic behaviour.

4.3.1 Classical elasticity

By virtue of the monotonicity of the graph $\mathcal{G}(\mathcal{E})$, the potentials ϕ and ψ turn out to be convex but, in general, they do not result neither strictly convex nor differentiable.

In the classical theory of elasticity, the graph $\mathcal{G}(\mathcal{E})$ is assumed strictly monotone, i.e. monotone and such that:

$$\langle \boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1, \boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_1 \rangle = 0 \implies \{\boldsymbol{\varepsilon}_1 , \boldsymbol{\sigma}_1\} = \{\boldsymbol{\varepsilon}_2 , \boldsymbol{\sigma}_2\}.$$

Hence the complementary potentials ϕ and ψ are both strictly convex. Since strictly convexity of one of them implies the differentiability of the other one, it follows that both potentials turn out to be differentiable. The elastic behaviour is then one-to-one and we can write:

$$\boldsymbol{\sigma} = \mathcal{E}(\boldsymbol{\varepsilon}) \quad \text{and} \quad \boldsymbol{\varepsilon} = \mathcal{E}^{-1}(\boldsymbol{\sigma}).$$

Figure 4.9: Vito Volterra (1860 - 1940)

If the map $\mathcal{E}$ is continuously differentiable, the conservativity property can be ensured by imposing VOLTERRA's *symmetry condition* [188]:

$$\langle T_{\boldsymbol{\varepsilon}}\mathcal{E} \cdot \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2 \rangle = \langle T_{\boldsymbol{\varepsilon}}\mathcal{E} \cdot \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_1 \rangle \qquad \forall\, \boldsymbol{\varepsilon} \in \operatorname{dom} \mathcal{E} \qquad \forall\, \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2 \in \mathbf{D}\,.$$

If $\mathcal{E}$ is conservative, the inverse map $\mathcal{E}^{-1}$ is conservative too and its derivative is symmetric.

The strict convexity of the potentials implies that their second derivatives are positive definite (see fig.4.10).

If the elastic operator $\mathcal{E} \in \mathrm{C}^1(\mathbf{D}\,; \mathbf{S})$ is linear, we have that

$$T_{\boldsymbol{\varepsilon}}\mathcal{E} = \mathcal{E}\,,$$

and the conservativity of the operator $\mathcal{E}$ implies its symmetry by virtue of VOLTERRA's condition. Further, the strict convexity of the elastic potentials implies that $\mathcal{E}$ and $\mathcal{E}^{-1}$ are positive definite. Accordingly, the potentials ϕ and ψ are the positive definite quadratic forms:

$$\phi(\boldsymbol{\varepsilon}) = \frac{1}{2}\,\langle\, \mathcal{E}\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}\,\rangle\,, \qquad \psi(\boldsymbol{\sigma}) = \frac{1}{2}\,\langle\, \boldsymbol{\sigma}, \mathcal{E}^{-1}\boldsymbol{\sigma}\,\rangle\,,$$

which assume the same value when evaluated at any point of the graph $\mathcal{G}(\mathcal{E})$.

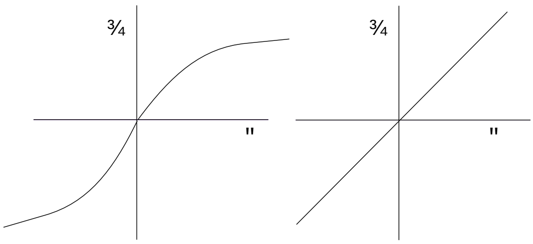

Figure 4.10: Nonlinear and linear elasticity

4.4 Global and local potentials

In order to analysing the equilibrium properties of a structural model with a generalized elastic behaviour, the constitutive relations must be written in global form, i.e. in terms of fields defined in the whole structure.

In a continuous model the global constitutive strain energy in the body with a generalized elastic behaviour, is a continuous functional $\varphi_{\mathcal{E}} \in \mathrm{C}^0(\mathcal{H}_{\mathbf{D}}\,;\Re)$ defined in $\mathcal{H}_{\mathbf{D}} = \mathcal{L}^2(\boldsymbol{\Omega}\,;\mathbf{D})$, the HILBERT space of square integrable strain fields over the domain $\boldsymbol{\Omega}$ occupied by the body.

We show that local subdifferential relations, enforced almost everywhere in $\boldsymbol{\Omega}$, are equivalently expressed in global form by integrating the relevant convex functions over the domain $\boldsymbol{\Omega}$.

To this end, let us defined the global elastic energy as the functional over the elastic strain fields $\boldsymbol{\varepsilon} \in \mathcal{H}_{\mathbf{D}}$ expressed by the integral of the specific elastic energy $\phi_{\mathbf{x}}$ over the whole body domain:

$$\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon}) = \int_{\boldsymbol{\Omega}} \phi_{\mathbf{x}}(\boldsymbol{\varepsilon}_{\mathbf{x}})\; d\mathbf{x}\,,$$

where the subscript $\boldsymbol{\varepsilon}_{\mathbf{x}}$ is the value of the field $\boldsymbol{\varepsilon} \in \mathcal{H}_{\mathbf{D}}$ at the point $\mathbf{x} \in \boldsymbol{\Omega}$.

Note that, whenever the local functions are convex, the corresponding global functional is convex as well, in the relevant fields.

Denoting by d^+ the *one-sided derivative* [149], the subdifferential of the global generalized elastic energy is locally defined by

$$\boldsymbol{\sigma} \in \partial\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon}) \iff d^+\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon};\boldsymbol{\eta}) \geq (\!(\, \boldsymbol{\sigma} \,,\, \boldsymbol{\eta} - \boldsymbol{\varepsilon} \,)\!)\,, \quad \forall\, \boldsymbol{\eta} \in \mathcal{H}_{\mathbf{D}}\,,$$

where:

$$d^+\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon};\boldsymbol{\eta}) = \int_{\boldsymbol{\Omega}} d^+\phi_{\mathbf{x}}(\boldsymbol{\varepsilon}_{\mathbf{x}};\boldsymbol{\eta}_{\mathbf{x}})\, d\mathbf{x}, \qquad (\!(\, \boldsymbol{\sigma} \,,\, \boldsymbol{\eta} - \boldsymbol{\varepsilon} \,)\!) = \int_{\boldsymbol{\Omega}} \boldsymbol{\sigma}_{\mathbf{x}} : (\boldsymbol{\eta}_{\mathbf{x}} - \boldsymbol{\varepsilon}_{\mathbf{x}})\, d\mathbf{x},$$

and the symbol $:$ denotes the scalar product between the local values of dual fields. The subdifferential of the local elastic energy is given by:

$$\boldsymbol{\sigma}_{\mathbf{x}} \in \partial\phi(\boldsymbol{\varepsilon}_{\mathbf{x}}) \iff d\phi(\boldsymbol{\varepsilon}_{\mathbf{x}};\boldsymbol{\eta}_{\mathbf{x}}) \geq \boldsymbol{\sigma}_{\mathbf{x}} : (\boldsymbol{\eta}_{\mathbf{x}} - \boldsymbol{\varepsilon}_{\mathbf{x}})\,, \quad \forall\, \boldsymbol{\eta}_{\mathbf{x}} \in \mathbf{D},$$

for almost every $\mathbf{x} \in \boldsymbol{\Omega}$ and the following equivalence are easily proved [140]:

$$\boldsymbol{\sigma} \in \partial\varphi_{\mathcal{E}}(\boldsymbol{\varepsilon}) \iff \boldsymbol{\sigma}_{\mathbf{x}} \in \partial\phi_{\mathbf{x}}(\boldsymbol{\varepsilon}_{\mathbf{x}}) \quad \text{a.e. in } \boldsymbol{\Omega}.$$

4.5 Elastic structures

In a structural model the state variables are given by two dual pairs:

- force systems $\mathbf{f} \in \mathcal{F}$ and displacement fields $\mathbf{u} \in \mathcal{V}$,
- stress fields $\boldsymbol{\sigma} \in \mathcal{H}_{\mathbf{S}}$ and strain fields $\boldsymbol{\varepsilon} \in \mathcal{H}_{\mathbf{D}}$,

where $\mathcal{H}_{\mathbf{D}} = \mathcal{L}^2(\boldsymbol{\Omega}\,;\mathbf{D})$ and $\mathcal{H}_{\mathbf{S}} = \mathcal{L}^2(\boldsymbol{\Omega}\,;\mathbf{S})$ are respectively the Hilbert spaces of square integrable stress and strain fields in $\boldsymbol{\Omega}$, $\mathcal{V}$ is the Hilbert space of Green-regular displacement fileds in $\boldsymbol{\Omega}$ and $\mathcal{F}$ is its topological dual, the linear space of force systems.

Between dual variables a regular generalized elastic relation of the type previously discussed is imposed.

As schematically depicted in fig.4.11, the relation between the internal variables $\{\boldsymbol{\varepsilon}, \boldsymbol{\sigma}\}$ is monotone nondecreasing while the relation between the external variables $\{\mathbf{u}, \mathbf{f}\}$ is monotone nonincreasing. We denote by:

$$\phi : \mathcal{H}_{\mathbf{D}} \mapsto \mathcal{R} \cup \{+\infty\}\,, \qquad \phi^* : \mathcal{H}_{\mathbf{S}} \mapsto \mathcal{R} \cup \{+\infty\}\,,$$

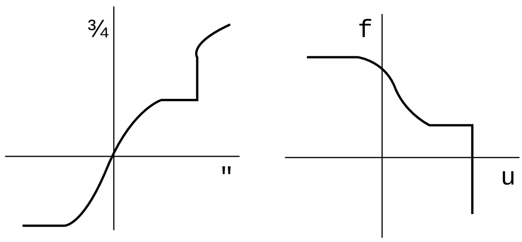

Figure 4.11: Constraint relations.

the convex conjugate potentials associated with the relation between the internal variables and by:

$$\gamma : \mathcal{V} \mapsto \mathcal{R} \cup \{-\infty\} \,, \qquad \gamma^* : \mathcal{F} \mapsto \mathcal{R} \cup \{-\infty\} \,,$$

the concave conjugate potentials associated with the relation between the external variables.

For simplicity of notation, we denote by the same symbol ∂ both the subdifferential operator of a convex functional and the supdifferential of a concave functional.

The problem of the elastic equilibrium can be written as follows [140], [151]:

$$\begin{cases} \mathbf{B}'\boldsymbol{\sigma} = \mathbf{f} \\ \mathbf{B}\,\mathbf{u} = \boldsymbol{\varepsilon} \\ \boldsymbol{\sigma} \in \partial\phi(\boldsymbol{\varepsilon}) \\ \mathbf{u} \in \partial\gamma^*(\mathbf{f}) \end{cases} , \quad \{\mathbf{u}, \mathbf{f}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}\} \in \mathcal{V} \times \mathcal{F} \times \mathcal{H}_{\mathbf{D}} \times \mathcal{H}_{\mathbf{S}} \,,$$

which, in terms of displacements and stresses becomes:

$$\begin{cases} \mathbf{B}'\boldsymbol{\sigma} \in \partial\gamma(\mathbf{u}) \,, \\ \mathbf{B}\,\mathbf{u} \in \partial\phi^*(\boldsymbol{\sigma}) \,. \end{cases}$$

Let us then consider the convex admissible domains of the state variables:

$$\begin{aligned} &\mathcal{U}_a &&= \text{dom}\ \gamma \subseteq \mathcal{V} && \text{admissible displacements,} \\ &\mathcal{F}_a &&= \text{dom}\ \gamma^* \subseteq \mathcal{F} && \text{admissible forces,} \\ &\mathcal{D}_a &&= \text{dom}\ \phi \subseteq \mathcal{H}_{\mathbf{D}} && \text{admissible strains,} \\ &\mathcal{S} - a &&= \text{dom}\ \phi^* \subseteq \mathcal{H}_{\mathbf{S}} && \text{admissible stresses.} \end{aligned}$$

Moreover let us define the domains:

$$\mathcal{C}_a = \{\mathbf{u} \in \mathcal{V} \mid \mathbf{B}\mathbf{u} \in \mathcal{D}_a\}\,, \qquad \Sigma_a = \{\boldsymbol{\sigma} \in \mathcal{H}_{\mathbf{S}} \mid \mathbf{B}'\boldsymbol{\sigma} \in \mathcal{F}_a\}\,,$$

of the displacements compatible with the admissible strains and of the stresses in equilibrium with the admissible forces.

4.6 Existence of a solution

In this section we prove the following result.

Theorem 4.6.1 (Existence conditions) *The problem of the elastic equilibrium admits a solution if and only if the constraint conditions are statically and kinematically admissible, i.e.*

$$\mathbf{B}'\mathcal{S} - a \cap \mathcal{F}_a \neq \emptyset\,, \qquad \mathbf{B}\mathcal{U}_a \cap \mathcal{D}_a \neq \emptyset\,,$$

or equivalently:

$$\mathcal{S} - a \cap \Sigma_a \neq \emptyset\,, \qquad \mathcal{U}_a \cap \mathcal{C}_a \neq \emptyset\,.$$

The condition of static compatibility $\mathbf{B}'\mathcal{S}-a\cap\mathcal{F}_a \neq \emptyset$ states that there exists at least an external admissible force in equilibrium with an internal admissible force.

The condition of kinematic compatibility $\mathbf{B}\mathcal{U}_a \cap \mathcal{D}_a \neq \emptyset$ states that there exists at least an admissible strain which is compatible with an admissible displacement.

The condition of static compatibility in the form $\mathcal{S} - a \cap \Sigma_a \neq \emptyset$ states that there exists at least an internal admissible force in equilibrium with an external admissible force.

The condition of kinematic compatibility in the form $\mathcal{U}_a \cap \mathcal{C}_a \neq \emptyset$ states that there exists at least an admissible displacement which corresponds to an admissible strain.

If there exists a solution, it is apparent that the two conditions of compatibility must be satisfied.

The proof that the two conditions above are also sufficient for the existence of a solution, is much more challenging. We provide here only a possible path of reasoning.

Firstly, it is convenient to re-state the problem in terms of one state variable: the displacement $\mathbf{u} \in \mathcal{V}$. To this end, substituting the condition of elastic compatibility:

$$\boldsymbol{\sigma} \in \partial\phi(\mathbf{B}\mathbf{u})\,,$$

in the equilibrium condition we get:

$$\mathbf{B}'\partial\phi(\mathbf{Bu}) \cap \partial\gamma(\mathbf{u}) \neq \emptyset \,.$$

Enforcing the subdifferential chain rule [72], [149]:

$$\partial(\phi \circ \mathbf{B})(\mathbf{u}) = \mathbf{B}'\partial\phi(\mathbf{Bu}) \,,$$

we can write the elastic equilibrium condition in the form:

$$\partial(\phi \circ \mathbf{B})(\mathbf{u}) \cap \partial\gamma(\mathbf{u}) \neq \emptyset \,,$$

or equivalently:

$$0 \in \partial(\phi \circ \mathbf{B})(\mathbf{u}) - \partial\gamma(\mathbf{u}) \,, .$$

By means of the additivity rule of the subdifferentials [155], the relation above becomes:

$$0 \in \partial(\phi \circ \mathbf{B} - \gamma)(\mathbf{u}) \,.$$

To prove that this subdifferential inclusion admits at least one solution we conjecture the following property.

Lemma 4.6.1 (Property of extension) *Let* $\mathbf{f} : \mathcal{X} \mapsto \mathcal{R} \cup \{+\infty\}$ *be a regular convex potential and* $\mathbf{f}_r : \mathcal{X} \mapsto \mathcal{R} \cup \{+\infty\}$ *its restriction to a closed convex set* $C \subseteq$ *dom* $\mathbf{f}$ *:*

$$\mathbf{f}_r(\mathbf{x}) = \begin{cases} \mathbf{f}(\mathbf{x}) & \text{if } \mathbf{x} \in C \,, \\ +\infty & \text{otherwise} \,. \end{cases}$$

Then the following formula holds:

$$im\, \partial\mathbf{f} \subseteq im\, \partial\mathbf{f}_r \,.$$

Proof. In fig. 4.12 it is shown how the property $\mathrm{im}\, \partial\mathbf{f} \subseteq \mathrm{im}\, \partial\mathbf{f}_r$ can be conjectured by observing the graphs of $\mathbf{f}$ and $\mathbf{f}_r$. ■

We can now prove proposition 4.6.1.

To prove the existence of a solution, let us consider the restrictions of ϕ to $\mathbf{B}\mathcal{U}_a \cap \mathcal{D}_a$ and of γ to $\mathcal{U}_a \cap \mathcal{C}_a$. The condition of the kinematic compatibility ensures that these restrictions have nonempty domains. Further, the extension property ensures that:

$$\mathbf{B}'\partial\phi(\mathcal{D}_a) \subseteq \mathbf{B}'\partial\phi(\mathbf{B}\mathcal{U}_a \cap \mathcal{D}_a) \,, \qquad \partial\gamma(\mathcal{U}_a) \subseteq \partial\gamma(\mathcal{U}_a \cap \mathcal{C}_a) \,,$$

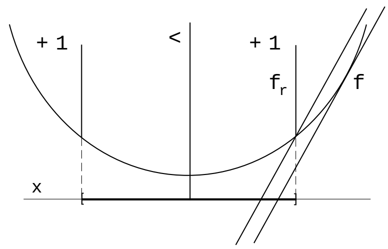

Figure 4.12: Extension property.

and the condition of static compatibility imposes:

$$\mathbf{B}'\partial\phi(\mathcal{D}_a) \cap \partial\gamma(\mathcal{U}_a) \neq \emptyset\,,$$

so that, *a fortiori*, we have: $\mathbf{B}'\partial\phi(\mathbf{B}\mathcal{U}_a \cap \mathcal{D}_a) \cap \partial\gamma(\mathcal{U}_a \cap \mathcal{C}_a) \neq \emptyset\,.$ The chain rule of subdifferential calculus allows us to write the equality:

$$\mathbf{B}'\partial\phi(\mathbf{B}\mathcal{U}_a \cap \mathcal{D}_a) = \partial(\phi\mathbf{B})(\mathcal{U}_a \cap \mathcal{C}_a)\,,$$

so that we have:

$$\partial(\phi \circ \mathbf{B})(\mathcal{U}_a \cap \mathcal{C}_a) \cap \partial\gamma(\mathcal{U}_a \cap \mathcal{C}_a) \neq \emptyset\,.$$

Finally, the additivity rule of the subdifferential calculus yields:

$$0 \in \partial(\phi \circ \mathbf{B} - \gamma)(\mathcal{U}_a \cap \mathcal{C}_a)$$

and the proposition 4.6.1 is proved. ■

It is worth noting that an analogous process can be repeated by stating the problem in terms of stresses.

4.7 Limit analysis

The following variational form is thus entailed for the static compatibility condition $\mathbf{B}'\mathcal{S} - a \cap \mathcal{F}_a \neq \emptyset$, [152]:

$$\inf_{\mathbf{f} \in \mathcal{F}_a} \langle \mathbf{f}, \mathbf{v} \rangle \leq \sup_{\boldsymbol{\sigma} \in \mathcal{S}-a} \langle \boldsymbol{\sigma}, \mathbf{B}\mathbf{v} \rangle \qquad \forall -\mathbf{v} \in \mathcal{N}_{\mathcal{F}_a}\,; \quad \mathbf{B}\mathbf{v} \in \mathcal{N}_{\mathcal{S}-a}$$

while the kinematic compatibility condition $\mathbf{B}\mathcal{U}_a \cap \mathcal{D}_a \neq \emptyset$ becomes:

$$\sup_{\boldsymbol{\varepsilon} \in \mathcal{D}_a} \langle \boldsymbol{\tau}, \boldsymbol{\varepsilon} \rangle \geq \inf_{\mathbf{u} \in \mathcal{U}_a} \langle \boldsymbol{\tau}, \mathbf{Bu} \rangle \qquad \forall\, \boldsymbol{\tau} \in \mathcal{N}_{\mathcal{D}_a} ; -\mathbf{B}'\boldsymbol{\tau} \in \mathcal{N}_{\mathcal{U}_a} .$$

Analogously it turns out to be:

$$\mathcal{S} - a \cap \Sigma_a \neq \emptyset \iff \sup_{\boldsymbol{\sigma} \in \mathcal{S}-a} \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle \geq \inf_{\boldsymbol{\sigma} \in \Sigma_a} \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle \qquad \forall\, \boldsymbol{\varepsilon} \in \mathcal{N}_{\mathcal{S}-a} \cap -\mathcal{N}_{\Sigma_a}$$

$$\mathcal{U}_a \cap \mathcal{C}_a \neq \emptyset \iff \sup_{\mathbf{u} \in \mathcal{U}_a} \langle \mathbf{f}, \mathbf{u} \rangle \geq \inf_{\mathbf{u} \in \mathcal{C}_a} \langle \mathbf{f}, \mathbf{u} \rangle \qquad \forall\, \mathbf{f} \in \mathcal{N}_{\mathcal{U}_a} \cap -\mathcal{N}_{\mathcal{C}_a} .$$

The discussion of the static compatibility condition is the field of the static limit analysis which is synthetically expounded in the sequel. In perfect duality an analogous treatment can be carried out for the kinematic compatibility condition which is the object of the kinematic limit analysis. Whenever the static compatibility condition is fulfilled, the set $\mathcal{S} - a \cap \Sigma_a$ is non-empty. Let then $\boldsymbol{\sigma}_o \in \mathcal{S} - a$ be an admissible stress which is in equilibrium with an admissible external force $\mathbf{B}'\boldsymbol{\sigma}_o = \mathbf{f_o} \in \mathcal{F}_a$. Further, let $\mathrm{Lin}\mathcal{F}_a$ and $\mathrm{Lin}\mathcal{S} - a$ be the subspaces parallel to the linear varieties generated by $\mathcal{F}_a$ and $\mathcal{S} - a$.

Let us introduce the definition of collapse mechanism. We shall say that $\mathbf{u_o} \in \mathcal{V}$ is a collapse mechanism if it turns out to be a free mechanism:

$$-\mathbf{u_o} \in \mathcal{N}_{\mathcal{F}_a}(\mathbf{f_o})$$

which is compatible with a collapse free deformation:

$$\boldsymbol{\varepsilon}_o = \mathbf{Bu_o} \in \mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma}_o)$$

$$\boldsymbol{\varepsilon}_o \notin (\mathrm{Lin}\mathcal{S} - a)^{\perp} \quad \mathbf{u_o} \notin (\mathrm{Lin}\mathcal{F}_a)^{\perp}$$

Three different kinds of mechanisms can be distinguished:

$$\begin{aligned} &\mathbf{u_o} \notin (\mathrm{Lin}\,\mathbf{B}'\mathcal{S} - a)^{\perp} && \text{internal collapse}\,, \\ &\mathbf{u_o} \notin (\mathrm{Lin}\,\mathcal{F}_a)^{\perp} && \text{external collapse}\,, \\ &\mathbf{u_o} \notin (\mathrm{Lin}\,\mathbf{B}'\mathcal{S} - a)^{\perp} \cap (\mathrm{Lin}\,\mathcal{F}_a)^{\perp} && \text{simultaneous collapse}\,. \end{aligned}$$

Let us prove the following fundamental result:

Proposition 4.7.1 (Fundamental theorem of limit analysis) *A collapse mechanism does exist if and only if the structure attains a static limit state, that is if and only if the admissible convex sets* $\mathbf{B}'\mathcal{S} - a$ *and* $\mathcal{F}_a$ *are separate.*

Proof. The equation of a hyperplane separating the sets $\mathcal{S} - a$ and Σ_a which contains a point $\boldsymbol{\sigma}_o \in \mathcal{S} - a \cap \Sigma_a$ is:

$$\langle \boldsymbol{\sigma} - \boldsymbol{\sigma}_o, \boldsymbol{\varepsilon}_o \rangle = 0, \quad \boldsymbol{\sigma} \in \mathcal{S}$$

so that the following inequalities do hold

$$\langle \boldsymbol{\sigma} - \boldsymbol{\sigma}_o, \boldsymbol{\varepsilon}_o \rangle \leq 0 \qquad \forall\, \boldsymbol{\sigma} \in \mathcal{S} - a$$

$$\langle \boldsymbol{\sigma} - \boldsymbol{\sigma}_o, \boldsymbol{\varepsilon}_o \rangle \geq 0 \qquad \forall\, \boldsymbol{\sigma} \in \Sigma_a;$$

hence:

$$\sup_{\boldsymbol{\sigma} \in \mathcal{S}-a} \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon}_o \rangle = \langle \boldsymbol{\sigma}_o, \boldsymbol{\varepsilon}_o \rangle = \inf_{\boldsymbol{\sigma} \in \Sigma_a} \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon}_o \rangle$$

or equivalently:

$$\boldsymbol{\varepsilon}_o \in \mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma}_o) \cap -\mathcal{N}_{\Sigma_a}(\boldsymbol{\sigma}_o)$$

The result follows then from the formula:

$$\mathcal{N}_{\Sigma_a}(\boldsymbol{\sigma}_o) = \mathbf{B}\mathcal{N}_{\mathcal{F}_a}(\mathbf{B}'\boldsymbol{\sigma}_o)$$

which can be obtained by recalling that:

$$\mathcal{N}_{\Sigma_a}(\boldsymbol{\sigma}_o) = \partial \sqcup_{\Sigma_a} (\boldsymbol{\sigma}_o)$$

and observing that, by definition, it turns out to be:

$$\sqcup_{\Sigma_a}(\boldsymbol{\sigma}_o) = \sqcup_{\mathcal{F}_a}(\mathbf{B}'\boldsymbol{\sigma}_o) = (\sqcup_{\mathcal{F}_a}\mathbf{B}')(\boldsymbol{\sigma}_o)\,.$$

By virtue of the chain rule of the subdifferentials we finally get:

$$\partial \sqcup_{\Sigma_a} (\boldsymbol{\sigma}_o) = \partial(\sqcup_{\mathcal{F}_a}\mathbf{B}')(\boldsymbol{\sigma}_o) = \mathbf{B}\partial \sqcup_{\mathcal{F}_a} (\mathbf{B}'\boldsymbol{\sigma}_o)$$

which provides the result. ■

Hence there exists $\mathbf{u_o} \in \mathcal{N}_{\mathcal{F}_a}(\mathbf{B}'\boldsymbol{\sigma}_o)$ such that $\boldsymbol{\varepsilon}_o = \mathbf{Bu_o}$.

The strict separation of the domains $\mathcal{F}_a$ and $\mathbf{B}'\mathcal{S} - a$ requires further that $\mathbf{u_o} \in (\mathrm{Lin}\,\mathcal{F}_a)^\perp$ or equivalently $\mathbf{Bu_o} \in (\mathrm{Lin}\,\mathcal{S} - a)^\perp$, i.e. $\mathbf{u_o}$ is a collapse mechanism.

The strain $\boldsymbol{\varepsilon}_o = \mathbf{Bu_o}$ represents the normal to the separating hyperplane of the convex sets $\mathcal{S} - a$ and Σ_a, while the collapse mechanism $\mathbf{u_o}$ represents the normal to the hyperplane separating the convex sets $\mathbf{B}'\mathcal{S} - a$ and $\mathcal{F}_a$.

The result is sketched in figs. 4.13 and 4.14.

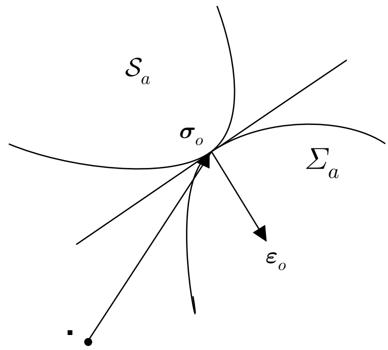

Figure 4.13:

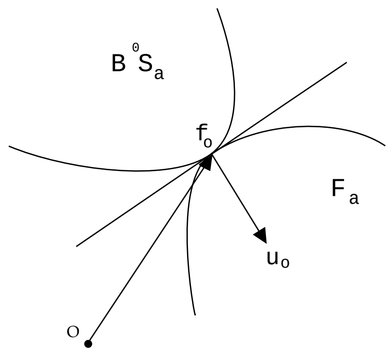

Figure 4.14:

4.7.1 Loading processes

Let us consider the case in which the convex set of the admissible external forces is defined as sum of a load variable with an affine law $\ell = \ell_o + \lambda\,\ell_d$ and of a fixed convex set of constraint reactions R_a :

$$\mathcal{F}_a = \ell_o + \lambda\ell_d + R_a$$

.

We are interested to evaluate the values of the loading parameter $\lambda \in \Re$ corresponding to the static limit conditions.

The condition expressing the admissibility of the load is given by:

$$\langle \ell_o, \mathbf{v} \rangle + \langle \lambda \ell_d, \mathbf{v} \rangle + \inf_{\mathbf{r} \in R_a} \langle \mathbf{r}, \mathbf{v} \rangle \leq \sup_{\boldsymbol{\sigma} \in \mathcal{S}-a} \langle \boldsymbol{\sigma}, \mathbf{B}\mathbf{v} \rangle \qquad \forall -\mathbf{v} \in \mathcal{N}_{R_a} \, ; \quad \mathbf{B}\mathbf{v} \in \mathcal{N}_{\mathcal{S}-a} .$$

that is:

$$\lambda \langle \ell_d, \mathbf{v} \rangle \leq \sup_{\boldsymbol{\sigma} \in \mathcal{S}-a} \langle \boldsymbol{\sigma}, \mathbf{B}\mathbf{v} \rangle - \langle \ell_o, \mathbf{v} \rangle - \inf_{\mathbf{r} \in R_a} \langle \mathbf{r}, \mathbf{v} \rangle \qquad \forall -\mathbf{v} \in \mathcal{N}_{R_a} \, ; \quad \mathbf{B}\mathbf{v} \in \mathcal{N}_{\mathcal{S}-a} \, .$$

Defining the convex set of the admissible loads as:

$$\Lambda_a = \mathbf{B}' \mathcal{S} - a - R_a$$

the convex set of the trial mechanisms:

$$\mathcal{N}_{\Lambda_a} = \{ \mathbf{v} \in \mathcal{V} \mid -\mathbf{v} \in \mathcal{N}_{R_a} , \, \mathbf{B}\mathbf{v} \in \mathcal{N}_{\mathcal{S}-a} \}$$

and the sublinear functional of the virtual dissipation:

$$\mathbf{D}(\mathbf{v}) = \sup_{\ell \in \Lambda_a} \langle \ell, \mathbf{v} \rangle - \langle \ell_o, \mathbf{v} \rangle = \sup_{\boldsymbol{\sigma} \in \mathcal{S}-a} \langle \boldsymbol{\sigma}, \mathbf{B}\mathbf{v} \rangle - \inf_{\mathbf{r} \in R_a} \langle \mathbf{r}, \mathbf{v} \rangle - \langle \ell_o, \mathbf{v} \rangle ,$$

we can express the admissibility condition of the loading parameter by means of the following three variational conditions:

$$\begin{aligned} 0 &\leq \mathbf{D}(\mathbf{v}) && \forall \, \mathbf{v} \in \mathcal{N}_{\Lambda_a} , \; \langle \ell_d, \mathbf{v} \rangle = 0 \, , \\ \lambda &\leq \mathbf{D}(\mathbf{v}) && \forall \, \mathbf{v} \in \mathcal{N}_{\Lambda_a} , \; \langle \ell_d, \mathbf{v} \rangle = 1 \, , \\ \lambda &\geq -\mathbf{D}(\mathbf{v}) && \forall \, \mathbf{v} \in \mathcal{N}_{\Lambda_a} , \; \langle \ell_d, \mathbf{v} \rangle = -1 \, , \end{aligned}$$

which can be equivalently stated in mechanical terms:

- It is *non-negative* the virtual dissipation associated with any trial mechanism for which the unit vector of the loading process performs *a null virtual power.*
- The loading parameter must be *not greater than* the virtual dissipation associated with every trial mechanism for which the unit vector of the loading process performs *a virtual power of unit value.*
- The loading parameter must be *not less than* the opposite of the virtual dissipation associated with every trial mechanism for which the unit vector of the loading process performs *a negative virtual power of unit value.*

The first variational condition amounts to imposing that:

$$\ell_o \in \Lambda_a + \operatorname{Lin} \ell_d$$

or equivalently that the reference load ℓ_o belongs to the cylinder having directrix Λ_a and generatrix ℓ_d. Provided that the previous condition is satisfied, we set:

$$\begin{aligned} \lambda^+ &= \inf\{\mathbf{D}(\mathbf{v}) \mid \mathbf{v} \in \mathcal{N}_{\Lambda_a},\ \langle \ell_d, \mathbf{v} \rangle = 1\} \\ \lambda^- &= \sup\{-\mathbf{D}(\mathbf{v}) \mid \mathbf{v} \in \mathcal{N}_{\Lambda_a},\ \langle \ell_d, \mathbf{v} \rangle = -1\}\,. \end{aligned}$$

The loading parameter will then turn out to be admissible if and only if:

$$\lambda^- \leq \lambda \leq \lambda^+$$

and it will yield a limit static condition when it does attain one of the extremal values. The contents of the previous discussion are exemplified in fig. 4.15.

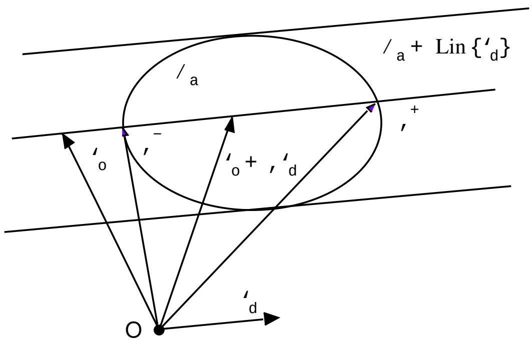

Figure 4.15:

4.7.2 Variational principles

We have shown how the problem of elastic equilibrium can be naturally expressed through the *subdifferential inclusions* [153]:

$$\begin{cases} \mathbf{B}'\boldsymbol{\sigma} \in \partial\gamma(\mathbf{u})\,, \\ \mathbf{B}\,\mathbf{u} \in \partial\phi^*(\boldsymbol{\sigma})\,. \end{cases}$$

By invoking FENCHEL's relations we realize that the two previous inclusions can be equivalently written as:

$$\phi(\mathbf{Bu}) + \phi^*(\boldsymbol{\sigma}) = \langle \boldsymbol{\sigma}, \mathbf{Bu} \rangle \,,$$

$$\gamma(\mathbf{u}) + \gamma^*(\mathbf{B}'\boldsymbol{\sigma}) = \langle \boldsymbol{\sigma}, \mathbf{Bu} \rangle \,.$$

Recalling that, for every $\mathbf{v} \in \mathcal{V}$ and $\boldsymbol{\tau} \in \mathcal{H}_{\mathbf{S}}$, one has:

$$\phi(\mathbf{Bv}) + \phi^*(\boldsymbol{\tau}) \geq \langle \boldsymbol{\tau}, \mathbf{Bv} \rangle \,,$$

$$\gamma(\mathbf{v}) + \gamma^*(\mathbf{B}'\boldsymbol{\tau}) \leq \langle \boldsymbol{\tau}, \mathbf{Bv} \rangle \,,$$

it follows that, for every $\mathbf{v} \in \mathcal{V}$, $\boldsymbol{\tau} \in \mathcal{H}_{\mathbf{S}}$, it is:

$$\phi(\mathbf{Bv}) - \gamma(\mathbf{v}) \geq -\phi^*(\boldsymbol{\tau}) + \gamma^*(\mathbf{B}'\boldsymbol{\tau}) \,,$$

$$\phi(\mathbf{Bu}) - \gamma(\mathbf{u}) = -\phi^*(\boldsymbol{\sigma}) + \gamma^*(\mathbf{B}'\boldsymbol{\sigma}) \,,$$

if and only if a pair $\{\mathbf{u}, \boldsymbol{\sigma}\}$ is a solution of the problem of the elastic equilibrium.

We define the convex functional *potential energy* and the concave functional *complementary energy* of the structural model as:

$$\mathrm{F}(\mathbf{u}) = \phi(\mathbf{Bu}) - \gamma(\mathbf{u}) \,,$$

$$\mathrm{G}(\boldsymbol{\sigma}) = -\phi^*(\boldsymbol{\sigma}) + \gamma^*(\mathbf{B}'\boldsymbol{\sigma}) \,.$$

A solution in terms of displacements and stresses can then be characterized as a minimum or maximum of these functionals.

$$\begin{cases} \mathbf{u} = \arg\,\min \mathrm{F} \\ \boldsymbol{\sigma} = \arg\,\max \mathrm{G} \end{cases}$$

and they assume the same value at a solution point:

$$\mathrm{F}(\mathbf{u}) = \mathrm{G}(\boldsymbol{\sigma}) \,.$$

Such classical extremum principles represent two elements of a family of stationarity principles equivalent to the problem of elastic equilibrium.

Their expression can be obtained by a direct approach based on the potential theory for multivalued monotone operators [154].

4.7.3 Hellinger-Reissner functional

Let us first examine the problem of elastic equilibrium formulated in terms of displacements and stresses.

Defining the dual spaces $\mathcal{X} = \mathcal{V} \times \mathcal{H}_{\mathbf{S}}$ and $\mathcal{X}' = \mathcal{F} \times \mathcal{H}_{\mathbf{D}}$, the structural problem assumes the form:

$$\begin{vmatrix} \mathbf{o} \\ \mathbf{o} \end{vmatrix} \in \mathbf{A} \begin{vmatrix} \mathbf{u} \\ \boldsymbol{\sigma} \end{vmatrix}$$

where the operator $\mathbf{A} : \mathcal{X} \mapsto \mathcal{X}'$ is defined by

$$\mathbf{A} = \begin{vmatrix} -\partial\gamma & \mathbf{B}' \\ \mathbf{B} & -\,\partial\phi^* \end{vmatrix} .$$

The operator $\mathbf{A}$ is sum of a linear symmetric operator and of two conservative monotone multivalued operators, respectively nonincreasing and nondecreasing, in the state variables $\mathbf{u}$ and $\boldsymbol{\sigma}$.

It follows that the operator $\mathbf{A}$ is conservative.

The relevant potential can be obtained in a direct way by integrating along the ray individuated by the point $\{\mathbf{u}\,, \boldsymbol{\sigma}\}$:

$$\int_{\{\mathbf{o},\mathbf{o}\}}^{\{\mathbf{u},\boldsymbol{\sigma}\}} \langle \mathbf{A} \begin{vmatrix} \bar{\mathbf{u}} \\ \bar{\boldsymbol{\sigma}} \end{vmatrix} , \begin{vmatrix} d\bar{\mathbf{u}} \\ d\bar{\boldsymbol{\sigma}} \end{vmatrix} \rangle =$$

$$= \int_{\{\mathbf{o},\mathbf{o}\}}^{\{\mathbf{u},\boldsymbol{\sigma}\}} \{ -\langle \partial\gamma(\bar{\mathbf{u}}), d\bar{\mathbf{u}} \rangle + \langle \mathbf{B}'\bar{\boldsymbol{\sigma}}, d\bar{\mathbf{u}} \rangle + \langle \mathbf{B}\bar{\mathbf{u}}, d\bar{\boldsymbol{\sigma}} \rangle - \langle \partial\phi^*(\bar{\boldsymbol{\sigma}}), d\bar{\boldsymbol{\sigma}} \rangle \}$$

which provides the expression of the potential:

$$R(\mathbf{u}, \boldsymbol{\sigma}) = -\gamma(\mathbf{u}) - \phi^*(\boldsymbol{\sigma}) + \langle \boldsymbol{\sigma}, \mathbf{B}\mathbf{u} \rangle \,.$$

This potential, which is the generalization of the potential known in the literature as the HELLINGER-REISSNER functional, is convex in $\mathbf{u}$ and concave in $\boldsymbol{\sigma}$. A solution of the generalized elastic problem is then a saddle point of $R(\mathbf{u}, \boldsymbol{\sigma})$:

$$\{\mathbf{u}, \boldsymbol{\sigma}\} = \arg\min\max R \,.$$

4.7.4 The variational tree

Formulating the problem of generalized elastic equilibrium in terms of all state variables, we get

$$\begin{cases} \mathbf{B}'\boldsymbol{\sigma} = \mathbf{f} \\ \mathbf{B}\,\mathbf{u} = \boldsymbol{\varepsilon} \\ \boldsymbol{\sigma} \in \partial\phi(\boldsymbol{\varepsilon}) \\ \mathbf{u} \in \partial\gamma^*(\mathbf{f}) \end{cases}$$

The dual product spaces are $\mathcal{X} = \mathcal{V} \times \mathcal{H}_{\mathbf{S}} \times \mathcal{H}_{\mathbf{D}} \times \mathcal{F}$ and $\mathcal{X}\,' = \mathcal{F} \times \mathcal{H}_{\mathbf{D}} \times \mathcal{H}_{\mathbf{S}} \times \mathcal{V}$ and the operator $\mathbf{A} : \mathcal{X} \mapsto \mathcal{X}\,'$ governing the structural problem is given by:

$$\mathbf{A} = \begin{bmatrix} \mathbf{O} & \mathbf{B}' & \mathbf{O} & -\mathbf{I}_{\mathcal{F}} \\ \mathbf{B} & \mathbf{O} & -\mathbf{I}_{\mathcal{D}} & \mathbf{O} \\ \mathbf{O} & -\mathbf{I}_{\mathcal{S}} & \partial\varphi_{\mathcal{E}} & \mathbf{O} \\ -\mathbf{I}_{\mathcal{V}} & \mathbf{O} & \mathbf{O} & \partial J^* \end{bmatrix}$$

By integrating along a ray in $\mathcal{X}$, we get the expression of the potential:

$$L(\boldsymbol{\varepsilon}, \boldsymbol{\sigma}, \mathbf{u}, \mathbf{f}) = \phi(\boldsymbol{\varepsilon}) + \gamma^*(\mathbf{f}) + \langle \boldsymbol{\sigma}, \mathbf{B}\mathbf{u} \rangle - \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle - \langle \mathbf{f}, \mathbf{u} \rangle\,,$$

which is convex in $\boldsymbol{\varepsilon}$, concave in $\mathbf{f}$ and linear in $\mathbf{u}$ and $\boldsymbol{\sigma}$. A solution $\{\boldsymbol{\varepsilon}, \boldsymbol{\sigma}, \mathbf{u}, \mathbf{f}\}$ then is a minimum point with respect to $\boldsymbol{\varepsilon}$, a maximum point with respect to $\mathbf{f}$ and a stationarity point with respect to $\mathbf{u}$ e $\boldsymbol{\sigma}$.

By properly eliminating the state variables, a family of ten potentials are generated according to the following tree-shaped scheme:

$$\{\boldsymbol{\varepsilon}, \boldsymbol{\sigma}, \mathbf{u}, \mathbf{f}\}$$

$$\{\boldsymbol{\varepsilon}, \boldsymbol{\sigma}, \mathbf{u}\} \qquad \{\boldsymbol{\sigma}, \mathbf{u}, \mathbf{f}\}$$

$$\{\boldsymbol{\varepsilon}, \boldsymbol{\sigma}\} \qquad \{\boldsymbol{\sigma}, \mathbf{u}\} \qquad \{\mathbf{u}, \mathbf{f}\}$$

$$\{\boldsymbol{\varepsilon}\} \qquad \{\boldsymbol{\sigma}\} \qquad \{\mathbf{u}\} \qquad \{\mathbf{f}\}$$

The variational family consists of the following ten potentials:

$$
\begin{aligned}
L(\boldsymbol{\varepsilon},\boldsymbol{\sigma},\mathbf{u},\mathbf{f}) &= \phi(\boldsymbol{\varepsilon}) + \gamma^*(\mathbf{f}) + \langle \boldsymbol{\sigma}, \mathbf{Bu} \rangle - \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle - \langle \mathbf{f}, \mathbf{u} \rangle, \\
H_1(\boldsymbol{\varepsilon},\boldsymbol{\sigma},\mathbf{u}) &= \phi(\boldsymbol{\varepsilon}) - \gamma(\mathbf{u}) + \langle \boldsymbol{\sigma}, \mathbf{Bu} \rangle - \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle, \\
H_2(\boldsymbol{\sigma},\mathbf{u},\mathbf{f}) &= -\phi^*(\boldsymbol{\sigma}) + \gamma^*(\mathbf{f}) + \langle \boldsymbol{\sigma}, \mathbf{Bu} \rangle - \langle \mathbf{f}, \mathbf{u} \rangle, \\
R_1(\boldsymbol{\varepsilon},\boldsymbol{\sigma}) &= \phi(\boldsymbol{\varepsilon}) + \gamma^*(\mathbf{B}'\boldsymbol{\sigma}) - \langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle, \\
R_2(\boldsymbol{\sigma},\mathbf{u}) &= -\phi^*(\boldsymbol{\sigma}) - \gamma(\mathbf{u}) + \langle \boldsymbol{\sigma}, \mathbf{Bu} \rangle, \\
R_3(\mathbf{u},\mathbf{f}) &= \phi(\mathbf{Bu}) + \gamma^*(\mathbf{f}) - \langle \mathbf{f}, \mathbf{u} \rangle, \\
P_1(\boldsymbol{\varepsilon}) &= \phi(\boldsymbol{\varepsilon}) - (\gamma^*\mathbf{B}')^*(\boldsymbol{\varepsilon}), \\
P_2(\boldsymbol{\sigma}) &= -\phi^*(\boldsymbol{\sigma}) + \gamma^*(\mathbf{B}'\boldsymbol{\sigma}), \\
P_3(\mathbf{u}) &= \phi(\mathbf{Bu}) - \gamma(\mathbf{u}), \\
P_4(\mathbf{f}) &= -(\phi \circ \mathbf{B})^*(\mathbf{f}) + \gamma^*(\mathbf{f}).
\end{aligned}
$$

All the potentials of the family do assume the same value at a solution point.

The extremum properties of each potential can be easily deduced by evaluating the convexity or concavity property with respect to each argument.

Remark 4.7.1 *The expression of the potentials P_1 and P_4 requires the evaluation of the conjugate functionals $(\gamma^* \circ \mathbf{B}')^*$ and $(\phi \circ \mathbf{B})^*$. Since they contain the deformation operator $\mathbf{B}$ of the structure, the evaluation of $(\gamma^* \circ \mathbf{B}')^*$ and $(\phi \circ \mathbf{B})^*$ requires the solution of an auxiliary problem of elastic equilibrium. For this reason the potentials P_1 and P_4 are not classically quoted in the literature. The extremum principle corresponding to P_4 can be applied in special circumstances in which the non-linearity of the problem is confined to the external constraint relation.*

4.7.5 Variational inequalities

The extremum properties of the functionals of the variational family can be expressed by requiring that the partial sub(sup)differentials with respect to each argument contain the null vector of the dual space.

Let us examine the case of the elastic potential functional whose minimum condition can be written as:

$$\mathbf{o} \in \partial \mathrm{F}(\mathbf{u}) \iff \mathrm{F}(\mathbf{v}) - \mathrm{F}(\mathbf{u}) \geq 0 \qquad \forall\, \mathbf{v} \in \mathcal{V}$$
$$\iff d\mathrm{F}(\mathbf{u};\mathbf{h}) \geq 0 \qquad \forall\, \mathbf{h} \in \mathcal{V}\,,$$

where d denotes the one-side derivative [149].

By making explicit the expression of F the extremum condition becomes

$$\phi(\mathbf{Bv}) - \phi(\mathbf{Bu}) \geq \gamma(\mathbf{v}) - \gamma(\mathbf{u}) \qquad \forall\, \mathbf{v} \in \mathcal{V}\,,$$

or equivalently, by virtue of the additivity property of the subdifferentials

$$d\phi(\mathbf{Bu};\mathbf{Bh}) \geq d\gamma(\mathbf{u};\mathbf{h}) \qquad \forall\, \mathbf{h} \in \mathcal{V}\,.$$

In the case of linear elasticity the elastic potential turns out to be quadratic. Denoting by b the bilinear form of the elastic energy, so that $\phi(\mathbf{u}) = 1/2\, b\,(\mathbf{u},\mathbf{u})\,,$ the variational inequality becomes:

$$b(\mathbf{u},\mathbf{h}) \geq d\gamma(\mathbf{u};\mathbf{h}) \qquad \forall\, \mathbf{h} \in \mathcal{V}.$$

Similar results can be written for all the functionals of the family.

4.7.6 Uniqueness of the solution

It has been shown how the issue of the existence of a solution for the problem of the generalized elastic equilibrium is amenable to a precise answer under very general hypotheses.

On the contrary the analysis of the properties of the solution set $\{\mathbf{u},\mathbf{f},\boldsymbol{\varepsilon},\boldsymbol{\sigma}\}$ can be performed more effectively by making use of the peculiar properties of the problem under examination. Actually, only simple considerations can be made in the general context.

If the elastic relation $\mathcal{E}$ is strictly monotone the potential ϕ turns is strictly convex and the solution in terms of deformations is unique. Uniqueness of the solution in terms of displacements is then guaranteed and it is defined apart from an additional rigid body which leaves the potential ϕ unaltered. If the elastic law $\mathcal{E}^{-1}$ is strictly monotone the solution in terms of stresses is unique; clearly this implies the uniqueness of the external forces at a solution point.

4.7.7 Barrier functionals

The constraint relations which have been considered thus far are described by potentials whose domains represent the admissible convex sets of the state variables.

In the applications the potentials are usually defined as sum of a regular convex potential defined on the whole space and the indicator functional of an admissible convex set.

For instance let us consider the case of an elastic relation in which a limitation is imposed to the range of the stresses.

The elastic complementary energy functional is then written as:

$$\phi^*(\boldsymbol{\sigma}) = \varphi^*(\boldsymbol{\sigma}) + \sqcup_{\mathcal{S}-a}(\boldsymbol{\sigma})\,,$$

where φ^* is strictly convex, hence differentiable, on $\mathcal{S}$. The elastic energy is provided by the conjugate potential:

$$\phi(\boldsymbol{\varepsilon}) = \sup\{\langle \bar{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon} \rangle - \phi^*(\bar{\boldsymbol{\sigma}}) \mid \bar{\boldsymbol{\sigma}} \in \mathcal{S}\}.$$

By recalling that the conjugate potential of the sum of convex functionals is given by the *inf-convolution* [149] of the conjugates of the addends, the following explicit expression can be obtained:

$$\phi(\boldsymbol{\varepsilon}) = \inf\{\varphi(\bar{\mathbf{e}}) + D(\bar{\boldsymbol{\delta}}) \mid \bar{\mathbf{e}} + \bar{\boldsymbol{\delta}} = \boldsymbol{\varepsilon}\}\,,$$

where $D = \sqcup^*_{\mathcal{S}-a}$ is the support functional of the convex set $\mathcal{S} - a$, defined by

$$\mathbf{D}(\boldsymbol{\varepsilon}) = \sqcup^*_{\mathcal{S}-a}(\boldsymbol{\varepsilon}) := \sup\{\langle \bar{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon} \rangle \mid \bar{\boldsymbol{\sigma}} \in \mathcal{S} - a\}\,.$$

The infimum appearing in the previous formula is attained in correspondence of the pairs $\{\mathbf{e}\,, \boldsymbol{\delta}\}$ such that:

$$\mathbf{e} \in \partial\varphi^*(\boldsymbol{\sigma})\,, \qquad \boldsymbol{\delta} \in \mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma})\,, \qquad \mathbf{e} + \boldsymbol{\delta} = \boldsymbol{\varepsilon}\,,$$

with $\boldsymbol{\sigma}$ conjugate of $\boldsymbol{\varepsilon}$ with respect to ϕ, i.e. : $\boldsymbol{\varepsilon} \in \partial\phi^*(\boldsymbol{\sigma})$.

The potential ϕ can then be expressed in the form:

$$\phi(\boldsymbol{\varepsilon}) = \varphi(\mathbf{e}) + D(\boldsymbol{\delta})\,.$$

The admissible domain $\mathcal{S} - a$ is usually defined as level set of a barrier convex functional $g : \mathcal{S} \mapsto \mathcal{R} \cup \{+\infty\}$:

$$\mathcal{S} - a = \{\boldsymbol{\sigma} \in \mathcal{S} \mid g(\boldsymbol{\sigma}) \leq 0\}\,.$$

The relevant indicator can then be rewritten as:

$$\sqcup_{\mathcal{S}-a}(\boldsymbol{\sigma}) = \sqcup_{R^-}[g(\boldsymbol{\sigma})] = (\sqcup_{R^-} \circ g)(\boldsymbol{\sigma}) \,,$$

so that:

$$\mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma}) = \partial \sqcup_{\mathcal{S}-a} (\boldsymbol{\sigma}) = \partial(\sqcup_{R^-} \circ g)(\boldsymbol{\sigma}).$$

The subdifferential of the functional $\sqcup_{R^-} \circ g$ can be evaluated by virtue of the following result contributed by the author in [155]

Let $m : \Re \mapsto \mathcal{R} \cup \{+\infty\}$ be a monotone convex function and $g : \mathcal{X} \mapsto \mathcal{R} \cup \{+\infty\}$ a continuous convex functional.

The composition $m \circ g : \mathcal{X} \mapsto \mathcal{R} \cup \{+\infty\}$ is then a convex functional and the relevant subdifferential at a point $\mathbf{x} \in \mathcal{X}$, which is not a minimum for g, is given by:

$$\partial(m \circ g)(x) = \partial m\,(g(x))\,\partial g(x)\,.$$

Applying the chain rule with $m = \sqcup_{R^-}$ one obtains:

$$\mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma}) = \partial \sqcup_{R^-} [g(\boldsymbol{\sigma})] \partial g(\boldsymbol{\sigma}) = \mathcal{N}_{R^-}[g(\boldsymbol{\sigma})] \partial g(\boldsymbol{\sigma}) \,,$$

and hence:

$$\delta \in \mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma}) \iff \delta \in \lambda\, \partial g(\boldsymbol{\sigma}) \quad \text{where} \quad \lambda \in \partial \sqcup_{R^-} [g(\boldsymbol{\sigma})]\,.$$

The parameter λ is the multiplier associated with the barrier functional g.

Let us notice that the following conditions are equivalent one another:

$$\begin{aligned}
&\lambda \in \partial \sqcup_{R^-} [g(\boldsymbol{\sigma})]\,,\\
&g(\boldsymbol{\sigma}) \in \partial \sqcup_{R^+} (\lambda)\,,\\
&\lambda \geq 0\,, \quad g(\boldsymbol{\sigma}) \leq 0\,, \quad \lambda g(\boldsymbol{\sigma}) = 0\,.
\end{aligned}$$

The last relations are referred to in the literature as *complementarity conditions*.

By making use of the chain rule, the generalized elastic relation $\boldsymbol{\varepsilon} \in \partial\phi^*(\boldsymbol{\sigma})$, can be written, in terms of the indicator of the admissible domain for the stresses, in the form:

$$\boldsymbol{\varepsilon} \in \partial\varphi^*(\boldsymbol{\sigma}) + \partial \sqcup_{\mathcal{S}-a} (\boldsymbol{\sigma})$$

and can be expressed in terms of a multiplier as:

$$\begin{cases} 0 \in g(\boldsymbol{\sigma}) - \partial \sqcup_{R^+} (\lambda) \\ 0 \in \partial\varphi^*(\boldsymbol{\sigma}) + \lambda \partial g(\boldsymbol{\sigma}) - \boldsymbol{\varepsilon}\,. \end{cases}$$

We consider then two multivalued operators $\mathcal{M}-\boldsymbol{\lambda}:\Re\times\mathcal{S}\mapsto\Re$ and $\mathcal{M}_{\boldsymbol{\sigma}}:\Re\times\mathcal{S}\mapsto\mathcal{D}$ defined by:

$$\begin{cases}\mathcal{M}-\boldsymbol{\lambda}(\lambda,\boldsymbol{\sigma})=g(\boldsymbol{\sigma})-\partial\sqcup_{R^+}(\lambda)\,,\\ \mathcal{M}_{\boldsymbol{\sigma}}(\lambda,\boldsymbol{\sigma})=\partial\varphi^*(\boldsymbol{\sigma})+\lambda\partial g(\boldsymbol{\sigma})-\boldsymbol{\varepsilon}\,.\end{cases}$$

The operator $\mathcal{M}-\boldsymbol{\lambda}(\lambda,\boldsymbol{\sigma})$ is monotone nonincreasing in λ and conservative, for any given $\boldsymbol{\sigma}\in\operatorname{dom}\varphi^*\cap\operatorname{dom}g$. The operator $\mathcal{M}_{\boldsymbol{\sigma}}(\lambda,\boldsymbol{\sigma})$ is monotone nondecreasing in $\boldsymbol{\sigma}$ and conservative, for any given $\lambda\in R^+$.

The previous inclusions can be written in symbolic form as:

$$\{0,0\}\in\mathcal{M}(\lambda,\boldsymbol{\sigma})\,,$$

with the operator $\mathcal{M}:\Re\times\mathcal{S}\mapsto\Re\times\mathcal{D}$ defined by:

$$\mathcal{M}(\lambda,\boldsymbol{\sigma})=\mathcal{M}-\boldsymbol{\lambda}(\lambda,\boldsymbol{\sigma})\times\mathcal{M}_{\boldsymbol{\sigma}}(\lambda,\boldsymbol{\sigma})\,.$$

The relevant Lagrangian potential can be evaluated by two successive integrations with respect to the two variables since the same result is obtained if the order of integration is inverted.

Apart from inessential integration constants, one obtains the following expression for the potential:

$$\mathcal{L}(\lambda,\boldsymbol{\sigma})=\varphi^*(\boldsymbol{\sigma})+\lambda\,g(\boldsymbol{\sigma})-\langle\boldsymbol{\sigma},\boldsymbol{\varepsilon}\rangle-\sqcup_{R^+}(\lambda)\,,$$

which is convex in $\boldsymbol{\sigma}$ and concave in λ.

The pair $\{\lambda,\boldsymbol{\sigma}\}$ associated with the deformation $\boldsymbol{\varepsilon}$ is thus a saddle point of the potential $\mathcal{L}$:

$$\{\lambda,\boldsymbol{\sigma}\}=\max_{\bar{\lambda}\in R^+}\min_{\bar{\boldsymbol{\sigma}}}\mathcal{L}(\bar{\lambda},\bar{\boldsymbol{\sigma}}).$$

The equivalence of the two problems expressed in terms of inclusions and of saddle point is inferred from:

$$\begin{cases}\mathcal{M}-\boldsymbol{\lambda}(\lambda,\boldsymbol{\sigma})=\partial_\lambda\,\mathcal{L}(\lambda,\boldsymbol{\sigma})\\[2ex] \mathcal{M}_{\boldsymbol{\sigma}}(\lambda,\boldsymbol{\sigma})=\partial_{\boldsymbol{\sigma}}\,\mathcal{L}(\lambda,\boldsymbol{\sigma})\end{cases}$$

Remark 4.7.2 *A generalized elastic model of the kind just described has been proposed in [150] to model the behaviour of structures made of elastic materials* *with no tensile strenght. In this case, the domain* $\mathcal{S} - a$ *is a convex cone and the functional* g *is the indicator of the negative polar* $\mathcal{S} - a^-$ *. Hence it is:*

$$\delta \in \mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma}) \iff \boldsymbol{\sigma} \in \mathcal{N}_{\mathcal{S}-a^-}(\delta)$$

$$\iff \boldsymbol{\sigma} \in \mathcal{S} - a, \quad \delta \in \mathcal{S} - a^-, \quad \langle \boldsymbol{\sigma}, \delta \rangle = 0\,.$$

The case of unilateral external constraints may be similarly dealt with.

4.8 Viscosity, viscoplasticity and plasticity

The generalized elastic model addressed in the paper represents a valuable reference model for treating the structural problems more usually referenced in the literature [178].

Actually a thorough analysis shows that the formal structure of such problems is completely similar to the ones considered in the previous paragraphs.

It is then possible to deduce in a systematic way the conditions on the existence and uniqueness of the solution and the variational formulations of the problem at hand by properly changing the formal treatment of the generalized elastic model.

Let us explicitly show how it is possible to frame within the reference formal context described in the paper the following structural models:

i) incremental plasticity,

ii) finite step viscosity, plasticity and visco-plasticity.

4.8.1 Incremental plasticity

In associated plasticity the plastic strain increment fulfills the normality rule to the convex elastic domain [100]:

$$\dot{\boldsymbol{\varepsilon}}_p \in \mathcal{N} \quad \text{with} \quad \mathcal{N} = \mathcal{N}_{\mathcal{S}-a}(\boldsymbol{\sigma})\,,$$

where the dot denotes right time derivative.

Assuming that $\dot{\boldsymbol{\varepsilon}}_p$ is continuous from the right it can be proved that the normality rule is equivalent to the following incremental law:

$$\dot{\boldsymbol{\varepsilon}}_p \in \mathcal{N}_{\mathcal{T}}(\dot{\boldsymbol{\sigma}}) \iff \dot{\boldsymbol{\sigma}} \in \mathcal{N}_{\mathcal{N}}(\dot{\boldsymbol{\varepsilon}}_p)$$

$$\iff \dot{\boldsymbol{\sigma}} \in \mathcal{T}\,, \quad \dot{\boldsymbol{\varepsilon}}_p \in \mathcal{N}\,, \quad \langle \dot{\boldsymbol{\sigma}}, \dot{\boldsymbol{\varepsilon}}_p \rangle = 0\,,$$

where $\mathcal{T} = \mathcal{N}^{-}$ is the tangent cone to the elastic domain $\mathcal{S} - a$ at the point $\boldsymbol{\sigma}$.

On expressing the original evolutive form in the incremental one, the convex domain $\mathcal{S} - a$ is thus substituted by the convex cone $\mathcal{T}$.

Hence the formulation of an elastoplastic problem in incremental terms leads to a structural model completely analogous to the one formulated for elastic no tension materials.

4.8.2 Finite step viscosity, plasticity and visco-plasticity

Let us consider a model which stands as a generalization of the one proposed by Perzyna [142] and which allows to cast the three kinds of constitutive behaviours within a unitary framework.

To this end let us imagine to assign a yield criterion defined by a continuous convex functional $y : \mathcal{S} \mapsto \mathcal{R} \cup \{+\infty\}$ with $y(0) = 0$, a scalar k which defines the yield limit and a flow function $m : \Re \mapsto \mathcal{R} \cup \{+\infty\}$ which is assumed to be convex, monotone nondecreasing and vanishing on $\Re^{-}$.

The barrier functional g, whose zero level set defines the elastic domain, is given by the difference between the functional y and the threshold value k:

$$g(\boldsymbol{\sigma}) = y(\boldsymbol{\sigma}) - k\,.$$

The potential is then expressed as chain of the barrier functional and the flow function:

$$\phi^{*} = m \circ g\,.$$

The viscoplastic flow rule is then given by:

$$\dot{\boldsymbol{\varepsilon}}_{vp} \in \partial\phi^{*}(\boldsymbol{\sigma}) = \partial(m \circ g)(\boldsymbol{\sigma}) = \partial m[g(\boldsymbol{\sigma})]\partial y(\boldsymbol{\sigma}).$$

The formulation of viscoplastic problems in kinematic terms requires the inversion of the constitutive relation and hence the evaluation of the functional of viscoplastic dissipation:

$$\phi(\dot{\boldsymbol{\varepsilon}}_{vp}) = (m \circ g)^{*}(\dot{\boldsymbol{\varepsilon}}_{vp})\,,$$

which is the conjugate of the viscoplastic potential ϕ^{*}.

The inverse relation will assume the form:

$$\boldsymbol{\sigma} \in \partial\phi(\dot{\boldsymbol{\varepsilon}}_{vp}) = \partial(m \circ g)^{*}(\dot{\boldsymbol{\varepsilon}}_{vp}).$$

The expression of the functional $(m \circ g)^*$ in terms of the conjugates of m and g is provided by the relation:

$$(mg)^*(\dot{\boldsymbol{\varepsilon}}_{vp}) = \inf_{\alpha} \{m^*(\alpha) + \sqcup^*_{\mathrm{epi}\,(g)}(\dot{\boldsymbol{\varepsilon}}_{vp}, -\alpha)\} \,,$$

where the support functional of the epigraph of g is given by:

$$\sqcup^*_{\mathrm{epi}\,(g)}(\dot{\boldsymbol{\varepsilon}}_{vp}, -\alpha) = \begin{cases} \alpha g^* \left(\dfrac{\dot{\boldsymbol{\varepsilon}}_{vp}}{\alpha} \right), & \text{if } \alpha > 0 \\ \sqcup^*_{\mathrm{dom}\ (g)}(\dot{\boldsymbol{\varepsilon}}_{vp}), & \text{if } \alpha = 0 \\ +\infty, & \text{if } \alpha < 0 \end{cases}$$

Hence the following formula does hold:

$$\phi(\dot{\boldsymbol{\varepsilon}}_{vp}) = (mg)^*(\dot{\boldsymbol{\varepsilon}}_{vp}) = \inf_{\alpha \geq 0} m^*(\alpha) + \begin{cases} \alpha g^* \left(\dfrac{\dot{\boldsymbol{\varepsilon}}_{vp}}{\alpha} \right) & \text{if } \alpha > 0 \\ \sqcup^*_{\mathrm{dom}\ (g)}(\dot{\boldsymbol{\varepsilon}}_{vp}) & \text{if } \alpha = 0 \end{cases}$$

The three different kinds of constitutive behaviour, namely viscoplastic, viscous and perfectly plastic, can be simulated by properly defining the threshold value k and the flow function m.

In the following schematic representations we will consider a sublinear yield functional y as it does occur in the von Mises criterion.

- A viscoplastic flow can be simulated by setting $k > 0$:
- A viscous flow of the Norton-Hoff kind can be simulated by setting $k = 0$ and hence $g = y$:
- A perfectly plastic law can be simulated by setting $k > 0$ and assuming as flow function the indicator of $\Re^-$.

Formulating the evolutive viscoplastic problem in terms of finite steps and adopting a backward time integration scheme [100], the constitutive law assumes the form:

$$\frac{\boldsymbol{\varepsilon}_{vp} - \boldsymbol{\varepsilon}_{vpo}}{\Delta t} \in \partial \phi^*(\boldsymbol{\sigma}) = \partial m[g(\boldsymbol{\sigma})] \partial y(\boldsymbol{\sigma}) \,,$$

while the inverse one becomes

$$\boldsymbol{\sigma} \in \partial \phi \left(\frac{\boldsymbol{\varepsilon}_{vp} - \boldsymbol{\varepsilon}_{vpo}}{\Delta t} \right) .$$

They are completely equivalent to the ones of the generalized elastic model, under the condition that the term $\boldsymbol{\varepsilon}_{vpo}$, which records the viscoplastic deformation at the end of the previous step, is properly accounted for.

4.9 Conclusions

The theory of generalized elasticity illustrated in this chapter embodies the principal features of inelastic behaviours and provides a simple and unifying framework for addressing the issues of existence and uniqueness of the solution and its variational characterizations.

The brief discussion on the viscoplastic behaviours exemplified in the last paragraph is not exhaustive but provides a hint for the application of the results of the theory of generalized elasticity to model the inelastic behaviours of materials and structures.

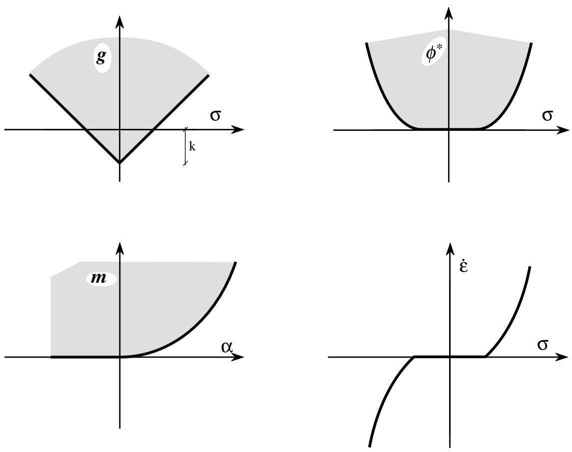

Figure 4.16: Viscoplasticity: direct potentials.

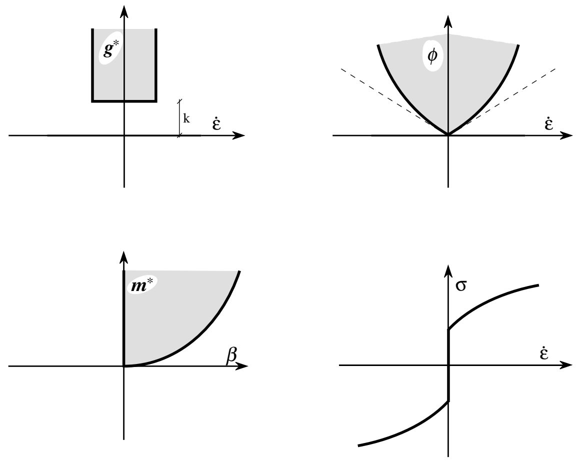

Figure 4.17: Viscoplasticity: conjugate potentials.

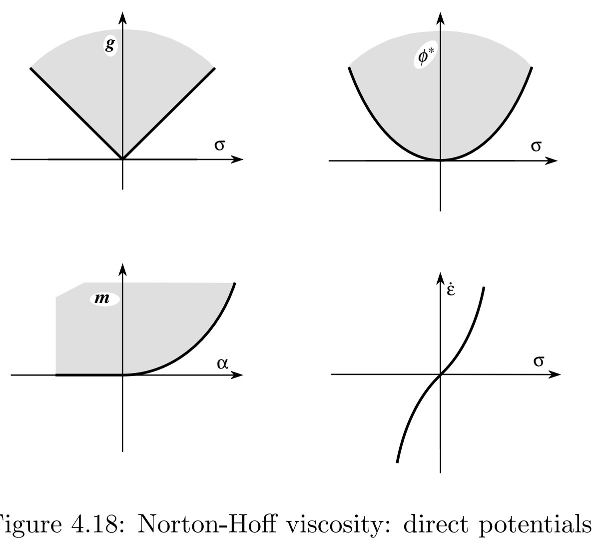

Figure 4.18: Norton-Hoff viscosity: direct potentials.

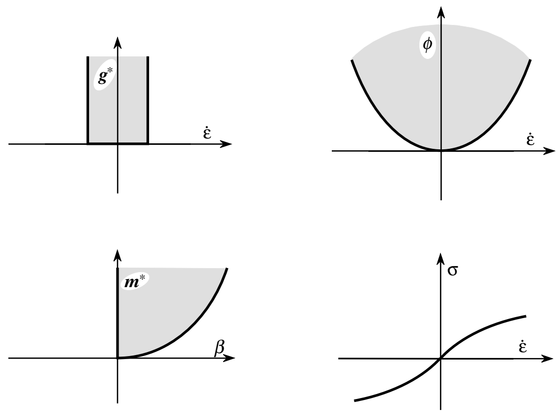

Figure 4.19: Norton-Hoff viscosity: conjugate potentials.

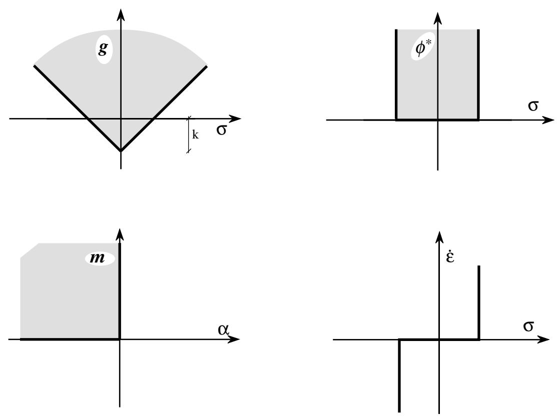

Figure 4.20: Perfect plasticity: direct potentials.

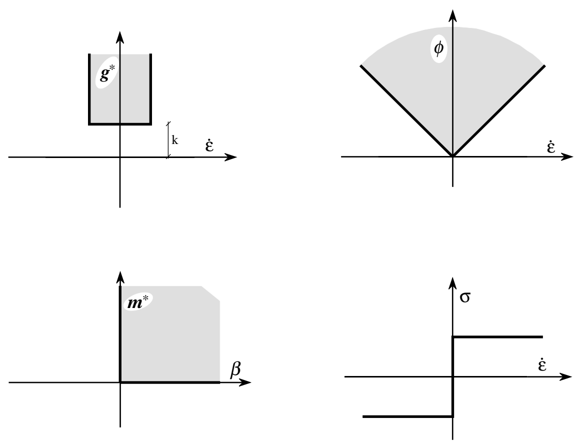

Figure 4.21: Perfect plasticity conjugate potentials.

Chapter 5

Constitutive behavior

The description of the mechanical properties of materials is a most challenging task for the design and the safety control of the dynamical behavior of a body subject to an history of actions. Materials respond in very different ways to the action of forces acted upon them by external agencies or by neighbouring bodies and often the response is time dependent in a very complex way.

By far the most important scheme of material behavior is the elastic model. A naïve description was due to the great experimentist Robert Hooke in 1676 in the form of the famous anagram *ceiiinosssttvv* of his statement that: *vt tensio sic vis* which asserts in latin language the proportionality between the elongation (*(ex)-tensio*) and the force (*vis*).

The modern form of the elastic law was envisaged by George Green in 1841 who, in his work on the propagation of light in cristallized media [65], conceived the existence of an elastic potential.

The elastic behavior is caracterized by the reversibility of the strain upon removal of the action (the stress field) and by the vanishing of the work stored in the material in a closed loop in the strain space.

Material behaviors other than elastic are usually dubbed anelastic. They include many important physical phenomena such as linear and nonlinear viscosity, plastic strains, thermally induced strains and phase-transition fronts propagating in the material.

We will not deal with the physics underlying these complex phenomena but will instead provide a synthetic treatment of the mathematical models envisaged for the description, at the continuum scale, of their most significant features.

5.1 Single-phase materials

The constitutive behavior of a single-phase material body is characterized by a natural placement $\mathbb{B}$, a differentiable submanifold embedded in the euclidean space $\{\mathbb{S}, \mathbf{g}\}$, by a square integrable metric tensor field $\mathbf{g_a} : \mathbb{B} \mapsto BL(\mathbb{TM}^2 ; \Re)$, describing the *anelastic* deformation field in $\mathbb{B}$, and by a differentiable scalar-valued function

$$W_{\mathbf{p}}((\boldsymbol{\varphi}{\downarrow}\mathbf{g})(\mathbf{p}), \mathbf{g_a}(\mathbf{p})) ,$$

which provides the elastic energy per unit volume at $\mathbf{p} \in \mathbb{B}$.

Dropping the explicit dependence on $\mathbf{p} \in \mathbb{B}$ of the arguments, the elastic energy in $\mathbb{B}$ is then given by

$$\int_{\mathbb{B}} W_{\mathbf{p}}(\boldsymbol{\varphi}{\downarrow}\mathbf{g}, \mathbf{g_a}) \ \boldsymbol{\mu} .$$

If the anelastic metric tensor field $\mathbf{g_a}$ coincides with the standard euclidean metric $\mathbf{g}$, the material behavior is in the *elastic range*. If the elastic energy density is independent of the position $\mathbf{p} \in \mathbb{B}$, the material is said to be elastically homogeneous.

The effects of irreversible changes in the microstructure of the material are taken into account by a assuming constitutive laws describing the variation of the anelastic metric tensor as a function of other state parameters, such as temperature, stress, time and a suitable set of internal variables.

The elastic behavior of the material is characterized by the requirement that, at any point $\mathbf{p} \in \mathbb{B}$, the PIOLA-KIRCHHOFF stress $\mathbf{s}^*(\mathbf{p}) := J_{\boldsymbol{\varphi}} \, \boldsymbol{\varphi}{\downarrow}\boldsymbol{\sigma}^*$ be the partial derivative of the elastic energy density $W_{\mathbf{p}}(\boldsymbol{\varphi}{\downarrow}\mathbf{g}, \mathbf{g_a})$ with respect to the configuration induced metric tensor:

$$\mathbf{s}^*(\mathbf{p}) = \partial_1 W_{\mathbf{p}}(\boldsymbol{\varphi}{\downarrow}\mathbf{g}, \mathbf{g_a}) .$$

This elastic law is due in essence to GEORGE GREEN [65].

The anelastic stress is defined to be the opposite of partial derivative of the elastic energy density $W_{\mathbf{p}}(\boldsymbol{\varphi}{\downarrow}\mathbf{g}, \mathbf{g_a})$ with respect to the anelastic metric tensor:

$$\mathbf{s}^*_{\mathbb{M}}(\mathbf{p}) := -\partial_2 W_{\mathbf{p}}(\boldsymbol{\varphi}{\downarrow}\mathbf{g}, \mathbf{g_a}) .$$

One would like that the constitutive laws retain the same formal expression when written in terms of the pushed-forward tensors, the KIRCHHOFF stress tensor $\boldsymbol{\tau}^* := \boldsymbol{\varphi}{\uparrow}\mathbf{s}^*$ and the anelastic stress tensor $\boldsymbol{\tau}^*_{\mathbb{M}} := \boldsymbol{\varphi}{\uparrow}\mathbf{s}^*_{\mathbb{M}}$ in the current placement:

$$\boldsymbol{\tau}^*(\boldsymbol{\varphi}(\mathbf{p})) = \partial_1 W_{\boldsymbol{\varphi}(\mathbf{p})}(\mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a})\,,$$

$$\boldsymbol{\tau}^*_{\mathbb{M}}(\boldsymbol{\varphi}(\mathbf{p})) = -\partial_2 W_{\boldsymbol{\varphi}(\mathbf{p})}(\mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a})\,.$$

The result provided in section 1.1.5 on page 14 shows that, to get these expressions, we have to define the elastic energy density, in the current placement, as

$$W_{\boldsymbol{\varphi}(\mathbf{p})} := W_{\mathbf{p}} \circ \boldsymbol{\varphi}\!\downarrow\,,$$

so that the elastic energy in $\boldsymbol{\varphi}(\mathbb{B})$ is given by

$$\int_{\boldsymbol{\varphi}(\mathbb{B})} W_{\boldsymbol{\varphi}(\mathbf{p})}(\mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a})\ \boldsymbol{\varphi}\!\uparrow\!\boldsymbol{\mu}\,,$$

In passing from a configuration $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ to a configuration $\boldsymbol{\xi} \circ \boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$, the transformation rule is accordingly given by

$$W_{(\boldsymbol{\xi}\circ\boldsymbol{\varphi})(\mathbf{p})} = \boldsymbol{\xi}\!\uparrow\! W_{\boldsymbol{\varphi}(\mathbf{p})} := W_{\boldsymbol{\varphi}(\mathbf{p})} \circ \boldsymbol{\xi}\!\downarrow\,,$$

or explicitly

$$W_{(\boldsymbol{\xi}\circ\boldsymbol{\varphi})(\mathbf{p})}(\boldsymbol{\xi}\!\uparrow\!\mathbf{g}, (\boldsymbol{\xi}\circ\boldsymbol{\varphi})\!\uparrow\!\mathbf{g_a}) := W_{\boldsymbol{\varphi}(\mathbf{p})}(\mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a})\,,$$

for any diffeomorphism $\boldsymbol{\xi} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{S})$. This transformation rule, dictated by form invariance of the constitutive laws, may also be written as:

$$W_{\boldsymbol{\varphi}(\mathbf{p})} = \boldsymbol{\xi}\!\downarrow\! W_{(\boldsymbol{\xi}\circ\boldsymbol{\varphi})(\mathbf{p})} := W_{(\boldsymbol{\xi}\circ\boldsymbol{\varphi})(\mathbf{p})} \circ \boldsymbol{\xi}\!\uparrow\,.$$

5.2 The covariance constitutive axiom

In the wake of the treatment developed by Marsden and Hughes in [106], some authors (see e.g. Simo [179]) prefer to start with a seemingly more general approach to constitutive relations in thermoelasticity and in elastoplasticity. In the present context, stated in precise terms, their proposal consists in assuming that the elastic energy density $W_{\boldsymbol{\varphi}(\mathbf{p})}$ at a placement $\boldsymbol{\varphi}(\mathbb{B})$ depends on the configuration map, on the euclidean metric and on the push-forward of the anelastic metric tensor:

$$W_{\boldsymbol{\varphi}(\mathbf{p})}(\boldsymbol{\varphi}, \mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a})\,.$$

To get a physically acceptable expression of the elastic energy density, this approach compels to invoke a covariance constitutive axiom which imposes that, at any configuration $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ and for any diffeomorphism $\boldsymbol{\xi} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{S})$, the following equality must hold:

$$W_{\boldsymbol{\varphi}(\mathbf{p})}(\boldsymbol{\varphi}, \mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a}) := W_{(\boldsymbol{\xi}\circ\boldsymbol{\varphi})(\mathbf{p})}(\boldsymbol{\xi}\circ\boldsymbol{\varphi}, \boldsymbol{\xi}\!\uparrow\!\mathbf{g}, (\boldsymbol{\xi}\circ\boldsymbol{\varphi})\!\uparrow\!\mathbf{g_a})\,.$$

Defining, for notational convenience, the push-forward of a diffeomorphism by $\boldsymbol{\xi}\!\uparrow\!\boldsymbol{\varphi} := \boldsymbol{\xi}\circ\boldsymbol{\varphi}$, the covariance constitutive axiom may be written in the simple form

$$W_{\boldsymbol{\varphi}(\mathbf{p})} = \boldsymbol{\xi}\!\downarrow\! W_{(\boldsymbol{\xi}\circ\boldsymbol{\varphi})(\mathbf{p})}\,,$$

with the pull-back $\boldsymbol{\xi}\!\downarrow\! W$ defined as usual by:

$$(\boldsymbol{\xi}\!\downarrow\! W)_{\boldsymbol{\varphi}(\mathbf{p})}(\boldsymbol{\varphi}, \mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a}) := W_{(\boldsymbol{\xi}\circ\boldsymbol{\varphi})(\mathbf{p})}(\boldsymbol{\xi}\!\uparrow\!\boldsymbol{\varphi}, \boldsymbol{\xi}\!\uparrow\!\mathbf{g}, \boldsymbol{\xi}\!\uparrow\!\boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a})\,.$$

This pull-back operation consists in changing the configuration $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ into $\boldsymbol{\xi}\circ\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ and the metric tensor $\mathbf{g}$ into its push-forward $\boldsymbol{\xi}\!\uparrow\!\mathbf{g}$ according to the diffeomorphism $\boldsymbol{\xi} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{S})$.

The covariance axiom states an invariance property under arbitrary diffeomorphisms $\boldsymbol{\xi} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{S})$ and is then a generalized version of the principle of material frame indifference, which states invariance under isometric diffeomorphisms, characterized by the property $\boldsymbol{\xi}\!\uparrow\!\mathbf{g} = \mathbf{g}$.

Setting $\boldsymbol{\xi} = \boldsymbol{\varphi}^{-1}$ the covariance axiom implies that

$$W_{\boldsymbol{\varphi}(\mathbf{p})}(\boldsymbol{\varphi}, \mathbf{g}, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a}) = W_{\mathbf{p}}(\mathbf{i}, \boldsymbol{\varphi}\!\downarrow\!\mathbf{g}, \mathbf{g_a})\,,$$

with $\mathbf{i} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{B})$ the identity map. The dependence of the elastic energy density $W_{\mathbf{p}}$ on the pull-back of the metric tensor, along the configuration map, and on the anelastic metric tensor, is thus recovered.

We may then conclude that the requirement of form invariance of the constitutive law, with the elastic energy density $W_{\boldsymbol{\varphi}(\mathbf{p})}$ at a placement $\boldsymbol{\varphi}(\mathbb{B})$ written in terms of the tensors $\mathbf{g}\,, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a}$ is equivalent to the constitutive covariance axiom with $W_{\boldsymbol{\varphi}(\mathbf{p})}$ written in terms of $\boldsymbol{\varphi}, \mathbf{g}\,, \boldsymbol{\varphi}\!\uparrow\!\mathbf{g_a}$.

Remark 5.2.1 *By giving to the elastic energy density the special form:*

$$W_{\mathbf{p}}(\boldsymbol{\varphi}\!\downarrow\!\mathbf{g} - \mathbf{g_a})\,,$$

we have that

$$\mathbf{s}^*(\mathbf{p}) = \mathbf{s}^*_{\mathbb{M}}(\mathbf{p}) = dW_{\mathbf{p}}(\boldsymbol{\varphi}\!\downarrow\!\mathbf{g} - \mathbf{g_a})\,.$$

Remark 5.2.2 *The configuration-induced metric and the anelastic metric are conveniently descibed as linear operators by means of the representation induced by the euclidean metric* $\mathbf{g} \in BL(\mathbb{TS}^2 ; \Re)$ *. We then get the (* $\mathbf{g}$*-symmetric and positive definite) linear* $\mathbf{g}$*-operators:*

$$d\boldsymbol{\varphi}^T d\boldsymbol{\varphi} \in BL(\mathbb{TM} ; \mathbb{TM}) , \qquad \textit{Piola-Green operator},$$
$$\mathbf{G}_\mathbf{a} \in BL(\mathbb{TM} ; \mathbb{TM}) , \qquad \textit{reference anelastic operator},$$

defined by

$$\boldsymbol{\varphi}{\downarrow}\mathbf{g} = \mathbf{g}\,(d\boldsymbol{\varphi}^T d\boldsymbol{\varphi}) , \qquad \mathbf{g}_\mathbf{a} = \mathbf{g}\,\mathbf{G}_\mathbf{a} ,$$

Different operator representations of the metric tensors may be obtained by choosing a metric tensor other than $\mathbf{g}$ *. The choice of the two metrics* $\mathbf{g}_\mathbf{a}$ *and* $\boldsymbol{\varphi}{\downarrow}\mathbf{g}$ *leads to the representation formulae:*

$$\boldsymbol{\varphi}{\downarrow}\mathbf{g} = \mathbf{g}_\mathbf{a}\,\mathbf{E} , \qquad \mathbf{g}_\mathbf{a} = (\boldsymbol{\varphi}{\downarrow}\mathbf{g})\,\mathbf{E}^{-1} ,$$

with $d\boldsymbol{\varphi}^T d\boldsymbol{\varphi} = \mathbf{G}_\mathbf{a}\,\mathbf{E}$ *.*

It is worth noting that the definition of the tensor $\mathbf{E}$ *could suggest a kind of chain decomposition of the Piola-Green operator into an elastic and a plastic part with the elastic operator* $\mathbf{E}$ *acting before the plastic one* $\mathbf{G}_\mathbf{a}$ *. Anyway these representations may not be convenient since the metrics* $\mathbf{g}_\mathbf{a}$ *and* $\boldsymbol{\varphi}{\downarrow}\mathbf{g}$ *are time dependent in an evolutive process.*

5.3 Multi-phase materials

To describe the evolution of phase transition phenomena in multi-phase material bodies, we consider a partition of the natural placement $\mathbb{B}$ of the body into a finite family $\mathcal{T}(\mathbb{B})$ of non-overlapping submanifolds. Each element of the partition $\mathcal{T}(\mathbb{B})$ is constituted by a single-phase material Accordingly, the elastic energy density of the multi-phase material at a particple $\mathbf{p} \in \mathbb{B}$ is given by

$$\hat{W}(\mathbf{p}) = W((\boldsymbol{\varphi}{\downarrow}\mathbf{g})(\mathbf{p}), \mathbf{g}_\mathbf{a}(\mathbf{p}), \mathbf{p}) .$$

Phase-transition phenomena are described by a flow $\boldsymbol{\chi}_{\tau,t} \in \mathrm{C}^1(\mathbb{B} ; \mathbb{B})$ which modifies the reference partition $\mathcal{T}(\mathbb{B})$ into an evolving one $\boldsymbol{\chi}_{\tau,t}(\mathcal{T}(\mathbb{B}))$ at time $\tau \in I$.

5.4 Material symmetry

Material symmetry at a point of the natural placement of a body is a measure of the indifference, of the material response due to a change of placement, to a linear pre-transformation of the tangent space before that the change of placement takes place.

For an elastic material behavior this property depends on whether the value of elastic potential be modified or not by a linear pre-transformation of the tangent space. In the next section we discuss some basic properties of the set of linear pre-transformation fulfilling the symmetry property for an elastic material behavior.

Symmetry Groups

Let us preliminarily give a useful definition. If $\mathbb{B}$ is a natural placement of a material body and $\mathbf{Q}, \mathbf{R} \in BL(\mathbb{T}_{\mathbf{p}}\mathbb{B}\,; \mathbb{T}_{\mathbf{p}}\mathbb{B})$ are linear isomorphisms and $\mathbf{g} \in BL(\mathbb{T}_{\mathbf{p}}\mathbb{B}^2\,; \Re)$ any tensor at $\mathbf{p} \in \mathbb{B}$, we define define the tensor $\mathbf{Qg} \in BL(\mathbb{T}_{\mathbf{p}}\mathbb{B}^2\,; \Re)$ by the identity

$$(\mathbf{Qg})(\mathbf{a}, \mathbf{b}) := \mathbf{g}(\mathbf{Q}^{-1}\mathbf{a}, \mathbf{Q}^{-1}\mathbf{b})\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{T}_{\mathbf{p}}\mathbb{B}\,,$$

The operation $\mathbf{Qg}$, which is reminiscent of the push forward operation, meets the property:

$$\mathbf{Q}(\mathbf{Rg}) = (\mathbf{RQ})\mathbf{g}\,.$$

The symmetry group G of the elastic body at $\mathbf{p} \in \mathbb{B}$ is then defined as the set of linear isomorphisms $\mathbf{R} \in BL(\mathbb{T}_{\mathbf{p}}\mathbb{B}\,; \mathbb{T}_{\mathbf{p}}\mathbb{B})$ such that

$$W(\boldsymbol{\varphi}{\downarrow}(\mathbf{Rg}) - \mathbf{g}) = W(\boldsymbol{\varphi}{\downarrow}\mathbf{g} - \mathbf{g})\,, \qquad \forall\, \boldsymbol{\varphi}(\mathbf{p}) \in BL(\mathbb{B}\,; \mathbb{S})\,.$$

Apparently $\mathbf{I}, -\mathbf{I} \in G$ and

$$\mathbf{Q}, \mathbf{R} \in G \implies \mathbf{QR}, \mathbf{RQ} \in G\,.$$

Indeed

$$\begin{aligned} W(\boldsymbol{\varphi}{\downarrow}((\mathbf{QR})\mathbf{g}) - \mathbf{g}) &= W(\boldsymbol{\varphi}{\downarrow}\mathbf{R}(\mathbf{Qg}) - \mathbf{g}) \\ &= W(\boldsymbol{\varphi}{\downarrow}(\mathbf{Rg}) - \mathbf{g}) \\ &= W(\boldsymbol{\varphi}{\downarrow}\mathbf{g} - \mathbf{g})\,, \end{aligned}$$

which proves also that

$$\mathbf{Q} \in G\,, \quad \mathbf{QR} \in G \implies \mathbf{R} \in G\,.$$

Hence $\mathbf{R} \in G$ and $\mathbf{R}^{-1}\mathbf{R} = \mathbf{I} \in G$ imply that $\mathbf{R}^{-1} \in G$.

We may then state that:

- the symmetries of an elastic material form a subgroup G of the algebra of linear isomorphisms, under the operation of composition. The group G includes the opposite of each of its elements.
- If the symmetry group is the whole unimodular group (that is the subgroup of isochoric isomorphisms) the material is an *elastic fluid*.
- If the symmetry group is included in the orthogonal group (that is the subgroup of isometric isomorphisms) the material is an *elastic solid*.
- If the symmetry group includes the whole orthogonal group the material is said to be *isotropic*, otherwise *aelotropic* (or *anisotropic*).

It can be shown that an elastic material, as defined above, can be either a fluid or a solid. The mathematical statement consists in the property that the orthogonal group is a maximal subgroup of the unimodular group [16], [131].

- The elastic energy density of an elastically isotropic material can be expressed in terms of the principal invariants of the $\mathbf{g}$-symmetric elastic deformation tensor.

5.4.1 Thermally isotropic materials

In an analogous way, we may define the symmetry group $G_{\mathbb{B}}$ of an anelastic constitutive relation. For example sake, let the thermal deformation of a material be described by the linear incremental law

$$\dot{\mathbf{G}}_{\mathbf{a}} = \mathbf{A}_\theta \, \dot{\theta} \,,$$

where $\mathbf{A}_\theta \in BL(\mathbb{T}_{\mathbf{p}}\mathbb{B}\,; \mathbb{T}_{\mathbf{p}}\mathbb{B})$ is the $\mathbf{g}$-symmetric operator of thermal expansion and $\dot{\theta}$ is the temperature rate of variation. The metric tensor $\mathbf{G}_{\mathbf{a}}$ is uniquely defined by requiring that, when applied to perform length measurements of the edges of a non-degenerated symplex at a point in the reference configuration, it provides the length of the edges of the thermally deformed symplex. An isotropic thermal expansion requires that

$$\mathbf{R}\,\mathbf{A}_\theta \mathbf{R}^T = \mathbf{A}_\theta \,, \quad \forall\, \mathbf{R} \in \mathrm{SO}(3) \,.$$

Any eigenvector $\mathbf{e} \in \mathbb{T}_{\mathbf{p}}\mathbb{B}$ of $\mathbf{A}_\theta$ transforms into an eigenvector $\mathbf{Re} \in \mathbb{T}_{\mathbf{p}}\mathbb{B}$ of $\mathbf{R}\,\mathbf{A}_\theta\mathbf{R}^T = \mathbf{A}_\theta$ with the same eigenvalue. Hence any nonzero vector is an eigenvector, that is

$$\mathbf{A}_\theta(\mathbf{e}) = \alpha_\theta\,\mathbf{e}\,, \quad \forall\,\mathbf{e} \in \mathbb{T}_{\mathbf{p}}\mathbb{B}\,.$$

so that $\mathbf{A}_\theta = \alpha_\theta\,\mathbf{I}$ and hence $\mathbf{G}_{\mathbf{a}} = \alpha_\theta\dot{\theta}\,\mathbf{I}$ where $\alpha_\theta \in \Re$ is the, temperature dependent, thermal expansion coefficient of the thermally isotropic material.

5.5 Elastic energy rate due to phase transition

To provide a mathematical formulation of the dissipation phenomena due to phase transition, we consider a virtual motion of the body in the ambient space $\mathbb{S}$ described by a flow $\boldsymbol{\psi}_{\tau,t} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{S})$ starting at the current configuration $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{S})$ at time $t \in I$.

The time dependence of the free energy density is expressed by

$$W_\tau := W((\boldsymbol{\psi}_{\tau,t} \circ \boldsymbol{\varphi}){\downarrow}\mathbf{g}\,, \mathbf{g}_{\mathbf{a}\tau})\,,$$

and the free energy of the body at time $\tau \in I$ is

$$\mathbf{E}_\tau := \int_{\mathbb{B}} W_\tau\,\boldsymbol{\mu}\,.$$

Let us now evaluate the time-rate of the free energy of the body.

In this respect it is important to notice that the time derivative of the free energy density W_τ cannot be performed in a classical way since the configuration-induced metric $(\boldsymbol{\psi}_{\tau,t} \circ \boldsymbol{\varphi}){\downarrow}\mathbf{g}$ and the phase-describing field $p_t \circ \boldsymbol{\chi}_{t,\tau}$ undergo a jump at the points $\mathbf{p} \in \mathbb{B}$ which are crossed by the evolving interfaces at time $\tau \in I$.

The corresponding DIRAC's impulses at the interfaces may be conveniently evaluated by adopting the following procedure.

By additivity, the integral over $\mathbb{B}$ is written as the sum of integrals over the elements $\mathcal{P}_\tau$ of the partition $\mathcal{T}_\tau(\mathbb{B}) = \boldsymbol{\chi}_{\tau,t}(\mathcal{T}(\mathbb{B}))$ at time $\tau \in I$:

$$\mathbf{E}_\tau = \int_{\mathcal{T}_\tau(\mathbb{B})} W_\tau\,\boldsymbol{\mu} := \sum \int_{\mathcal{P}_\tau} W_\tau\,\boldsymbol{\mu}\,.$$

Then the time derivative is evaluated by making recourse to the transport formula:

$$\partial_{\tau=t} \int_{\mathcal{P}_\tau} W_\tau\,\boldsymbol{\mu} = \int_{\mathcal{P}} \dot{W}\boldsymbol{\mu} + \int_{\mathcal{P}} \mathcal{L}_{\dot{\boldsymbol{\chi}}}(W\boldsymbol{\mu})\,.$$

where $\mathcal{L}_{\dot{\boldsymbol{\chi}}}(W\boldsymbol{\mu})$ is the Lie derivative of the free energy volume-form $W\boldsymbol{\mu}$ along the phase-transition describing flow $\boldsymbol{\chi}_{\tau,t} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{B})$, starting at time $t \in I$ with propagation speed $\dot{\boldsymbol{\chi}} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{TM})$.

By formula $v)$ in proposition 1.3.11

$$\mathcal{L}_{\dot{\boldsymbol{\chi}}}(W\boldsymbol{\mu}) = \mathcal{L}_{(W\dot{\boldsymbol{\chi}})}\boldsymbol{\mu} = \mathrm{div}\,(W\dot{\boldsymbol{\chi}})\boldsymbol{\mu}\,,$$

and the divergence theorem, we get the expression

$$\begin{aligned}\dot{\mathbf{E}} &= \int_{\mathcal{T}(\mathbb{B})} \dot{W}\boldsymbol{\mu} + \int_{\mathcal{T}(\mathbb{B})} \mathcal{L}_{\dot{\boldsymbol{\chi}}}(W\boldsymbol{\mu}) \\ &= \int_{\mathcal{T}(\mathbb{B})} \dot{W}\boldsymbol{\mu} + \int_{\mathcal{T}(\mathbb{B})} \mathcal{L}_{(W\dot{\boldsymbol{\chi}})}\,\boldsymbol{\mu} \\ &= \int_{\mathcal{T}(\mathbb{B})} \dot{W}\boldsymbol{\mu} + \int_{\mathcal{T}(\mathbb{B})} \mathrm{div}\,(W\dot{\boldsymbol{\chi}})\,\boldsymbol{\mu} \\ &= \int_{\mathcal{T}(\mathbb{B})} \dot{W}\boldsymbol{\mu} + \int_{\partial\mathcal{T}(\mathbb{B})} W\mathbf{g}(\dot{\boldsymbol{\chi}},\mathbf{n})\,(\boldsymbol{\mu}\mathbf{n})\,,\end{aligned}$$

where $\boldsymbol{\mu}\mathbf{n}$ is the area-form induced on the surfaces $\partial\mathcal{T}(\mathbb{B})$ by the volume form $\boldsymbol{\mu}$ in $\mathbb{B}$. Since the flow $\boldsymbol{\chi}_{\tau,t} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{B})$ leaves the boundary $\partial\mathbb{B}$ invariant, we have that $\mathbf{g}(\dot{\boldsymbol{\chi}},\mathbf{n}) = 0$ on $\partial\mathbb{B}$.

Then, defining the jump $[[W]] = W^+ - W^-$ across the phase-transition interfaces and setting $\mathbf{n} = \mathbf{n}^-$, the outward normal to $\partial\mathcal{P}^-$, we get the final result:

$$\dot{\mathbf{E}} = \int_{\mathcal{T}(\mathbb{B})} \dot{W}\,\boldsymbol{\mu} - \int_{\mathcal{I}} [[W]]\,v_{\boldsymbol{\chi}}\,(\boldsymbol{\mu}\mathbf{n})\,,$$

where $\mathcal{I}$ is the set of phase-transition interfaces travelling with normal speed $v_{\boldsymbol{\chi}} = \mathbf{g}(\dot{\boldsymbol{\chi}},\mathbf{n})$.

Since the normal speed points towards the $\mathcal{P}^+$ phase, the impulsive term, provided by the integral over the interfaces, measures the rate of decrease of the free energy due to the motion of phase-transition fronts.

5.5.1 Dissipation due to phase transition

Phase-transition phenomena are characterized by the continuity of the configuration map with a possible finite jump of its differential across the transition fronts. These singular surfaces are shock-waves and their propagation requires a dissipation of energy.

Phase-transition phenomena are dealt with by relying upon the theory of singular surfaces travelling in the material, in which MAXWELL's jump condition and HADAMARD's condition for shock waves are the main analytical tools.

Kinematics of shock waves

To deal with discontinuity surfaces travelling in the material body, we consider the general case in which the configuration map $\boldsymbol{\varphi} \in \mathrm{C}^0(\mathbb{B}\,;\mathbb{S}) \cap \mathrm{C}^1(\mathrm{PAT}(\mathbb{B})\,;\mathbb{S})$ is continuous on $\mathbb{B}$ and continuously differentiable in each element of the partition $\mathrm{PAT}(\mathbb{B})$ whose interfaces may travel in the material according to a flow $\boldsymbol{\chi}_{\tau,t} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{B})$. By continuity, the derivatives of $\boldsymbol{\varphi}$ along tangent directions on each side of the interfaces $\mathcal{I}$ of $\mathcal{T}(\mathbb{B})$ are equal:

$$d_{\mathbf{t}}\, \boldsymbol{\varphi}^{+}(\mathbf{p}) = d_{\mathbf{t}}\, \boldsymbol{\varphi}^{-}(\mathbf{p})\,, \quad \forall\, \mathbf{t} \in \mathbb{T}_{\mathbf{p}}\mathcal{I}\,.$$

It follows that the differential $d\boldsymbol{\varphi}(\mathbf{p}) \in BL\,(\mathbb{T}_{\mathbf{p}}\mathbb{B}\,;\mathbb{T}_{\boldsymbol{\varphi}(\mathbf{p})}\mathbb{S})$ must meet at the interfaces MAXWELL*'s jump condition*:

$$[[d\boldsymbol{\varphi}]] = [[d\boldsymbol{\varphi}]]\,\mathbf{n} \otimes \mathbf{n}\,.$$

Then, across a shock wave front, the configuration map is continuous and its differential may undergo a finite jump.

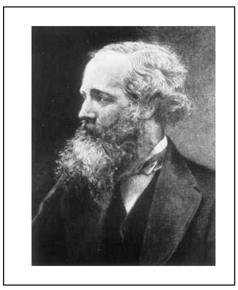

Figure 5.1: James Clerk Maxwell (1831 - 1879)

The spatial speed $(\boldsymbol{\varphi} \circ \boldsymbol{\chi})^{\cdot}$ of the points travelling on the shock wave, propagating in the material with speed $\dot{\boldsymbol{\chi}} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{TM})$, may be evaluated, by the Leibniz rule, on each side of the shock wave to get:

$$(\boldsymbol{\varphi} \circ \boldsymbol{\chi})^{\cdot} = \dot{\boldsymbol{\varphi}} + d_{\dot{\boldsymbol{\chi}}}\boldsymbol{\varphi} = \mathbf{v} \circ \boldsymbol{\varphi} + d_{\dot{\boldsymbol{\chi}}}\boldsymbol{\varphi}\,.$$

Since the l.h.s. is continuous across the interface, the following jump condition must be met:

$$[[\mathbf{v}]] \circ \boldsymbol{\varphi} + [[d\boldsymbol{\varphi}]]\, \dot{\boldsymbol{\chi}} = 0\,.$$

From MAXWELL's jump condition we then get HADAMARD*'s condition* for shock waves:

$$[[\mathbf{v}]] \circ \boldsymbol{\varphi} + v_{\boldsymbol{\chi}}\, [[d\boldsymbol{\varphi}]]\, \mathbf{n} = 0\,.$$

This condition tells us that the velocity field will undergo, across the shock wave front, a finite jump equal to the opposite of the finite jump of the normal derivative of the configuration map times the normal speed of propagation of the shock wave.

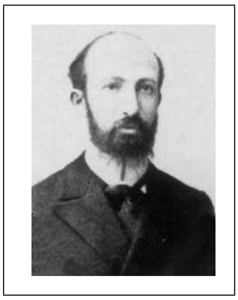

Figure 5.2: Jacques Salomon Hadamard (1865 - 1963)

As shown below, HADAMARD's condition plays a basic role in the evaluation of the dissipation induced by evolving phase transition interfaces.

Evolution problem

The equilibrium of the body at the current configuration is expressed by the virtual work condition, which, in the reference placement, is written as

$$\langle \mathbf{f}, \mathbf{v} \rangle = \int_{\mathcal{T}_{(\mathbf{v}\circ\boldsymbol{\varphi})}(\mathbb{B})} \langle \mathbf{S}, \mathbf{D}(\boldsymbol{\varphi})^{\cdot} \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,, \quad \forall\ \mathbf{v} \in \mathrm{KIN}$$

where $\mathbf{v} = \dot{\boldsymbol{\psi}} \in \mathrm{KIN}$ is the initial speed along the virtual trajectory described by the flow $\boldsymbol{\psi}_{\tau,t} \in \mathrm{C}^1(\boldsymbol{\varphi}(\mathbb{B})\,;\mathbb{S})$ and $\mathbf{D}(\boldsymbol{\varphi})^{\cdot} := \partial_{\tau=t}\,(\mathbf{D}(\boldsymbol{\psi}_{\tau,t} \circ \boldsymbol{\varphi})\,.$

Let us now assume that the virtual speed $\mathbf{v} \in \mathrm{KIN}$ be compatible with the normal speed of the phase-transition interfaces travelling according to the flow

$\boldsymbol{\chi}_{\tau,t} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{B})$. This means that the phase-describing partition $\mathcal{T}(\mathbb{B})$ is a regularity patchwork for the virtual speed $\mathbf{v} \in \mathrm{C}^1(\mathcal{T}(\mathbb{B})\,;\mathbb{S})$ and that it fulfils the HADAMARD condition for shock waves:

$$[[\mathbf{v}]] \circ \boldsymbol{\varphi} + v_{\boldsymbol{\chi}}\, [[d\boldsymbol{\varphi}]]\, \mathbf{n} = 0\,,$$

at the interfaces $\mathcal{I}$ of the partition $\mathcal{T}(\mathbb{B})$. Let us express the free energy in terms of operators $W(\mathbf{D}(\boldsymbol{\varphi}), \boldsymbol{\Delta}, p)$, with $\mathbf{g}\boldsymbol{\Delta} = \mathbf{g}_{\mathbf{a}}$. The equilibrium condition is then obtained by imposing the constitutive requirement $\mathbf{S} = d_1 W$ and setting $\mathbf{S}_{\mathbb{B}} := -d_2 W$. The time derivative in each element $\mathcal{P}$ of $\mathcal{T}(\mathbb{B})$ is given by

$$\begin{aligned}(W(\mathbf{D}(\boldsymbol{\varphi}), \boldsymbol{\Delta}, p))^{\cdot} &:= \partial_{\tau=t}(W(\mathbf{D}(\boldsymbol{\psi}_{\tau,t} \circ \boldsymbol{\varphi}), \boldsymbol{\Delta}_\tau, p_\tau)) \\ &= \langle\, d_1 W, \mathbf{D}(\boldsymbol{\varphi})^{\cdot} \rangle_{\mathbf{g}} + \langle\, d_2 W, \dot{\boldsymbol{\Delta}}\, \rangle_{\mathbf{g}} \\ &= \langle\, \mathbf{S}, \mathbf{D}(\boldsymbol{\varphi})^{\cdot} \rangle_{\mathbf{g}} - \langle\, \mathbf{S}_{\mathbb{B}}, \dot{\boldsymbol{\Delta}}\, \rangle_{\mathbf{g}}\,,\end{aligned}$$

since $\dot{p} = 0$ due to the constancy of p_τ in each $\mathcal{P}_\tau$ at any time $\tau \in I$.

The equilibrium condition may then be written as

$$\begin{aligned}\langle\, \mathbf{f}, \mathbf{v}\, \rangle &= \int_{\mathcal{T}(\mathbb{B})} \langle\, \mathbf{S}, \mathbf{D}(\boldsymbol{\varphi})^{\cdot}\, \rangle_{\mathbf{g}}\, \boldsymbol{\mu} \\ &= \int_{\mathcal{T}(\mathbb{B})} \dot{W}\, \boldsymbol{\mu} + \int_{\mathbb{B}} \langle\, \mathbf{S}_{\mathbb{B}}, \dot{\boldsymbol{\Delta}}\, \rangle_{\mathbf{g}}\, \boldsymbol{\mu} \\ &= \dot{\mathbf{E}} + \int_{\mathcal{I}} [[W]]\, v_{\boldsymbol{\chi}}\, (\boldsymbol{\mu}\mathbf{n}) + \int_{\mathbb{B}} \langle\, \mathbf{S}_{\mathbb{B}}, \dot{\boldsymbol{\Delta}}\, \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,.\end{aligned}$$

The virtual work of the force system acting on the body can be split into the sum of two contributions.

The former is the virtual work performed by the loading $\ell \in$ LOAD in correspondence of the virtual velocity $\mathbf{v} \circ \boldsymbol{\varphi} \in \mathrm{C}^1(\mathrm{PAT}(\mathbb{B})\,;\mathbb{S})$. The latter is the virtual work performed by the reactive forces $\mathbf{r}$ acting on the faces of each phase-transition interface due to the finite jump of the virtual velocity across the phase-transition interfaces:

$$\langle\, \mathbf{f}, \mathbf{v}\, \rangle = \langle\, \ell, \mathbf{v}\, \rangle + \langle\, \mathbf{r}, \mathbf{v}\, \rangle\,.$$

The boundary equilibrium condition at the interfaces implies that $[[\mathbf{P}\mathbf{n}]] = 0$, and hence the reactive term is given by

$$\langle\, \mathbf{r}, \mathbf{v}\, \rangle = -\int_{\mathcal{I}} [[\langle\, \mathbf{P}\mathbf{n}, \mathbf{v} \circ \boldsymbol{\varphi} \rangle]]_{\mathbf{g}}\, (\boldsymbol{\mu}\mathbf{n}) = -\int_{\mathcal{I}} \langle\, \mathbf{P}\mathbf{n}, [[\mathbf{v}]] \circ \boldsymbol{\varphi}\, \rangle_{\mathbf{g}}\, (\boldsymbol{\mu}\mathbf{n})\,.$$

The minus sign above is due to the usual notation $[[\mathbf{v}]] = \mathbf{v}^+ - \mathbf{v}^-$ with $\mathbf{n} = \mathbf{n}^-$ the outward normal to $\partial\mathcal{P}^-$. The equilibrium condition may then be written as

$$\langle \ell, \mathbf{v} \rangle = \dot{\mathbf{E}} + \int_{\mathcal{I}} [[W]]\, v_{\boldsymbol{\chi}}\, (\boldsymbol{\mu}\mathbf{n}) + \int_{\mathcal{I}} \langle \mathbf{Pn}, [[\mathbf{v}]] \circ \boldsymbol{\varphi} \rangle_{\mathbf{g}}\, (\boldsymbol{\mu}\mathbf{n}) + \int_{\mathbb{B}} \langle \mathbf{S}, \dot{\boldsymbol{\Delta}} \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,.$$

Imposing the fulfilment of HADAMARD's condition for shock waves at the interfaces $\mathcal{I}$ of phase-transition:

$$[[\mathbf{v}]] \circ \boldsymbol{\varphi} + v_{\boldsymbol{\chi}}\, [[d\boldsymbol{\varphi}]]\, \mathbf{n} = 0\,.$$

we get the following formula for the virtual power balance law:

$$\langle \ell, \mathbf{v} \rangle = \dot{\mathbf{E}} + \int_{\mathcal{I}} ([[W]] - \mathbf{g}(\mathbf{Pn}, [[d\boldsymbol{\varphi}]]\, \mathbf{n}))\, v_{\boldsymbol{\chi}}\, (\boldsymbol{\mu}\mathbf{n}) + \int_{\mathbb{B}} \langle \mathbf{S}_{\mathbb{B}}, \dot{\boldsymbol{\Delta}} \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,,$$

to hold for all spatial speed $\mathbf{v} \in \mathrm{C}^1(\mathrm{PAT}(\mathbb{B})\,; \mathbb{S})$ and for all phase-transition speed $\dot{\boldsymbol{\chi}} \in \mathrm{C}^1(\mathbb{B}\,; \mathbb{TM})$ fulfilling HADAMARD's condition.

Now, observing that $W = \mathbf{g}(W\mathbf{n}, \mathbf{n})$, we introduce ESHELBY's tensor:

$$\mathbf{Y} := W\mathbf{I} - d\boldsymbol{\varphi}^T \mathbf{P} = W\mathbf{I} - d\boldsymbol{\varphi}^T d\boldsymbol{\varphi}\, \mathbf{S}\,,$$

and write the virtual power balance law as

$$\langle \ell, \mathbf{v} \rangle = \dot{\mathbf{E}} + \int_{\mathcal{I}} \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{n})\, v_{\boldsymbol{\chi}}\, (\boldsymbol{\mu}\mathbf{n}) + \int_{\mathbb{B}} \langle \mathbf{S}_{\mathbb{B}}, \dot{\boldsymbol{\Delta}} \rangle \boldsymbol{\mu}\,.$$

Then from the properties

$$\begin{aligned}
&\mathbf{g}(\mathbf{n}, \mathbf{t}) = 0 \implies \mathbf{g}([[W]]\, \mathbf{n}, \mathbf{t}) = [[W]]\, \mathbf{g}(\mathbf{n}, \mathbf{t}) = 0\,,\\
&\mathbf{g}(\mathbf{n}, \mathbf{t}) = 0 \implies \mathbf{g}([[d\boldsymbol{\varphi}^T]]\, \mathbf{Pn}, \mathbf{t}) = \mathbf{g}(\mathbf{Pn}, [[d\boldsymbol{\varphi}]]\, \mathbf{t})\\
&\qquad\qquad\qquad = \mathbf{g}(\mathbf{Pn}, [[d\boldsymbol{\varphi}]]\, (\mathbf{n} \otimes \mathbf{n})\, \mathbf{t}) = \mathbf{g}(\mathbf{Pn}, [[d\boldsymbol{\varphi}]]\mathbf{n})\, \mathbf{g}(\mathbf{n}, \mathbf{t}) = 0\,,
\end{aligned}$$

the latter being a consequence of MAXWELL's jump condition, we infer that

$$\mathbf{g}(\mathbf{n}, \mathbf{t}) = 0 \implies \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{t}) = 0\,,$$

that is, the flux of the jump of ESHELBY's tensor at an interface is directed along the normal to the interface.

Hence, being $v_{\boldsymbol{\chi}} = \mathbf{g}(\dot{\boldsymbol{\chi}}, \mathbf{n})$, we infer the equality

$$\mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{n})\, v_{\boldsymbol{\chi}} = \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \dot{\boldsymbol{\chi}})\,,$$

and the virtual power balance law may be rewritten as

$$\langle \ell, \mathbf{v} \rangle = \dot{\mathbf{E}} + \int_{\mathcal{I}} \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \dot{\boldsymbol{\chi}})\ (\boldsymbol{\mu}\mathbf{n}) + \int_{\mathbb{B}} \langle \mathbf{S}_{\mathbb{B}}, \dot{\boldsymbol{\Delta}} \rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,.$$

This result may be phrased by stating that the (virtual) power performed by the applied load is equal to the (virtual) increase in free energy plus the (virtual) dissipation due to the evolution of phase transition and to the anelastic deformation rate.

In the actual motion, we get a mechanical statement of the principle of conservation of the power expended.

Remark 5.5.1 ESHELBY*'s tensor* $\mathbf{Y} = W\mathbf{I} - d\boldsymbol{\varphi}^T d\boldsymbol{\varphi}\,\mathbf{S}$ *is not* $\mathbf{g}$*-symmetric, but symmetry holds with respect to the metric* $(d\boldsymbol{\varphi}^T \mathbf{g})(\mathbf{p}) \in BL(\mathbb{T}_{\mathbf{p}}\mathbb{B}, \mathbb{T}_{\mathbf{p}}\mathbb{B}\,; \Re)$ *defined at* $\mathbf{p} \in \mathbb{B}$ *by*

$$(d\boldsymbol{\varphi}^T \mathbf{g})(\mathbf{a}, \mathbf{b}) := \mathbf{g}(d\boldsymbol{\varphi}^{-T}\mathbf{a}, d\boldsymbol{\varphi}^{-T}\mathbf{b})\,, \quad \forall\, \mathbf{a}, \mathbf{b} \in \mathbb{T}_{\mathbf{p}}(\mathbb{B})\,.$$

This property is a direct consequence of the $\mathbf{g}$*-symmetry of the* PIOLA-KIRCHHOFF *stress tensor* $\mathbf{S}$ *since*

$$\begin{aligned}(d\boldsymbol{\varphi}^T \mathbf{g})(d\boldsymbol{\varphi}^T d\boldsymbol{\varphi}\,\mathbf{S}\,\mathbf{a}, \mathbf{b}) &= \mathbf{g}(d\boldsymbol{\varphi}^{-T} d\boldsymbol{\varphi}^T d\boldsymbol{\varphi}\,\mathbf{S}\,\mathbf{a}, d\boldsymbol{\varphi}^{-T}\mathbf{b}) \\ &= \mathbf{g}(d\boldsymbol{\varphi}\,\mathbf{S}\,\mathbf{a}, d\boldsymbol{\varphi}^{-T}\mathbf{b}) = \mathbf{g}(d\boldsymbol{\varphi}^{-1} d\boldsymbol{\varphi}\,\mathbf{S}\,\mathbf{a}, \mathbf{b}) = \mathbf{g}(\,\mathbf{S}\,\mathbf{a}, \mathbf{b})\,.\end{aligned}$$

ESHELBY*'s tensor is then symmetrizable and enjoys all the useful properties of a symmetric operator.*

It has a spectral representation with real eigenvalues since there exists in $\mathbb{T}_{\mathbf{p}}\mathbb{B}$ *a principal basis of mutually orthogonal eigenvectors according to the metric* $(d\boldsymbol{\varphi}^T \mathbf{g})(\mathbf{p})$*. Setting* $\mathbf{C} = d\boldsymbol{\varphi}^T d\boldsymbol{\varphi}$ *the symmetry of* ESHELBY*'s tensor can be written as* $\mathbf{YC} = \mathbf{CY}^T$*, a result quoted in [41].*

Remark 5.5.2 *The previous expression of the virtual power balance law is based on the analysis developed by* MORTON GURTIN *in discussing the role of what he calls configurational forces (see [70], formula 1-6).* GURTIN*'s formula is derived under the assumption of fixed kinematic boundary conditions, and vanishing body forces and anelastic deformation rate so that* $\langle \ell, \mathbf{v} \rangle = 0$ *and* $\dot{\boldsymbol{\Delta}} = 0$*. In our notations, his formula reads*

$$-\dot{\mathbf{E}} = -\int_{\mathcal{T}(\mathbb{B})} \langle \mathbf{S}, \mathbf{D}(\boldsymbol{\varphi})^{\cdot} \rangle_{\mathbf{g}}\, \boldsymbol{\mu} + \int_{\mathcal{I}} [[W]]\, \mathbf{g}(\dot{\boldsymbol{\chi}}, \mathbf{n})\, (\boldsymbol{\mu}\mathbf{n})\,,$$

to hold for all $\mathbf{v} \in \mathrm{C}^1(\mathcal{T}(\mathbb{B})\,;\mathbb{S})$ *and* $\dot{\boldsymbol{\chi}} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{TM})$ *fulfilling* HADAMARD*'s condition for shock waves on* $\mathcal{I}$ *. This is equivalent to*

$$-\dot{\mathbf{E}} = \int_{\mathcal{I}} \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \dot{\boldsymbol{\chi}})\ (\boldsymbol{\mu}\mathbf{n}) = \int_{\mathcal{I}} \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{n})\, v_{\boldsymbol{\chi}}\, (\boldsymbol{\mu}\mathbf{n})\,.$$

He then assumes that $\dot{\mathbf{E}} = 0$ *for all* $v_{\boldsymbol{\chi}}$ *concluding that* $\mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{n}) = 0$ *, a condition which he claims to be often referred to as* MAXWELL*'s relation (but it has in fact no connection with* MAXWELL*'s jump condition illustrated above). From the property* $\mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{t}) = 0$ *for all* $\mathbf{t}$ *such that* $\mathbf{g}(\mathbf{n}, \mathbf{t}) = 0$ *, he then concludes that* $[[\mathbf{Y}]]\mathbf{n} = 0$ *at phase-transition interfaces.*

We must confess to be unable to find a physical motivation for GURTIN*'s assumption that* $\dot{\mathbf{E}} = 0$ *for all* $v_{\boldsymbol{\chi}}$ *.*

As a consequence, his conclusion that $\mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{n}) = 0$ *and* $[[\mathbf{Y}]]\mathbf{n} = 0$ *cannot be agreed on, since it implies that the evolution of the phase-transition interfaces requires no power to be expended, despite of experimental evidences in solid state physics and fracture mechanics.*

Reasoning in the opposite direction, we are led to conclude that the singular term

$$\int_{\mathcal{I}} \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \dot{\boldsymbol{\chi}})\ (\boldsymbol{\mu}\mathbf{n})\,, \qquad \dot{\boldsymbol{\chi}} \in \mathrm{C}^1(\mathbb{B}\,;\mathbb{TM})\,,$$

provides the (virtual) power dissipated in the motion of the evolving phase-transition interfaces.

Small displacement formulation

Many engineering applications can be dealt with by a geometrically linearized formulation. To provide the specialization of the previous theory to this important class of problems, it is convenient to re-formulate the analysis in terms of the displacement field $\mathbf{u} \in \mathrm{C}^0(\mathbb{B}\,;\mathbb{TS}) \cap \mathrm{C}^1(\mathcal{T}(\mathbb{B})\,;\mathbb{TS})$ defined by

$$\mathbf{u}(\mathbf{p}) = \boldsymbol{\varphi}(\mathbf{p}) - \mathbf{p}\,,$$

so that $d\mathbf{u} = d\boldsymbol{\varphi} - \mathbf{I}$ in $\mathcal{T}(\mathbb{B})$. For the jump across the phase-transition interfaces $\mathcal{I}$ we have the equality $[[d\mathbf{u}]] = [[d\boldsymbol{\varphi}]]$ and hence the ESHELBY's tensor can be equivalently defined in terms of displacement field as

$$\mathbf{Y}_{\mathbf{u}} := W\mathbf{I} - d\mathbf{u}^T\mathbf{P} = W\mathbf{I} - d\boldsymbol{\varphi}^T\mathbf{P} + \mathbf{P} = \mathbf{Y} + \mathbf{P}\,,$$

with $[[\mathbf{Y}_{\mathbf{u}}]]\mathbf{n} = [[\mathbf{Y}]]\mathbf{n}$ since $[[\mathbf{P}\mathbf{n}]] = 0$.

In the geometrically linearized theory, the reference and the actual placements of the body are taken to be coincident so that the Piola stress $\mathbf{P}$ and the Cauchy stress $\mathbf{T}$ may be identified.

Accordingly ESHELBY's tensor takes the form

$$\mathbf{Y_u} = W\mathbf{I} - d\mathbf{u}^T\mathbf{T}\,.$$

5.5.2 Divergence of Eshelby's tensor

We provide hereafter the expression of the divergence of Eshelby's tensor in each phase of the multi-phase material, since the vanishing of the divergence is at the basis of the invariance property of the J-integral in fracture mechanics. In each material phase the free energy density is given by

$$\hat{W}(\mathbf{p}) = W(\mathbf{D}(\boldsymbol{\varphi})(\mathbf{p}), \boldsymbol{\Delta}(\mathbf{p}), \mathbf{p})\,.$$

Evaluating the spatial derivative in a direction $\mathbf{h} \in \mathbb{T}_{\mathbf{p}}\mathbb{B}$, by the LEIBNIZ's rule we have that

$$\mathbf{g}(d\hat{W}, \mathbf{h}) = \langle d_1 W, d_{\mathbf{h}}\mathbf{D}(\boldsymbol{\varphi})\rangle_{\mathbf{g}} + \langle d_2 W, d_{\mathbf{h}}\boldsymbol{\Delta}\rangle_{\mathbf{g}} + \mathbf{g}(d_3 W, \mathbf{h})\,,$$

where $d_i W$, $i = 1, 2, 3$ are the partial derivatives. By the formula $d\hat{W} = \operatorname{div}(W\mathbf{I})$ we may write

$$\mathbf{g}(d_3 W, \mathbf{h}) + \langle d_2 W, d_{\mathbf{h}}\boldsymbol{\Delta}\rangle_{\mathbf{g}} = \mathbf{g}(\operatorname{div}(W\mathbf{I}), \mathbf{h}) - \langle d_1 W, d_{\mathbf{h}}\mathbf{D}(\boldsymbol{\varphi})\rangle_{\mathbf{g}}\,,$$

which is the formula prodromic to ESHELBY's one.

In terms of PIOLA's tensor field $\mathbf{P}$ we have that

$$\langle d_1 W, d_{\mathbf{h}}\mathbf{D}(\boldsymbol{\varphi})\rangle_{\mathbf{g}} = \langle \mathbf{S}, d_{\mathbf{h}}\mathbf{D}(\boldsymbol{\varphi})\rangle_{\mathbf{g}} = \langle \mathbf{P}, d_{\mathbf{h}}\, d\boldsymbol{\varphi}\rangle_{\mathbf{g}}\,.$$

and, accordingly, the formula above becomes

$$\mathbf{g}(d_3 W, \mathbf{h}) - \langle \mathbf{S}_{\mathbb{B}}, d_{\mathbf{h}}\boldsymbol{\Delta}\rangle_{\mathbf{g}} = \mathbf{g}(\operatorname{div}(W\mathbf{I}), \mathbf{h}) - \langle \mathbf{P}, d_{\mathbf{h}}\, d\boldsymbol{\varphi}\rangle_{\mathbf{g}}\,.$$

Recalling that the divergence of a field of operators $\mathbf{A} \in \mathrm{C}^1(\mathbb{B}\,; BL(\mathbb{TM}\,; \mathbb{TM}))$ is the vector field $\operatorname{div}\mathbf{A} \in \mathrm{C}^0(\mathbb{B}\,; \mathbb{TM})$ defined by

$$\mathbf{g}(\operatorname{div}\mathbf{A}, \mathbf{v}) := \operatorname{div}(\mathbf{A}^T\mathbf{v}) - \langle \mathbf{A}, d\mathbf{v}\rangle_{\mathbf{g}}\,, \quad \forall\, \mathbf{v} \in \mathrm{C}^1(\mathbb{B}\,; \mathbb{TM})\,,$$

and observing that $d_{\mathbf{h}} d\boldsymbol{\varphi} = d(d_{\mathbf{h}}\boldsymbol{\varphi})$ and setting $\mathbf{A} = \mathbf{P}$ and $\mathbf{v} = d_{\mathbf{h}}\boldsymbol{\varphi}$, we get

$$\langle \mathbf{P}, d_{\mathbf{h}}\, d\boldsymbol{\varphi}\rangle_{\mathbf{g}} = \langle \mathbf{P}, d\, d_{\mathbf{h}}\boldsymbol{\varphi}\rangle_{\mathbf{g}} = \operatorname{div}(\mathbf{P}^T d_{\mathbf{h}}\boldsymbol{\varphi}) - \mathbf{g}(\operatorname{div}\mathbf{P}, d_{\mathbf{h}}\boldsymbol{\varphi})\,.$$

The differential equilibrium condition, under a body force field $\mathbf{b}$, and the divergence formula again, with $\mathbf{A} = d\boldsymbol{\varphi}^T \mathbf{P}$ and $\mathbf{v} = \mathbf{h}$, imply that

$$\operatorname{div} \mathbf{P} = -\mathbf{b}\,, \qquad \operatorname{div}(\mathbf{P}^T d_{\mathbf{h}}\boldsymbol{\varphi}) = \operatorname{div}((\mathbf{P}^T d\boldsymbol{\varphi})\mathbf{h}) = \mathbf{g}(\operatorname{div}(d\boldsymbol{\varphi}^T \mathbf{P}), \mathbf{h})\,.$$

It follows that

$$\mathbf{g}(d_3 W, \mathbf{h}) - \langle \mathbf{S}_{\mathbb{B}}, d_{\mathbf{h}}\boldsymbol{\Delta} \rangle_{\mathbf{g}} = \mathbf{g}(\operatorname{div}(W\mathbf{I} - d\boldsymbol{\varphi}^T \mathbf{P}), \mathbf{h}) - \mathbf{g}(\mathbf{b}, d_{\mathbf{h}}\boldsymbol{\varphi})\,,$$

and, in terms of the Eshelby's operator $\mathbf{Y} := W\mathbf{I} - d\boldsymbol{\varphi}^T \mathbf{P}$, we may write

$$\mathbf{g}(\operatorname{div} \mathbf{Y}, \mathbf{h}) = \mathbf{g}(d_3 W, \mathbf{h}) - \langle \mathbf{S}_{\mathbb{B}}, d_{\mathbf{h}}\boldsymbol{\Delta} \rangle_{\mathbf{g}} + \mathbf{g}(\mathbf{b}, d_{\mathbf{h}}\boldsymbol{\varphi})\,.$$

Hence, in an homogeneous elastic phase, under homogeneous anelastic metric and no body forces, we may conclude that Eshelby's operator is solenoidal, i.e. that $\operatorname{div} \mathbf{Y} = 0$.

5.5.3 Crack propagation

The evaluation of what in fracture mechanics is commonly dubbed the *driving force* on travelling cracks can be based on a suitable specialization of the general expression of the dissipation contributed above. To this end, we consider the motion of a crack travelling in the material.

Assuming that the crack-tip moves with a translational speed $\dot{\boldsymbol{\chi}}(\mathbf{p}) = \dot{\chi}\,\mathbf{d}$ directed along its axis (labeled by the unit vector $\mathbf{d}$), and writing the dissipation as $F\,\dot{\chi}$, the driving force F is given by the relation

$$F = \int_{\mathcal{I}} \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n}) = \int_{\mathcal{I}} \mathbf{g}(\mathbf{Y}^+\mathbf{n}^+ + \mathbf{Y}^-\mathbf{n}^-, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n})\,.$$

where $\mathbf{n}$ is the outward normal to the crack boundary $\mathcal{I}$, oriented from the crack (the minus side) towards the surrounding material (the plus side).

Eshelby's formula for the driving force on translating defects is recovered under the further assumption that the divergence of Eshelby's operator vanishes inside and outside the defect.

Indeed, denoting by Σ any closed surface surrounding the defect, with outward normal $\mathbf{n}$, we have that

$$\int_{\mathcal{I}} (\mathbf{Y}^+\mathbf{n}^+ + \mathbf{Y}^-\mathbf{n}^-)\ (\boldsymbol{\mu}\mathbf{n}) = \int_{\mathcal{I}} \mathbf{Y}^+\mathbf{n}^+\ (\boldsymbol{\mu}\mathbf{n}) = \int_{\Sigma} \mathbf{Y}\mathbf{n}\ (\boldsymbol{\mu}\mathbf{n})\,.$$

which is Eshelby's original formula [47], [49], [109].

As illustrated below, this general result finds application in fracture mechanics for the evaluation of the dissipation associated with non-cohesive and cohesive brittle crack propagation.

J-integral for non-cohesive cracks

Let us consider the motion of a non-cohesive crack travelling in the material. Since there is no material inside the crack, we may assume that there $W^- = 0$ so that $[[W]] = W^+ - W^- = W^+$. Non-cohesive cracks are characterized by the property that the interface between the crack and the surrounding material is traction-free. Denoting by on $\mathcal{I}$ the closed interface bounding the crack nose, that is the terminal crack zone, where $\mathbf{g}(\mathbf{n}, \mathbf{d})$ is non-vanishing, being $\mathbf{P}\mathbf{n} = 0$, we have that

$$[[\mathbf{Y}]]\mathbf{n} = \mathbf{Y}^+\mathbf{n}^+ = \mathbf{Y}^+\mathbf{n}^+ + \mathbf{Y}^-\mathbf{n}^- = W^+\mathbf{n}^+ ,$$

and the driving force takes the expression

$$F = \int_{\mathcal{I}} \mathbf{g}(\mathbf{Y}^+\mathbf{n}^+, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n}) = \int_{\mathcal{I}} \mathbf{g}(W^+\mathbf{n}^+, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n}) ,$$

Following James Rice [145] we consider any closed surface Σ enclosing a region $C(\Sigma)$ which includes the crack-nose.

The J-integral associated with the surface Σ is then defined as:

$$J(\Sigma) := \int_{\Sigma} \mathbf{g}(\mathbf{Y}\mathbf{n}, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n}) ,$$

so that $J(\mathcal{I}) = F$. By the divergence theorem and the formula for $\operatorname{div}\mathbf{Y}$ derived in section 5.5.2, we then get the following general invariance property:

$$\begin{aligned} F &= J(\Sigma) - \int_{C(\Sigma)} \mathbf{g}(\operatorname{div}\mathbf{Y}, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n}) \\ &= J(\Sigma) - \int_{C(\Sigma)} \mathbf{g}(d_3 W, \mathbf{d})\, \boldsymbol{\mu} + \int_{C(\Sigma)} \langle \mathbf{S}_{\mathbb{B}}, d_{\mathbf{d}}\boldsymbol{\Delta} \rangle_{\mathbf{g}}\, \boldsymbol{\mu} - \int_{C(\Sigma)} \mathbf{g}(\mathbf{b}, d_{\mathbf{d}}\boldsymbol{\varphi})\, \boldsymbol{\mu} . \end{aligned}$$

Special instances of this formula are quoted in [94], [95].

In an homogeneous phase, under homogeneous anelastic metric and no body forces, the divergence of Eshelby's tensor field vanishes, i.e. $\operatorname{div}\mathbf{Y} = 0$, and the driving force F is equal to the J-integral evaluated on any surface Σ. In plane problems of fracture mechanics, the invariance property $J(\Sigma) = J(\mathcal{I}) = F$ is commonly referred to as the path independence of the J-integral.

Remark 5.5.3 *In the literature on fracture mechanics (see e.g. [146] III-E), in the wake of* Griffith*'s treatment, crack propagation criteria are discussed* *in terms of an augmented total potential energy of the body which includes a so-called separation energy due to newly created crack faces. This is a nice example of a wrong way to a right result. Not completely right to be honest, since it is correct only if a geometrically linearized modelization is applicable. Indeed, in the nonlinear geometrical range, a total potential energy exists only under conservative loadings and such a requirement is completely extraneous to the physics of the problem at hand. Fortunately what really enters in the analysis is the (pseudo)-time derivative of the augmented total potential energy and this amounts in evaluating a virtual dissipation rate.*

Cohesive cracks

Cohesive cracks are characterized by a process zone, extending ahead the crack-tip, in which cohesive ties oppose the opening of the crack, till the separation of the crack faces reaches a characteristic value that breaks the cohesive bonds.

In Barenblatt's model for brittle fracture [13] a nonlinear relation is assumed between the cohesive restraining action and the separation between the crack faces.

The bond-reactions are variable with the opening, first increasing from the pointed nose of the process zone until a maximum is reached, and then decreasing to zero, in correspondence of a threshold value of the opening, where breaking of the bonds occurs, at the crack tip.

To provide the expression of the driving force F acting on cohesive cracks, propagating with a translational speed $\dot{\boldsymbol{\chi}}(\mathbf{p}) = \dot{\chi}\,\mathbf{d}$, we rely again upon the general expression of the driving force:

$$F = \int_{\mathcal{I}} \mathbf{g}([[\mathbf{Y}]]\mathbf{n}, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n})\,,$$

where the interface $\mathcal{I}$ is the closed contour of the process zone.

Following Rice [145], we make the simplifying assumption that, due to the slit-shape of the crack, we may set $\mathbf{g}(\mathbf{n}, \mathbf{d}) = 0$ along the crack faces. Since the flux-jump $[[\mathbf{Y}]]\mathbf{n}$ of Eshelby's tensor is directed along the normal $\mathbf{n}$ at the interface, the contribution of the crack faces to the driving force vanishes. Then the integral can be extended only to the back-portion $\mathcal{B}$ of the interface which cuts the crack in correspondence of the end of the process zone, where breaking of the bonds occurs. There $\mathbf{g}(\mathbf{n}^-, \mathbf{d}) = -1$ and $\mathbf{Y}^+ = 0\,,\ \mathbf{Pn} = 0\,$, so that:

$$F = \int_{\mathcal{B}} \mathbf{g}(-\mathbf{Y}^-\mathbf{n}^-, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n}) = \int_{\mathcal{B}} \mathbf{g}(-\mathbf{Y}^-\mathbf{n}^-, \mathbf{d})\ (\boldsymbol{\mu}\mathbf{n}) = \int_{\mathcal{B}} W^-\ (\boldsymbol{\mu}\mathbf{n})\,.$$

The energy W^- is the one accumulated in the cohesive bonds per unit volume in correspondence of the breaking surface $\mathcal{B}$. Its integral over the surface $\mathcal{B}$ is equal to the area of the Barenblatt diagram for the cohesive bond and its product by the propagation speed provides the energy release rate due to the bond breaking.

This result is in accordance with the conclusions obtained by Rice on the basis of an *a priori* definition of the J-integral [145].

5.5.4 Conclusions

We owe essentially to Morton Gurtin the approach followed for the description of phase-transition phenomena in which phase-transition fronts are considered as shock waves travelling in the material [70].

Gurtin's point of view appears to have been strongly influenced by the attempt to prove that configurational forces are basics concepts of continuum physics. His intention of endowing Eshelby's tensor with properties similar to Piola's stress led him to make the assumption that no free energy release rate is associated with the evolution of phase transition fronts ([70] chapter 1, section b, page 4). This ansatz cannot be agreed on since the physics of these phenomena tell us that a dissipation occurs at expenses of a free energy release rate. We have shown that the balance law, derived from the virtual work principle of mechanics by a suitable definition of the free energy density for multi-phase materials, provides the basic expression of the dissipation associated with the evolution of phase-transition fronts. By applying the theory to crack propagation phenomena in fracture mechanics, we have shown that the J-integral, introduced *a priori* in [145], is in fact a special expression of the general dissipation formula for phase-transition fronts travelling in the material. Both non-cohesive and cohesive crack propagation may be directly analyzed by the present theory.

5.5.5 Noll's theory of material behavior

The theoretical scheme in Noll's theory of material behavior is outlined hereafter. Reference is made to the exposition provided in [134] and in [35] with some modifications.

A material body is a fiber bundle $\mathbb{E}$ with base manifold $\mathbb{M}$ and typical fiber a manifold S whose elements are the *states* $\mathrm{s} \in \mathrm{S}$ of the base material point. The *states* manifold S is assumed to be a metric space.

A map $\boldsymbol{\varphi} \in C^1(S\,;\mathcal{D})$ assign a *condition* $\mathrm{E} \in \mathrm{E}$, which is observable and controllable, to each state. The map $\boldsymbol{\varphi} \in C^1(S\,;E)$ is not injective, in general, so that many states may be in relation with a given condition. The simplest instance is that in which states belong to a finite list of linear spaces and conditions are elements of a sublist.

Time changes of conditions are described by *process* maps $\mathbf{P} \in C^1(I\,;E)$ with $I = [t_i(\mathbf{P}), t_f(\mathbf{P})]$ a time interval. A process may also be considered as a transformation $\mathbf{P} \in C^1(E\,;E)$ which maps the condition $\mathbf{P}(t_i(\mathbf{P}))$ into the condition $\mathbf{P}(t_f(\mathbf{P}))$. Accordingly, the composition of two subsequent processes $\mathbf{P}_1 \in C^1(I_1\,;E)$ and $\mathbf{P}_2 \in C^1(I_2\,;E)$ with $t_f(\mathbf{P}_1) = t_i(\mathbf{P}_2)$ is denoted by $\mathbf{P}_2 \circ \mathbf{P}_1$. The space of admissible processes is denoted by Π.

Changes of *state* are assumed to be produced by *process* maps and are described by *evolution* maps which, to any admissible *process* $\mathbf{P} \in \Pi$ assign a *state* transformation $\boldsymbol{\psi}_{\mathbf{P}} \in C^1(S\,;S)$ from a state $\mathrm{s} \in S$ such that $\boldsymbol{\varphi}(\mathrm{s}) = \mathbf{P}(t_i(\mathbf{P})) \in E$ to the state $\boldsymbol{\psi}_{\mathbf{P}}(\mathrm{s}) \in S$ such that $\boldsymbol{\varphi}(\boldsymbol{\psi}_{\mathbf{P}}(\mathrm{s})) = \mathbf{P}(t_f(\mathbf{P})) \in E$.

Evolution maps are denoted alternatively by $\boldsymbol{\psi} \in C^1(\mathbf{P} \times S\,;S)$ or $\boldsymbol{\psi}_{\mathbf{P}} \in C^1(S\,;S)$ or $\boldsymbol{\psi}_{\mathrm{s}} \in C^1(\mathcal{P}\,;S)$ setting

$$\boldsymbol{\psi}(\mathbf{P}, \mathrm{s}) = \boldsymbol{\psi}_{\mathbf{P}}(\mathrm{s}) = \boldsymbol{\psi}_{\mathrm{s}}(\mathbf{P})\,.$$

Let us denote by $\mathrm{FR}_{\mathrm{E}} \in C^1(I\,;E)$ a *freezing* process with constant value $\mathrm{E} \in E$.

The *relaxed state* s_∞ corresponding to the *state* $\mathrm{s} \in S$ is the limit of the evolution along a freezing process as time goes by:

$$\mathrm{s}_\infty := \lim_{t_f(\mathrm{FR}_{\mathrm{E}}) \to \infty} \boldsymbol{\psi}_{\mathrm{s}}(\mathrm{FR}_{\mathrm{E}})\,,$$

where $\mathrm{E} = \boldsymbol{\varphi}(\mathrm{s})$. The existence of this limit may be directly assumed, as in [134] or deduced by other assumptions. In [35] it is proposed to endow the process space Π with a metric fulfilling the *fading distance property* that the distance between two processes, which coalesce after a finite time into a unique process, goes to zero as time goes by:

$$\left.\begin{aligned} &\mathbf{P}, \mathbf{P}_1, \mathbf{P}_2 \in \Pi \\ &t_f(\mathbf{P}) \to \infty \end{aligned}\right\} \implies \mathrm{DIST}_{\Pi}(\mathbf{P} \circ \mathbf{P}_1, \mathbf{P} \circ \mathbf{P}_2) \to 0\,.$$

Further the *elastic region* $\mathcal{E}(\mathrm{s}) \subset E$ pertaining to a state $\mathrm{s} \in S$ is defined as the maximal set of *conditions* such that the following properties hold true: *i)* the *condition* $\boldsymbol{\varphi}(\mathrm{s})$ belongs to the elastic region of the state $\mathrm{s} \in S$, i.e. $\boldsymbol{\varphi}(\mathrm{s}) \subset \mathcal{E}(\mathrm{s}) \subset E$, *ii)* the restriction of any *evolution* map $\boldsymbol{\psi}_{\mathrm{s}} \in C^0(\mathcal{P}\,;S)$

to the space $\Pi_{\mathcal{E}}(\mathrm{s})$ of *elastic processes*, that is those evolving into the elastic region $\mathcal{E}(\mathrm{s}) \subset \mathrm{E}$, is continuous with respect to the pair of metrics, DIST_{Π} in the process space Π and $\mathrm{DIST}_{\mathrm{S}}$ in the state space S:

$$\mathrm{DIST}_{\Pi}(\mathbf{P}_1, \mathbf{P}_2) \to 0 \implies \mathrm{DIST}_{\mathrm{S}}(\boldsymbol{\psi}_{\mathrm{S}}(\mathbf{P}_1), \boldsymbol{\psi}_{\mathrm{S}}(\mathbf{P}_2)) \to 0\,, \quad \mathbf{P}_1, \mathbf{P}_2 \in \Pi_{\mathcal{E}}(\mathrm{s})\,.$$

The following properties are readily proved.

Theorem 5.5.1 (Fading memory) *The evolution along two elastic processes, which coalesce after a finite time into a unique elastic process, tend to a unique state as time goes by*

$$\left.\begin{aligned} &\mathbf{P}, \mathbf{P}_1, \mathbf{P}_2 \in \Pi_{\mathcal{E}}(\mathrm{s}) \\ &t_f(\mathbf{P}) \to \infty \end{aligned}\right\} \implies \mathrm{DIST}_{\mathrm{S}}(\boldsymbol{\psi}_{\mathrm{S}}(\mathbf{P} \circ \mathbf{P}_1), \boldsymbol{\psi}_{\mathrm{S}}(\mathbf{P} \circ \mathbf{P}_2)) \to 0\,.$$

Proof. The *fading distance property* and the continuity of *evolutions* in the elastic range provide the result. ∎

A state $\boldsymbol{\psi}_{\mathbf{P}}(\mathrm{s})$ with $\mathbf{P} \in \Pi$ is said to be *accessible* from the state $\mathrm{s} \in \mathrm{S}$ via the process $\mathbf{P}$.

Theorem 5.5.2 (Relaxation) *All states accessible from a given state* $\mathrm{s} \in \mathrm{S}$ *via an elastic process admit the same relaxed limit:*

$$\mathrm{s}_{\infty} := \lim_{t_f(\mathrm{FR}_{\mathrm{E}}) \to \infty} \boldsymbol{\psi}_{\mathrm{FR}_{\mathrm{E}}}(\mathrm{s}) = \lim_{t_f(\mathrm{FR}_{\mathrm{E}}) \to \infty} (\boldsymbol{\psi}_{\mathrm{FR}_{\mathrm{E}}} \circ \boldsymbol{\psi}_{\mathbf{P}})(\mathrm{s})\,, \quad \mathbf{P} \in \Pi_{\mathcal{E}}(\mathrm{s})\,.$$

Proof. The *fading distance property* ensures that $\mathrm{DIST}_{\Pi}(\mathrm{FR}_{\mathrm{E}} - \mathrm{FR}_{\mathrm{E}} \circ \mathbf{P}) \to 0$ as $t_f(\mathrm{FR}_{\mathrm{E}}) \to \infty$. Moreover, by the continuity of *evolutions* in the elastic range:

$$\mathrm{DIST}_{\Pi}(\mathrm{FR}_{\mathrm{E}}, \mathrm{FR}_{\mathrm{E}} \circ \mathbf{P}) \to 0 \implies \mathrm{DIST}_{\mathrm{S}}(\boldsymbol{\psi}_{\mathrm{S}}(\mathrm{FR}_{\mathrm{E}}), \boldsymbol{\psi}_{\mathrm{S}}(\mathrm{FR}_{\mathrm{E}} \circ \mathbf{P})) \to 0\,,$$

and the chain of the two implications provides the result. ∎

Denoting by $\Sigma(\mathrm{s})$ the set of all states accessible from a given state $\mathrm{s} \in \mathrm{S}$ via an elastic process, the assumption that

$$\bar{\mathrm{s}} \in \Sigma(\mathrm{s}) \implies \mathcal{E}(\bar{\mathrm{s}}) = \mathcal{E}(\mathrm{s})\,, \quad \Sigma(\bar{\mathrm{s}}) = \Sigma(\mathrm{s})\,,$$

makes $\Sigma(\mathrm{s})$ a family of equivalence classes in S.

According to [35], the change of equivalence class may be seen as a signal that a plastic deformation has occurred.

Chapter 6

Thermodynamics

6.1 Thermodynamic state variables and related potentials

Let us consider a spatial placement of a body in the euclidean space $\{\mathbb{S}, \mathbf{g}\}$ described by a configuration map $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{B}; \mathbb{S})$ from a reference placement $\mathbb{B} \subset \mathbb{S}$, an embedded submanifold of $\mathbb{S}$.

We denote by $\boldsymbol{\Omega} = \boldsymbol{\varphi}(\mathbb{B}) \subset \mathbb{S}$ the actual placement of the body and by $\boldsymbol{\varepsilon}(\boldsymbol{\varphi}) \in \mathcal{L}^2(\mathbb{B}; D)$ the deformation field induced at $\mathbf{m} \in \mathbb{B}$ by the configuration map $\boldsymbol{\varphi} \in \mathrm{C}^1(\mathbb{B}; \mathbb{S})$ and by $\mathbf{a} \in \mathcal{L}^2(\mathbb{B}; D)$ the related field of anelastic strains. The internal energy density $\mathcal{U}$ of a body is the functional which provides the field of the elastic energy density as a function of the elastic strain field

$$\mathbf{e} = \boldsymbol{\varepsilon}(\boldsymbol{\varphi}) - \mathbf{a} \in \mathcal{L}^2(\mathbb{B}; D),$$

and of the entropy density field $\eta \in \mathcal{L}^2(\mathbb{B}; \Re)$, so that:

$$\hat{\mathcal{U}}(\mathbf{m}) := \mathcal{U}(\boldsymbol{\varepsilon}(\mathbf{m}), \mathbf{a}(\mathbf{m}), \eta(\mathbf{m})), \quad \forall\, \mathbf{m} \in \mathbb{B}.$$

The constitutive thermodynamic relations are expressed by

$$\boldsymbol{\sigma} \in \partial\mathcal{U}_{\{\eta, \mathbf{a}\}}(\boldsymbol{\varepsilon}), \qquad \theta \in \partial\mathcal{U}_{\{\boldsymbol{\varepsilon}, \mathbf{a}\}}(\eta),$$

where $\boldsymbol{\sigma} \in \mathcal{L}^2(\mathbb{B}; D)$ is the stress field conjugate to the strain measure $\boldsymbol{\varepsilon} \in \mathcal{L}^2(\mathbb{B}; D)$ and $\theta \in \mathcal{L}^2(\mathbb{B}; \Re)$ is the absolute temperature field.

By changing the choice of the state variables in the dual pairs $\{\boldsymbol{\varepsilon}\,,\boldsymbol{\sigma}\}$ and $\{\theta\,,\eta\}$ we generate other basic thermodynamic potentials.

The HELMHOLTZ free energy density $\mathcal{F}(\boldsymbol{\varepsilon},\theta)$, the enthalpy density $\mathcal{H}(\boldsymbol{\sigma},\eta)$ and the GIBBS free energy density $\mathcal{G}(\boldsymbol{\sigma},\theta)$, are the LEGENDRE-FENCHEL-conjugate of the internal energy density, according to the complementarity rules

$$\begin{aligned} \mathcal{U}(\boldsymbol{\varepsilon},\eta) - \mathcal{F}(\boldsymbol{\varepsilon},\theta) &= \langle \eta,\theta \rangle\,, \\ \mathcal{F}(\boldsymbol{\varepsilon},\theta) - \mathcal{G}(\boldsymbol{\sigma},\theta) &= \langle \boldsymbol{\sigma},\boldsymbol{\varepsilon} \rangle\,, \\ \mathcal{U}(\boldsymbol{\varepsilon},\eta) - \mathcal{H}(\boldsymbol{\sigma},\eta) &= \langle \boldsymbol{\sigma},\boldsymbol{\varepsilon} \rangle\,. \end{aligned}$$

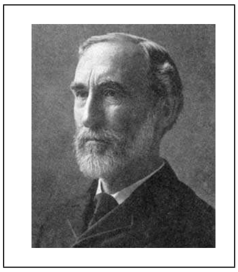

Figure 6.1: Josiah Willard Gibbs (1839 - 1903)

The LEGENDRE transformation rules hold pointwise, under the assumption that the internal energy density $\mathcal{U}_{\mathbf{a}}(\boldsymbol{\varepsilon},\eta)$ is a convex function of each of its two arguments. By the rules of convex analysis we then infer the following relations between the thermodynamic potentials:

$$-\mathcal{F}(\boldsymbol{\varepsilon},\theta) = \inf_{\bar{\eta}\in\Re}\{\langle \bar{\eta},\theta \rangle - \mathcal{U}(\boldsymbol{\varepsilon},\bar{\eta})\}\,, \qquad \text{HELMHOLTZ free energy density}$$

$$-\mathcal{G}(\boldsymbol{\sigma},\theta) = \inf_{\bar{\boldsymbol{\varepsilon}}\in D}\{\langle \boldsymbol{\sigma},\bar{\boldsymbol{\varepsilon}} \rangle - \mathcal{F}(\bar{\boldsymbol{\varepsilon}},\theta)\}\,, \qquad \text{GIBBS free energy density}$$

$$-\mathcal{H}(\boldsymbol{\sigma},\eta) = \inf_{\bar{\boldsymbol{\varepsilon}}\in D}\{\langle \boldsymbol{\sigma},\bar{\boldsymbol{\varepsilon}} \rangle - \mathcal{U}(\bar{\boldsymbol{\varepsilon}},\eta)\}\,, \qquad \text{enthalpy density}$$

The HELMHOLTZ free energy density $\mathcal{F}(\boldsymbol{\varepsilon},\theta)$ is convex-convex. Indeed its opposite is concave in $\boldsymbol{\varepsilon}$, being the infimum of a family of concave functions, and concave in θ being the infimum of a family of affine functions.
The GIBBS free energy density $\mathcal{G}(\boldsymbol{\sigma},\theta)$ is convex-convex. Indeed its opposite is concave in $\boldsymbol{\sigma}$, being the infimum of a family of affine functions, and concave in θ being the infimum of a family of concave functions.

The enthalpy density $\mathcal{H}(\boldsymbol{\sigma}, \eta)$ is convex-convex. Indeed its opposite is concave in $\boldsymbol{\sigma}$, being the infimum of a family of affine functions, and concave in η being the infimum of a family of concave functions.

Being

$$\mathcal{G}(\boldsymbol{\sigma}, \theta) - \mathcal{H}(\boldsymbol{\sigma}, \eta) = \mathcal{F}(\boldsymbol{\varepsilon}, \theta) - \mathcal{U}(\bar{\boldsymbol{\varepsilon}}, \eta) = -\langle \eta, \theta \rangle \,,$$

we have also that:

$$\mathcal{H}(\boldsymbol{\sigma}, \eta) \, = \inf_{\bar{\theta} \in \Re} \{ \langle \eta, \bar{\theta} \rangle + \mathcal{G}(\boldsymbol{\sigma}, \bar{\theta}) \} \,.$$

or

$$-\mathcal{G}(\boldsymbol{\sigma}, \theta) \, = \inf_{\bar{\eta} \in \Re} \{ \langle \bar{\eta}, \theta \rangle - \mathcal{H}(\boldsymbol{\sigma}, \bar{\eta}) \} \,.$$

6.2 Conservation of energy

Let us consider a spatial placement of a body $\mathcal{B}$ in the euclidean space $\{\mathbb{S}\,, \mathbf{g}\}$ described by a configuration map $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{B}\,; \mathbb{S})$ from a reference placement $\mathbb{B} \subset \mathbb{S}$, an embedded submanifold of $\mathbb{S}$.

We denote by $\mathbf{\Omega} = \boldsymbol{\varphi}_t(\mathbb{B}) \subset \mathbb{S}$ the placement of the body at time $t \in I$ and by $Sym \subset BL\,(V^2\,; \Re)$ the space of symmetric tensors on the translations space V of $\{\mathbb{S}\,, \mathbf{g}\}$.

The symmetric Green's tensor field $\frac{1}{2}(\boldsymbol{\varphi}_t {\downarrow} \mathbf{g} - \mathbf{g})$ is defined in the reference placement and measures the deformation of $\mathcal{B}$ induced by the configuration map $\boldsymbol{\varphi}_t \in \mathrm{C}^1(\mathbb{B}\,; \mathbb{S})$.

By adopting the subscript 0 to denote quantities pertaining to the reference placement, we denote the strain tensor field by

$$\boldsymbol{\varepsilon}_0 := \tfrac{1}{2}(\boldsymbol{\varphi}_t {\downarrow} \mathbf{g} - \mathbf{g}) \,.$$

Time-dependent anelastic phenomena are simulated, at the continuum level by a metric tensor field $\mathbf{g}_{\mathbb{B}t}$ in the reference placement. The anelastic strain tensor $\mathbf{a}_0$ is then defined by

$$\mathbf{a}_0 := \tfrac{1}{2}\,(\mathbf{g}_{\mathbb{B}t} - \mathbf{g}) \,.$$

The internal energy density $\mathcal{U}_{0t}$ per unit mass is assumed to be a pointwise function of the values of the strain field $\boldsymbol{\varepsilon}_0 \in \mathcal{L}^2(\mathbb{B}\,; D)$, of the anelastic strain $\mathbf{a}_0 \in \mathcal{L}^2(\mathbb{M}\,; D)$ and of the entropy density field $\eta_0 \in \mathcal{L}^2(\mathbb{M}\,; \Re)$ per unit mass:

$$\mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0(\mathbf{m}), \mathbf{a}_0(\mathbf{m}), \eta_0(\mathbf{m})) := \mathcal{U}_0(\boldsymbol{\varepsilon}_{0t}(\mathbf{m}), \mathbf{a}_{0t}(\mathbf{m}), \eta_{0t}(\mathbf{m})) \,, \quad \forall\, \mathbf{m} \in \mathbb{M} \,.$$

The constitutive relations are expressed pointwise, in the reference placement, by

$$\begin{aligned}\boldsymbol{\sigma}_{0t} &= \rho_{0t}\, d_1\mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0)\,,\\ \bar{\boldsymbol{\sigma}}_{0t} &= -\rho_{0t}\, d_2\mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0)\,,\\ \theta_{0t} &= d_3\mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0)\,,\end{aligned}$$

where $\boldsymbol{\sigma}_{0t}, \bar{\boldsymbol{\sigma}}_{0t} \in \mathcal{L}^2(\mathbb{M}\,; Sym)$ are the elastic and the anelastic stress fields and $\theta_{0t} \in \mathcal{L}^2(\mathbb{M}\,;\Re)$ is the absolute-temperature field in the reference placement. The global internal energy is the integral of its density per unit mass:

$$\mathcal{E}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0) := \int_{\mathbb{M}} \mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0)\, \rho_{0t}\, \boldsymbol{\mu}\,,$$

where ρ_{0t} is the mass-density per unit volume in the reference placement $\mathbb{M}$.

The internal energy density per unit mass in the placement $\boldsymbol{\varphi}_t(\mathbb{M})$ is defined as the push-forward of the one pertaining to the reference placement:

$$(\boldsymbol{\varphi}_t{\uparrow}\mathcal{U}_{0t})(\boldsymbol{\varphi}_t{\uparrow}\boldsymbol{\varepsilon}_0, \boldsymbol{\varphi}_t{\uparrow}\mathbf{a}_0, \boldsymbol{\varphi}_t{\uparrow}\eta_0) \circ \boldsymbol{\varphi}_t := \mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0)\,.$$

The point-value of the internal energy at $\mathbf{m} \in \mathbb{M}$, due to the state variables evaluated at time $t \in I$ in the reference placement, is then equal to the point-value of the internal energy at $\boldsymbol{\varphi}_t(\mathbf{m}) \in \boldsymbol{\varphi}_t(\mathbb{M})$, due to the push-forward to that point of the state variables. Setting

$$\begin{aligned}\eta_t &:= \boldsymbol{\varphi}_t{\uparrow}\eta_0 = \eta_0 \circ \boldsymbol{\varphi}_t^{-1}\,,\\ \boldsymbol{\varepsilon}_t &:= \boldsymbol{\varphi}_t{\uparrow}\boldsymbol{\varepsilon}_0 = \tfrac{1}{2}(\mathbf{g} - \boldsymbol{\varphi}_t{\uparrow}\mathbf{g})\,,\\ \mathbf{a}_t &:= \boldsymbol{\varphi}_t{\uparrow}\mathbf{a}_0 = \tfrac{1}{2}\boldsymbol{\varphi}_t{\uparrow}(\mathbf{g}_{\mathbb{M}t} - \mathbf{g})\,,\end{aligned}$$

we adopt the simplified notation $\mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t) := (\boldsymbol{\varphi}_t{\uparrow}\mathcal{U}_{0t})(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t)$.

The partial derivatives of $\mathcal{U}_t$ and $\mathcal{U}_{0t}$, with respect to their i-th argument, are related by

$$d_i\,\mathcal{U}_t := d_i\,(\boldsymbol{\varphi}_t{\uparrow}\mathcal{U}_{0t}) = \boldsymbol{\varphi}_t{\uparrow}(d_i\,\mathcal{U}_{0t})\,.$$

The constitutive relations are then expressed pointwise, in terms of fields in the current placement $\boldsymbol{\varphi}_t(\mathbb{M})$, by

$$\begin{aligned}\boldsymbol{\sigma}_t &= \rho_t\, d_1\mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t)\,,\\ \bar{\boldsymbol{\sigma}}_t &= -\rho_t\, d_2\mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t)\,,\\ \theta_t &= d_3\mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t)\,.\end{aligned}$$

The elastic and the anelastic stress fields $\boldsymbol{\sigma}_t, \bar{\boldsymbol{\sigma}}_t \in \mathcal{L}^2(\boldsymbol{\varphi}_t(\mathbb{M}) ; Sym)$ and the absolute-temperature field $\theta_t \in \mathcal{L}^2(\boldsymbol{\varphi}_t(\mathbb{M}) ; \Re)$, are related to the corresponding reference fields by

$$\begin{aligned}
&\boldsymbol{\sigma}_t \otimes \boldsymbol{\mu} = \boldsymbol{\varphi}_t\!\uparrow\!(\boldsymbol{\sigma}_{0t} \otimes \boldsymbol{\mu}) = (\boldsymbol{\varphi}_t\!\uparrow\!\boldsymbol{\sigma}_{0t}) \otimes (\boldsymbol{\varphi}_t\!\uparrow\!\boldsymbol{\mu})\,,\\
&\bar{\boldsymbol{\sigma}}_t \otimes \boldsymbol{\mu} = \boldsymbol{\varphi}_t\!\uparrow\!(\bar{\boldsymbol{\sigma}}_{0t} \otimes \boldsymbol{\mu}) = (\boldsymbol{\varphi}_t\!\uparrow\!\bar{\boldsymbol{\sigma}}_{0t}) \otimes (\boldsymbol{\varphi}_t\!\uparrow\!\boldsymbol{\mu})\,,\\
&\theta_t = \boldsymbol{\varphi}_t\!\uparrow\!\theta_{0t}\,.
\end{aligned}$$

The tensor $\boldsymbol{\sigma}_t \otimes \boldsymbol{\mu}$ is the TRUESDELL stress tensor. When contracted with the tangent-strain rate, it provides the volume-form of *mechanical working*:

$$(\boldsymbol{\sigma}_t \otimes \boldsymbol{\mu})(\tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}) = \langle \boldsymbol{\sigma}, \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g} \rangle\, \boldsymbol{\mu}\,.$$

By the transformation rule for integrals of volume-forms under a diffeomorphism and the principle of conservation of mass $\boldsymbol{\varphi}_{t*}(\rho_{0t}\boldsymbol{\mu}) = \rho_t \boldsymbol{\mu}$, the global internal energy may be written as an integral over the current placement:

$$\mathcal{E}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0) = \int_{\mathbb{M}} \mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0)\, \rho_{0t}\, \boldsymbol{\mu} = \int_{\boldsymbol{\varphi}_t(\mathbb{M})} \mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t)\, \rho_t\, \boldsymbol{\mu} = \mathcal{E}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t)$$

where ρ_t is the mass-density per unit volume in the current placement $\boldsymbol{\varphi}_t(\mathbb{M})$.

The *First Principle of Thermodynamics* asserts that, for any body $\mathcal{B}$ at any time $t \in I$ the *law of conservation of energy* holds:

$$\dot{\mathcal{E}}_t := \partial_{\tau=t}\, \mathcal{E}_\tau = \mathcal{M}_t + \mathcal{Q}_t\,,$$

where $\dot{\mathcal{E}}_t$ is the time-rate of change of the *internal energy*, $\mathcal{M}_t$ is the *mechanical working*, $\mathcal{Q}_t$ is the *heat working.*

In the sequel the subscript t will be often dropped when redundant.

The mechanical working is the power performed by the total force system acting on the body which includes body forces, inertia forces and boundary tractions. According to the principle of virtual works, it is given by

$$\begin{aligned}
\mathcal{M} &= \int_{\boldsymbol{\varphi}(\mathbb{M})} \mathbf{g}\,(\mathbf{b} - \rho\dot{\mathbf{v}}, \mathbf{v})\, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{M}))} \mathbf{g}\,(\mathbf{t}, \mathbf{v})\, \partial\boldsymbol{\mu}\\
&= \int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{M}))} \langle \boldsymbol{\sigma}, \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g} \rangle\, \boldsymbol{\mu} = \int_{\mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{M}))} \langle \mathbf{T}, \mathrm{sym}\,\partial\mathbf{v} \,\rangle_{\mathbf{g}}\, \boldsymbol{\mu}\,,
\end{aligned}$$

- ρ is the spatial mass density along the trajectory,

- $\mathbf{v}, \dot{\mathbf{v}}$ are the fields of velocities and accelerations,
- $\mathbf{b}$ is the field of body forces per unit volume,
- $\mathbf{t}$ is the field of boundary tractions,
- $\boldsymbol{\sigma} = \mathbf{gT}$ is the CAUCHY stress tensor field,
- $\frac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g} = \mathbf{g}(\mathrm{sym}\,\partial\mathbf{v})$ is tangent strain-rate tensor field,
- $\boldsymbol{\mu}$ is the volume form induced by the metric $\mathbf{g}$
- $\boldsymbol{\mu}\mathbf{n}$ is the associated surface area form.

The kinetic energy $\mathcal{K}$ and the power $\mathcal{W}$ performed by applied forces are

$$\mathcal{K} = \frac{1}{2} \int_{\boldsymbol{\varphi}(\mathbb{M})} \mathbf{g}(\mathbf{v}, \mathbf{v})\, \rho\, \boldsymbol{\mu} = \int_{\mathbb{M}} (\mathbf{g}(\mathbf{v}, \mathbf{v}) \circ \boldsymbol{\varphi})\ \rho\, \boldsymbol{\mu}\,,$$

$$\mathcal{W} = \int_{\boldsymbol{\varphi}(\mathbb{M})} \mathbf{g}\,(\mathbf{b}, \mathbf{v})\, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}(\boldsymbol{\varphi}(\mathbb{M}))} \mathbf{g}\,(\mathbf{t}, \mathbf{v})\, \partial\boldsymbol{\mu}\,.$$

The mechanical power is then given by $\mathcal{M} = \mathcal{W} - \dot{\mathcal{K}}$ and the law of conservation of the energy may be written as

$$\dot{\mathcal{E}} + \dot{\mathcal{K}} = \mathcal{W} + \mathcal{Q}\,.$$

The principle of conservation of energy is a balance law prescribing a rule to be fulfilled by any thermodynamical process which evolves starting from a given thermodynamical status $\{\boldsymbol{\varphi}\,, \eta\}$ of the body.

Our main target is to show how to adapt the proof of the virtual work theorem in mechanics to formulate the principle of conservation of energy as a virtual temperature theorem. This result assesses the existence of a vector field, representing the *cold-flow* in the body, which fulfils a virtual balance law. The analogy with the equilibrium condition of mechanics permits to extend, *mutatis mutandis*, propositions and results from one context to the other, once it has been recognized that both rely upon the same formal mathematical base. This is indeed the peculiar task of *Mathematical Physics*.

6.2.1 Virtual temperatures

The linear space $\text{TEMP}(\boldsymbol{\Omega})$ of GREEN-*regular temperatures* is composed by square integrabile scalar fields $\theta \in \text{SQIF}(\boldsymbol{\Omega})$ whose generalized derivatives are piecewice regular in $\boldsymbol{\Omega}$ according to a regularity patchwork $\text{PAT}_\theta(\boldsymbol{\Omega})$ with boundary $\partial\text{PAT}_\theta(\boldsymbol{\Omega})$ and interfaces $\text{IF}(\text{PAT}_\theta(\boldsymbol{\Omega}))$.

To provide a precise definition, we recall that a distribution on $\boldsymbol{\Omega}$ is a linear functional on $C_0^\infty(\boldsymbol{\Omega}\,;V)$ which is continuous according to the topology induced by the uniform convergence of every derivative on any compact subset of the open set $\boldsymbol{\Omega}$. The distributional gradient ∇ is the linear operator which, to any $\theta \in \text{TEMP}(\boldsymbol{\Omega})$, associates the distribution defined by

$$\langle \nabla\theta, \boldsymbol{\lambda} \rangle := -\int_{\boldsymbol{\Omega}} (\operatorname{div} \boldsymbol{\lambda})\, \theta\, \boldsymbol{\mu}\,, \qquad \forall\, \boldsymbol{\lambda} \in C_0^\infty(\boldsymbol{\Omega}\,;V)\,,$$

and $-\text{div}$ is called the formal dual of the gradient operator.

The piecewice regularity consists in requiring that the distributional gradient $\nabla\theta$ be represented by a GREEN's formula:

$$\begin{aligned}\langle \nabla\theta, \boldsymbol{\lambda} \rangle &= -\int_{\boldsymbol{\Omega}} (\operatorname{div} \boldsymbol{\lambda})\, \theta\, \boldsymbol{\mu} \\ &= \int_{\text{PAT}_\theta(\boldsymbol{\Omega})} \mathbf{g}(\nabla\theta, \boldsymbol{\lambda})\, \boldsymbol{\mu} + \int_{\mathcal{I}_\theta(\boldsymbol{\Omega})} [[\boldsymbol{\Gamma}\theta]]\, \mathbf{g}(\boldsymbol{\lambda}, \mathbf{n})\, \partial\boldsymbol{\mu}\,.\end{aligned}$$

where $\boldsymbol{\Gamma}\theta$ is the boundary value on $\partial\text{PAT}_\theta(\boldsymbol{\Omega})$ of the field $\theta \in \text{TEMP}(\boldsymbol{\Omega})$ and

$$[[\boldsymbol{\Gamma}\theta]] = \boldsymbol{\Gamma}\theta^+ - \boldsymbol{\Gamma}\theta^-\,,$$

is the jump across the interfaces $\mathcal{I}_\theta(\boldsymbol{\Omega})$ of the patchwork $\text{PAT}_\theta(\boldsymbol{\Omega})$ and $\mathbf{n} = \mathbf{n}^-$ is the outward normal pointing towards the $+$ face. The square integrable vector fields $\nabla\theta \in \text{SQIV}(\text{PAT}_\theta(\boldsymbol{\Omega}))$ and scalar fields $[[\boldsymbol{\Gamma}\theta]] \in \text{SQIF}(\text{IF}_\theta(\boldsymbol{\Omega}))$ are respectively said to be the *regular part* and the *singular part* of the distributional gradient.

The space $\text{TEMP}(\boldsymbol{\Omega})$ of virtual temperature fields is a pre-HILBERT space when endowed with the inner product and norm given by

$$\begin{aligned}\langle \theta_1, \theta_2 \rangle &:= \int_{\boldsymbol{\Omega}} \theta_1\, \theta_2\, \boldsymbol{\mu} + \int_{\text{PAT}_\theta(\boldsymbol{\Omega})} \mathbf{g}(\nabla\theta_1, \nabla\theta_2)\, \boldsymbol{\mu}\,, \\ \|\mathbf{u}\|^2 &:= \int_{\boldsymbol{\Omega}} \theta^2\, \boldsymbol{\mu} + \int_{\text{PAT}_\theta(\boldsymbol{\Omega})} \|\nabla\theta\|^2\, \boldsymbol{\mu}\,.\end{aligned}$$

GREEN-regular temperature fields will be dubbed *virtual temperatures*.

Given a patchwork $\text{PAT}(\mathbf{\Omega})$, the *conforming virtual temperatures* belong to a closed linear subspace $\text{CONF}(\mathbf{\Omega}) \subset \text{TEMP}(\mathbf{\Omega})$ of fields having $\text{PAT}(\mathbf{\Omega})$ as common regularity patchwork. It is a Hilbert space for the topology induced by $\text{TEMP}(\mathbf{\Omega})$.

Piecewice constant virtual temperature fields belong to a closed subspace $\text{CONST}(\mathbf{\Omega}) \subset \text{TEMP}(\mathbf{\Omega})$. The fields $\theta \in \text{CONST}(\mathbf{\Omega})$ are characterized by the property that the regular part $\nabla\theta$ of their distributional gradient vanishes on each element of the regularity patchwork $\text{PAT}_\theta(\mathbf{\Omega})$.

Entropy rate systems in the body at the placement $\mathbf{\Omega}$ are continuous linear functionals defined on $\text{TEMP}(\mathbf{\Omega})$ and hence belong to the pre-HILBERT space $\text{TEMP}(\mathbf{\Omega})^*$ topological dual of $\text{TEMP}(\mathbf{\Omega})$.

6.2.2 The variational form of the first principle

To get the existence result provided by the theorem of virtual thermal work, we need to re-formulate the first principle of thermodynamics as a variational condition. To this end we begin by providing a definition of the rate of increase of the inner energy $\dot{\mathcal{E}}$, the mechanical power $\mathcal{M}$ and the heat power $\mathcal{Q}$, in terms of bounded linear functionals over the space $\text{TEMP}(\mathbf{\Omega})$ of virtual temperature fields.

For any $\theta \in \text{TEMP}(\mathbf{\Omega})$ let us consider the characteristic functions of the elements of the partion $\text{PAT}_\theta(\mathbf{\Omega})$:

$$1_{\mathcal{P}}(\mathbf{m}) = \begin{cases} 1 & \mathbf{m} \in \mathcal{P} \\ 0 & \mathbf{m} \in \mathbf{\Omega} \backslash \mathcal{P} \end{cases}$$

with $\mathcal{P} \in \text{PAT}_\theta(\mathbf{\Omega})$. We then define

$$\mathcal{F}_{\dot{\mathcal{E}}}(1_{\mathcal{P}}) := \dot{\mathcal{E}}(\mathcal{P}),$$

$$\mathcal{F}_{\mathcal{M}}(1_{\mathcal{P}}) := \mathcal{M}(\mathcal{P}),$$

$$\mathcal{F}_{\mathcal{Q}}(1_{\mathcal{P}}) := \mathcal{Q}(\mathcal{P}).$$

An extension by linearity allows us to introduce the linear functionals $\mathcal{F}_{\dot{\mathcal{E}}}$, $\mathcal{F}_{\mathcal{M}}$ and $\mathcal{F}_{\mathcal{Q}}$ on the linear subspace $Ker\,\nabla \subseteq \text{TEMP}(\mathbf{\Omega})$ of piecewise constant virtual temperatures fields.

HAHN's extension theorem ensures that these bounded linear functionals may be extended (non-univocally) to bounded linear functionals on $\text{TEMP}(\mathbf{\Omega})$

without increasing their norm (see e.g. [192]). Anyway in section 6.2.4 it will be shown that in boundary value problems the linear functionals $\mathcal{F}_{\dot{\mathcal{E}}}$, $\mathcal{F}_{\mathcal{M}}$ and $\mathcal{F}_{\mathcal{Q}}$ can be univocally defined on the whole space $\text{TEMP}(\mathbf{\Omega})$.

The energy conservation law is then expressed by the variational condition:

$$\langle \mathcal{F}_{\dot{\mathcal{E}}}, \theta \rangle = \langle \mathcal{F}_{\mathcal{M}}, \theta \rangle + \langle \mathcal{F}_{\mathcal{Q}}, \theta \rangle , \quad \forall\, \theta \in Ker\, \nabla \,.$$

6.2.3 Virtual Thermal-Work Theorem

To stress the analogy between the energy conservation law and the equilibrium condition of a continuous body, we define the *energy-gap rate* as the difference between the time-rate of change of the *internal energy*, and the sum of the *mechanical working* plus the *heat working*:

$$\mathcal{G} := \dot{\mathcal{E}} - (\mathcal{M} + \mathcal{Q}) \,.$$

Then we consider the space $\text{TEMP}(\mathbf{\Omega})^*$, topological dual of $\text{TEMP}(\mathbf{\Omega})$, and introduce the *thermal force* as the linear functional $\mathcal{F}_{\mathcal{P}} \in \text{TEMP}(\mathbf{\Omega})^*$ given by:

$$\mathcal{F}_{\mathcal{G}} := \mathcal{F}_{\dot{\mathcal{E}}} - \mathcal{F}_{\mathcal{M}} - \mathcal{F}_{\mathcal{Q}} \,.$$

The energy conservation law $\mathcal{G} = 0$ takes then the variational form

$$\langle \mathcal{F}_{\mathcal{G}}, \theta \rangle = 0 \,, \quad \forall\, \theta \in Ker\, \nabla \,,$$

or in geometrical terms

$$\mathcal{F}_{\mathcal{G}} \in (Ker\, \nabla)^{\perp} \,.$$

This condition, which is perfectly analogous to the equilibrium condition in mechanics, states that the virtual power of the thermal force must vanish for any piecewise constant virtual temperature field.

Proceeding in analogy with the theory of equilibrium of continuous bodies as proposed in [165], we observe that the regular part of the distributional gradient trivially fulfills KORN's inequality:

$$\|\nabla\theta\|_0 + \|\theta\|_0 \geq \alpha\, \|\theta\|_1 \,, \quad \forall\, \theta \in H^1(\mathcal{P}\,; \Re) \,, \quad \mathcal{P} \in \text{PAT}(\mathbf{\Omega}) \,,$$

with $\alpha = 1$, where $\| \cdot \|_k$ denotes the mean square norm in $\mathcal{P}$ of the field and of all its derivatives up to the k order. Indeed we have that

$$\|\theta\|_1^2 := \|\nabla\theta\|_0^2 + \|\theta\|_0^2 \leq (\|\nabla\theta\|_0 + \|\theta\|_0)^2 \,.$$

For any closed linear subspace of conforming virtual temperatures

$$\textsc{Conf}(\mathbf{\Omega}) \subset \textsc{Temp}(\mathbf{\Omega}) \,,$$

the linear subspace $\nabla(\textsc{Conf}(\mathbf{\Omega})) \subset \textsc{Sqiv}(\mathbf{\Omega})$ is closed in $\textsc{Sqiv}(\mathbf{\Omega})$ and the kernel $Ker\, \nabla \cap \textsc{Conf}(\mathbf{\Omega})$ is finite dimensional [161].

Introducing the bounded linear operator $\nabla^* \in BL\,(\textsc{Sqiv}(\mathbf{\Omega})\,;\textsc{Conf}(\mathbf{\Omega})^*)$, dual of the operator $\nabla \in BL\,(\textsc{Conf}(\mathbf{\Omega})\,;\textsc{Sqiv}(\mathbf{\Omega}))$, defined by the property

$$\langle \nabla^* \mathbf{q}, \theta \rangle = \langle \mathbf{q}, \nabla\theta \rangle \,, \quad \forall\, \mathbf{q} \in \textsc{Sqiv}(\mathbf{\Omega}) \,, \quad \forall\, \theta \in \textsc{Conf}(\mathbf{\Omega}) \,,$$

we may invoke Banach's closed range theorem [192] to infer that the linear operator $\nabla^* \in BL\,(\textsc{Sqiv}(\mathbf{\Omega})\,;\textsc{Conf}(\mathbf{\Omega})^*)$ has also a closed range, and that

$$\nabla^*(\textsc{Sqiv}(\mathbf{\Omega})) = (Ker\, \nabla \cap \textsc{Conf}(\mathbf{\Omega}))^{\perp} \,.$$

Since $\mathcal{F}_{\mathcal{P}} \in (Ker\, \nabla)^{\perp} \subset (Ker\, \nabla_{\mathcal{L}})^{\perp}$ we may conclude that $\mathcal{F}_{\mathcal{P}} \in \mathrm{Im}\, \nabla^*_{\mathcal{L}}$ so that there exists at least a vector field $\mathbf{q} \in H(\mathbf{\Omega}\,; V)$, the *cold-flow vector field*, such that $\mathcal{F}_{\mathcal{P}} = \nabla^*_{\mathcal{L}} \mathbf{q}$ which is equivalent to the variational condition:

$$\langle \mathcal{F}_{\mathcal{P}}, \theta \rangle = \langle \nabla^*_{\mathcal{L}} \mathbf{q} \,, \theta \rangle = \int_{\mathbf{\Omega}} \mathbf{g}(\mathbf{q} \,, \nabla\theta)\, \boldsymbol{\mu} \,, \quad \forall\, \theta \in \mathcal{L}(\textsc{Pat}(\mathbf{\Omega})) \,.$$

This result is the statement of the *theorem of virtual thermal work*:

- The thermodynamical axiom (first principle of thermodynamics), stating that there is no energy creation in any part of a body, is equivalent to the assumption that in the body there exists a square integrable vector field $\mathbf{q} \in H(\mathbf{\Omega}\,; V)$, the *cold-flow vector field*, which performs, for the regular part of the distributional gradient of a conforming virtual temperature field, a thermal virtual work equal to the one that the thermal force performes for the conforming virtual temperature field.

The theorem of virtual thermal work provides a thermodynamical principle, analogous to the principle of virtual powers in mechanics, which we baptize the *principle of virtual temperatures*.

The cold-flow vector field plays in thermodynamics the same role as the one played by the stress tensor field in mechanics.

6.2.4 Boundary value problems

A boundary value problem (BVP) of thermodynamical equilibrium is characterized by the fulfillment of the following properties:

- the subspace of conforming virtual temperatures $\mathcal{L}(\mathrm{PAT}(\mathbf{\Omega})) \subset \Theta(\mathrm{PAT}(\mathbf{\Omega}))$ includes the subspace $\mathrm{C}_o^\infty(\mathrm{PAT}(\mathbf{\Omega})\,;\Re)$ of indefinitely differentiable temperature fields with compact support in every open element of the regularity patchwork $\mathrm{PAT}(\mathbf{\Omega})$. This means that, in defining the conforming subspace, linear conditions are imposed only to boundary values of virtual temperatures on the boundary of the elements of $\mathrm{PAT}(\mathbf{\Omega})$,
- the functionals $\mathcal{F}_{\dot{\mathcal{E}}}$, $\mathcal{F}_{\mathcal{M}}$ and $\mathcal{F}_{\mathcal{Q}}$ are defined in terms of densities.

Indeed, let the rate of heat supply $\mathcal{Q}$ be defined in terms of a bulk density q per unit mass and of a superficial density ∂q per unit area. Then, performing the time derivative of the global internal energy, making recourse to the transport theorem and to the principle of conservation of mass in the form $\mathcal{L}_{t,\mathbf{v}}(\rho_t\,\boldsymbol{\mu}) = 0$ and recalling that $\mathbf{\Omega} := \boldsymbol{\varphi}_t(\mathbb{M})$, we have:

$$\begin{aligned}
\dot{\mathcal{E}}_t &:= \partial_{\tau=t}\ \mathcal{E}(\boldsymbol{\varepsilon}_\tau, \mathbf{a}_\tau, \eta_\tau) = \int_{\mathbf{\Omega}} \mathcal{L}_{t,\mathbf{v}}\,(\mathcal{U}_t\,\rho_t\,\boldsymbol{\mu}) = \int_{\mathbf{\Omega}} (\mathcal{L}_{t,\mathbf{v}}\,\mathcal{U}_t)\,\rho_t\,\boldsymbol{\mu}\,,\\
\mathcal{M} &:= \int_{\mathbf{\Omega}} \langle\,\mathbf{T}, \mathrm{sym}\,\partial\mathbf{v}\,\rangle_{\mathbf{g}}\,\boldsymbol{\mu}\,,\\
\mathcal{Q} &:= \int_{\mathbf{\Omega}} q\,\rho\,\boldsymbol{\mu} + \int_{\partial\mathrm{PAT}_q(\mathbf{\Omega})} \partial q\,\partial\boldsymbol{\mu}\,,
\end{aligned}$$

with $\mathrm{PAT}_q(\mathbf{\Omega})$ finer than $\mathrm{PAT}(\mathbf{\Omega})$. Observing that $\mathcal{U}_\tau = \boldsymbol{\varphi}_\tau\!\uparrow\!\mathcal{U}_{0\tau}$, we have that

$$\mathcal{L}_{t,\mathbf{v}}\mathcal{U}_t = \partial_{\tau=t}\ (\boldsymbol{\varphi}_{t,\tau}\!\uparrow\mathcal{U}_\tau) = \partial_{\tau=t}\ (\boldsymbol{\varphi}_{t,\tau}\!\uparrow\!\boldsymbol{\varphi}_\tau\!\uparrow\!\mathcal{U}_\tau) = \boldsymbol{\varphi}_t\!\uparrow\!\partial_{\tau=t}\ \mathcal{U}_{0\tau} = \dot{\mathcal{U}}_{0t}\circ\boldsymbol{\varphi}_t^{-1}\,.$$

The functionals $\mathcal{F}_{\dot{\mathcal{E}}}$, $\mathcal{F}_{\mathcal{M}}$ and $\mathcal{F}_{\mathcal{Q}}$ on $\mathrm{TEMP}(\mathbf{\Omega})$ are then defined by the integrals:

$$\begin{aligned}
\langle\,\mathcal{F}_{\dot{\mathcal{E}}}, \theta\,\rangle &:= \int_{\mathbf{\Omega}} (\mathcal{L}_{t,\mathbf{v}}\,\mathcal{U}_t)\,\theta\,\rho\,\boldsymbol{\mu} = \int_{\mathbf{\Omega}} (\dot{\mathcal{U}}_{0t}\circ\boldsymbol{\varphi}_t^{-1})\,\theta\,\rho\,\boldsymbol{\mu}\,,\\
\langle\,\mathcal{F}_{\mathcal{M}}, \theta\,\rangle &:= \int_{\mathbf{\Omega}} \langle\,\mathbf{T}, \mathrm{sym}\,\partial\mathbf{v}\,\rangle_{\mathbf{g}}\,\theta\,\boldsymbol{\mu}\,,\\
\langle\,\mathcal{F}_{\mathcal{Q}}, \theta\,\rangle &:= \int_{\mathbf{\Omega}} q\,\theta\,\rho\,\boldsymbol{\mu} + \int_{\partial\mathrm{PAT}_q(\mathbf{\Omega})} \partial q\,\mathbf{\Gamma}(\theta)\,\partial\boldsymbol{\mu}\,.
\end{aligned}$$

Defining the bulk energy-gap field as

$$p := \langle \mathbf{T}, \operatorname{sym} \partial \mathbf{v} \rangle_{\mathbf{g}} + \rho \, q - \rho \, \mathcal{L}_{t,\mathbf{v}} \, \mathcal{U}_t \,,$$

the virtual work of the thermal force functional $\mathcal{F}_{\mathcal{G}}$ for a field $\theta \in \mathrm{TEMP}(\mathbf{\Omega})$ of GREEN-regular virtual temperatures, may be written as

$$\langle \mathcal{F}_{\mathcal{G}}, \theta \rangle = \int_{\mathbf{\Omega}} p \, \theta \, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}_q(\mathbf{\Omega})} \partial q \, \mathbf{\Gamma}(\theta) \, \partial \boldsymbol{\mu} \,.$$

The *principle of virtual temperatures* then states:

- The energy conservation law, written as $\langle \mathcal{F}_{\mathcal{G}}, \theta \rangle = 0 \,, \quad \forall \, \theta \in Ker \, \nabla$, is equivalent to assume the existence a vector field $\mathbf{q} \in H(\mathbf{\Omega} \,; V)$ such that

$$\int_{\mathbf{\Omega}} p \, \theta \, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}_q(\mathbf{\Omega})} \partial q \, \mathbf{\Gamma}(\theta) \, \partial \boldsymbol{\mu} = \int_{\mathbf{\Omega}} \mathbf{g}(\mathbf{q}, \nabla \theta) \, \boldsymbol{\mu} \,, \quad \forall \, \theta \in \mathcal{L}(\mathrm{PAT}(\mathbf{\Omega})) \,.$$

6.2.5 Local balance equations

Let us now apply an argument, analogous to that of CAUCHY's theorem in continuum mechanics, to show that any cold-flow vector field $\mathbf{q} \in H(\mathbf{\Omega} \,; V)$, fulfilling the principle of virtual temperatures, is GREEN-regular and admits $\mathrm{PAT}_q(\mathbf{\Omega})$ as regularity support.

This means that the restriction of its distributional divergence to each open element of the patchwork $\mathrm{PAT}_q(\mathbf{\Omega})$ is square integrabile.

To this end we recall that the distributional operator $\mathrm{DIV} = -\mathrm{div}$ is the linear operator which to any $\mathbf{q} \in \mathcal{L}^2(\mathbf{\Omega} \,; V)$ associates the distribution:

$$\langle \mathrm{DIV} \, \mathbf{q}, \theta \rangle := \int_{\mathbf{\Omega}} \mathbf{g}(\mathbf{q}, \nabla \theta) \, \boldsymbol{\mu} \,, \quad \forall \, \theta \in \mathrm{C}_o^{\infty}(\mathbf{\Omega} \,; \Re) \,.$$

If conforming virtual temperature fields belong to $\mathrm{C}_o^{\infty}(\mathrm{PAT}_q(\mathbf{\Omega}) \,; \Re)$, the boundary integral in the principle of virtual temperatures vanishes and we have that:

$$\langle \mathrm{DIV} \, \mathbf{q}, \theta \rangle = \int_{\mathbf{\Omega}} \mathbf{g}(\mathbf{q}, \nabla \theta) \, \boldsymbol{\mu} = \int_{\mathbf{\Omega}} p \, \theta \, \boldsymbol{\mu} \,, \quad \forall \, \theta \in \mathrm{C}_o^{\infty}(\mathrm{PAT}_q(\mathbf{\Omega}) \,; \Re) \,,$$

and the result is proven.

The regular part of $\text{Div} = -\text{div}$ is the formal adjoint ∇'_o of the regular part of the distributional gradient ∇ and is defined, for all $\theta \in \text{Temp}(\mathbf{\Omega})$, by Green's formula:

$$\int_{\mathbf{\Omega}} \mathbf{g}(\mathbf{q}, \nabla\theta)\,\boldsymbol{\mu} = \int_{\mathbf{\Omega}} \nabla'_o \mathbf{q}\,\theta\,\boldsymbol{\mu} + \int_{\partial\text{Pat}_{q\theta}(\mathbf{\Omega})} \mathbf{g}(\mathbf{q}, \mathbf{n})\,\mathbf{\Gamma}(\theta)\,\partial\boldsymbol{\mu}\,,$$

where $\mathbf{n}$ is the outward unit normal and $\text{Pat}_{q\theta}(\mathbf{\Omega})$ is any patchwork finer than $\text{Pat}_q(\mathbf{\Omega})$ and $\text{Pat}_\theta(\mathbf{\Omega})$.

The previous result may be expressed by stating that $\nabla'_o \mathbf{q} = p$.

The *reactive heat supply* is then defined as the difference

$$\langle \mathcal{F}_r, \theta \rangle := \int_{\mathbf{\Omega}} \mathbf{g}(\mathbf{q}, \nabla\theta)\,\boldsymbol{\mu} - \int_{\mathbf{\Omega}} p\,\theta\,\boldsymbol{\mu} - \int_{\partial\text{Pat}_{q\theta}(\mathbf{\Omega})} \partial q\,\mathbf{\Gamma}(\theta)\,\partial\boldsymbol{\mu}\,.$$

Being $\nabla'_o \mathbf{q} = p$, we have that

$$\langle \mathcal{F}_r, \theta \rangle = \int_{\partial\text{Pat}_{q\theta}(\mathbf{\Omega})} (\mathbf{g}(\mathbf{q}, \mathbf{n}) - \partial q)\,\mathbf{\Gamma}(\theta)\,\partial\boldsymbol{\mu}\,, \qquad \forall\,\theta \in \text{Temp}(\mathbf{\Omega})\,.$$

Defining the *reactive boundary heat supply* as $\partial r(\mathbf{q}, \partial q) := \mathbf{g}(\mathbf{q}, \mathbf{n}) - \partial q$, the theorem of virtual thermal work ensures that

$$\int_{\partial\text{Pat}_{q\theta}(\mathbf{\Omega})} \partial r(\mathbf{q}, \partial q)\,\mathbf{\Gamma}(\theta)\,\partial\boldsymbol{\mu} = 0\,, \quad \forall\,\theta \in \mathcal{L}(\text{Pat}(\mathbf{\Omega}))\,.$$

Hence, in particular, $\partial r(\mathbf{q}, \partial q) = 0$ on any piece of boundary where the virtual temperatures are not prescribed to vanish.

With this definition, the principle of virtual temperatures may be written as

$$\int_{\mathbf{\Omega}} p\,\theta\,\boldsymbol{\mu} + \int_{\partial\text{Pat}_{q\theta}(\mathbf{\Omega})} (\partial q + \partial r(\mathbf{q}, \partial q))\,\mathbf{\Gamma}(\theta)\,\partial\boldsymbol{\mu} = \int_{\mathbf{\Omega}} \mathbf{g}(\mathbf{q}, \nabla\theta)\,\boldsymbol{\mu}\,, \quad \forall\,\theta \in \text{Temp}(\mathbf{\Omega})\,,$$

and the corresponding local *balance equations*, analogous to Cauchy's equations of equilibrium, are

$$\begin{aligned} &\nabla'_o \mathbf{q} = p\,, && \text{volumetric heat supply}\,, \\ &\mathbf{g}(\mathbf{q}, \mathbf{n}) = \partial q + \partial r(\mathbf{q}, \partial q) && \text{boundary heat supply}\,. \end{aligned}$$

the former holding in $\mathbf{\Omega}$ and the latter on any $\partial\text{Pat}_{q\theta}(\mathbf{\Omega})$.

- The bulk equation $\nabla_o' \mathbf{q} = p$ is known in literature as the *reduced equation of conservation of the energy.*

Recalling that

$$p := \langle \mathbf{T}, \operatorname{sym} \partial \mathbf{v} \rangle_{\mathbf{g}} + \rho\, q - \rho\, \mathcal{L}_{t,\mathbf{v}} \mathcal{U}_t \,,$$

it may be more explicitly written in the form:

$$\begin{aligned} \rho\, \mathcal{L}_{t,\mathbf{v}} \mathcal{U}_t &= \rho\, q + \langle \mathbf{T}, \operatorname{sym} \partial \mathbf{v} \rangle_{\mathbf{g}} - \nabla_o' \mathbf{q} \\ &= \rho\, q + \langle \mathbf{T}, \operatorname{sym} \partial \mathbf{v} \rangle_{\mathbf{g}} + \operatorname{div} \mathbf{q} \,. \end{aligned}$$

in which $\mathcal{L}_{t,\mathbf{v}} \mathcal{U}_t$ is the total time-derivative of the internal energy per unit mass, q is the rate of heat supply per unit mass, $\operatorname{div} \mathbf{q}$ is the volumetric source of cold, $\langle \mathbf{T}, \operatorname{sym} \partial \mathbf{v} \rangle_{\mathbf{g}}$ it is the mechanical power per unit volume.

- The boundary relation $\partial q = \mathbf{g}(\mathbf{q}, \mathbf{n})$, in absence of reactive boundary heat supply, is known as the *heat flux principle* of Fourier-Stokes.

The existence of a cold-flow vector field is thus a consequence of the principle of conservation of energy, not the object of a further assumption.

Both the previous relations are mathematical results steming from the theorem of virtual thermal work, for thermal boundary value problems.

In thermodynamics it is customary to consider the vector field $-\mathbf{q} \in H(\mathbf{\Omega} \,; V)$ called the *heat flow vector field.* Accordingly, the heat flux principle asserts that the rate of heat supply per unit area of a boundary surface is equal to the flux of the incoming heat flow vector field (outcoming cold flow vector field).

6.2.6 Fourier's Law

Fourier*'s law*, the thermal analog of the linear elastic law in mechanics, is the pointwise constitutive relation of thermal conduction:

$$\mathbf{q}_{\mathbf{m}} = \nabla f_{\mathbf{m}}(\boldsymbol{\theta}_{\mathbf{m}}) \,, \quad \forall\, \mathbf{m} \in \mathbf{\Omega} \,,$$

expressed in terms of a scalar potential $f \in \mathrm{C}^1(V \,; \Re)$ where $\boldsymbol{\theta} \in \mathcal{L}^2(\mathbf{\Omega} \,; V)$ is a square integrable field of *thermal gradients.* A linear behavior is characterized by a strictly convex quadratic potential with a constant hessian, so that

$$\mathbf{q}_{\mathbf{m}} = \mathbf{K}_{\mathbf{m}}\, \boldsymbol{\theta}_{\mathbf{m}} \,, \quad \forall\, \mathbf{m} \in \mathbf{\Omega} \,,$$

with $\mathbf{K}_{\mathbf{m}} = \nabla^2 f_{\mathbf{m}}(\boldsymbol{\theta}_{\mathbf{m}}) \in BL(V \,; V)$ symmetric and positive definite *thermal conductivity* of the material.

Figure 6.2: Jean Baptiste Joseph Fourier (1768 - 1830)

For linear and thermally isotropic materials we have that $\mathbf{K_m} = k_\mathbf{m}\,\mathbf{I}$ and the FOURIER's law takes the form

$$\mathbf{q_m} = k_\mathbf{m}\,\boldsymbol{\theta}_\mathbf{m}\,,$$

where $k > 0$ it is the scalar field of *isotropic thermal conductivity.*

Thermal compatibility imposes that the field of thermal gradients must be the gradient of an admissible temperature field, i.e. $\boldsymbol{\theta} = \nabla\theta = \operatorname{grad}\theta$, where $\theta \in \bar{\theta} + \mathcal{L}(\mathrm{PAT}(\boldsymbol{\Omega}))$ with $\bar{\theta} \in \mathrm{TEMP}(\boldsymbol{\Omega})$ is a prescribed GREEN-regular temperature field. This condition is analogous to the kinematic compatibility for linearized strain fields in classical continuum mechanics.

6.3 Thermal conduction BVP's

FOURIER's linear constitutive law of heat conduction, when inserted in the principle of virtual temperatures, leads to the variational formulation of BVP's of thermal conduction:

$$\int_{\boldsymbol{\Omega}} p\,\delta\theta\,\boldsymbol{\mu} + \int_{\partial\mathrm{PAT}_q(\boldsymbol{\Omega})} \partial q\,\boldsymbol{\Gamma}\delta(\theta)\,\partial\boldsymbol{\mu} = \int_{\boldsymbol{\Omega}} \mathbf{g}(\mathbf{K}(\nabla\theta), \nabla\delta\theta)\,\boldsymbol{\mu}\,, \quad \forall\,\delta\theta \in \mathcal{L}(\mathrm{PAT}(\boldsymbol{\Omega}))\,,$$

in which the test fields are conforming temperature fields $\delta\theta \in \mathcal{L}(\mathrm{PAT}(\boldsymbol{\Omega}))$ and the basic unknown is an admissible temperature field $\theta \in \bar{\theta} + \mathcal{L}(\mathrm{PAT}(\boldsymbol{\Omega}))$ with $\bar{\theta} \in \mathrm{TEMP}(\boldsymbol{\Omega})$ a prescribed GREEN-regular temperature field.

The corresponding local balance equations are given by

$$\begin{aligned} &\nabla'_o(\mathbf{K}(\nabla\theta)) = \langle \mathbf{T}, \operatorname{sym}\partial\mathbf{v}\rangle_\mathbf{g} + \rho\,q - \rho\,\mathcal{L}_{t,\mathbf{v}}\,\mathcal{U}_t\,, && \text{volumetric heat supply}\,,\\ &\mathbf{g}(\mathbf{K}(\nabla\theta), \mathbf{n}) = \partial q + \partial r(\mathbf{K}(\nabla\theta), \partial q)\,, && \text{boundary heat supply}\,. \end{aligned}$$

By the constitutive relations, the material time-derivative of the internal energy density per unit mass in the current placement may be expressed, in terms of stresses and temperature fields in the reference placement, as

$$\begin{aligned}(\mathcal{L}_{t,\mathbf{v}}\,\mathcal{U}_t)(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t) \circ \boldsymbol{\varphi}_t &= \partial_{\tau=t}\, \mathcal{U}_{0t}(\boldsymbol{\varepsilon}_{0\tau}, \mathbf{a}_{0\tau}, \eta_{0\tau}) \\ &= \langle\, d_1\mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0), \dot{\boldsymbol{\varepsilon}}_{0t}\,\rangle \\ &+ \langle\, d_2\mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0), \dot{\mathbf{a}}_{0t}\,\rangle \\ &+ \langle\, d_3\mathcal{U}_{0t}(\boldsymbol{\varepsilon}_0, \mathbf{a}_0, \eta_0), \dot{\eta}_{0t}\,\rangle \\ &= \rho_0^{-1}\langle\, \boldsymbol{\sigma}_{0t}, \dot{\boldsymbol{\varepsilon}}_{0t}\,\rangle + \rho_0^{-1}\langle\, \bar{\boldsymbol{\sigma}}_{0t}, \dot{\mathbf{a}}_{0t}\,\rangle + \theta_{0t}\,\dot{\eta}_{0t}\,,\end{aligned}$$

and, in terms of stresses and temperature fields in the current placement, as

$$\begin{aligned}(\mathcal{L}_{t,\mathbf{v}}\,\mathcal{U}_t)(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t) &= \langle\, d_1\mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t), \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}\rangle \\ &+ \langle\, d_2\mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t), \boldsymbol{\varphi}_{t*}\dot{\mathbf{a}}_{0t}\,\rangle \\ &+ \langle\, d_3\mathcal{U}_t(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \eta_t), \boldsymbol{\varphi}_{t*}\dot{\eta}_{0t}\,\rangle \\ &= \rho_t^{-1}\langle\, \boldsymbol{\sigma}_t, \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}\rangle - \rho_t^{-1} D_t + \theta_t\, \boldsymbol{\varphi}_{t*}\dot{\eta}_{0t}\,,\end{aligned}$$

where we have made recourse to the fact that

$$\boldsymbol{\varphi}_{t*}\dot{\boldsymbol{\varepsilon}}_{0t} = \boldsymbol{\varphi}_{t*}(\boldsymbol{\varphi}_t^*\mathbf{g})^{\cdot} = \mathcal{L}_{\mathbf{v}_t}\mathbf{g}\,,$$

and the anelastic dissipation D_t per unit volume of the current placement is defined by

$$D_t := \rho_t\langle\, \bar{\boldsymbol{\sigma}}_t, \boldsymbol{\varphi}_*\dot{\mathbf{a}}_{0t}\,\rangle = \rho_t\rho_0^{-1}\langle\, \bar{\boldsymbol{\sigma}}_{0t}, \dot{\mathbf{a}}_{0t}\,\rangle \circ \boldsymbol{\varphi}_t^{-1}\,.$$

Being

$$\langle\, \boldsymbol{\sigma}_t, \tfrac{1}{2}\mathcal{L}_{\mathbf{v}}\mathbf{g}\,\rangle = \langle\, \mathbf{T}, \operatorname{sym}\partial\mathbf{v}\,\rangle_{\mathbf{g}}\,,$$

the differential balance equation may be rewritten as

$$\nabla_o'(\mathbf{K}(\nabla\theta)) = D + \rho\, q - \theta_t\, \boldsymbol{\varphi}_{t*}\dot{\eta}_{0t}\,.$$

The Helmholtz free energy $H_t(\boldsymbol{\varepsilon}, \mathbf{a}, \theta)$ is the opposite of the partial Legendre transform of the internal energy $\mathcal{U}_t(\boldsymbol{\varepsilon}, \mathbf{a}, \eta)$ with respect to the entropy:

$$\mathcal{U}_t(\boldsymbol{\varepsilon}, \mathbf{a}, \eta) - H_t(\boldsymbol{\varepsilon}, \mathbf{a}, \theta) = \theta\,\eta\,, \qquad \eta = -d_3H_t(\boldsymbol{\varepsilon}, \mathbf{a}, \theta)\,, \quad \theta = d_3\mathcal{U}_t(\boldsymbol{\varepsilon}, \mathbf{a}, \eta)\,.$$

Hence we have that

$$\begin{aligned} -\dot{\eta} &= \partial_{\tau=t}\, d_3 H(\boldsymbol{\varepsilon}_t, \mathbf{a}_t, \theta_t) \\ &= d_{13}H(\boldsymbol{\varepsilon}, \mathbf{a}, \theta) \cdot \dot{\boldsymbol{\varepsilon}} + d_{23}H(\boldsymbol{\varepsilon}, \mathbf{a}, \theta) \cdot \dot{\mathbf{a}} + d_{33}H(\boldsymbol{\varepsilon}, \mathbf{a}, \theta) \cdot \dot{\theta}\,, \end{aligned}$$

Then, defining the *specific heat at constant strain*:

$$c_v := -\theta\, d_{33}H(\boldsymbol{\varepsilon}, \mathbf{a}, \theta) = \theta\, d_3\eta(\boldsymbol{\varepsilon}, \mathbf{a}, \theta)\,.$$

the differential balance equation takes the form

$$\nabla'_o(\mathbf{K}(\nabla\theta)) = \theta\,(d_{13}H(\boldsymbol{\varepsilon}, \mathbf{a}, \theta) \cdot \dot{\boldsymbol{\varepsilon}} + d_{23}H(\boldsymbol{\varepsilon}, \mathbf{a}, \theta) \cdot \dot{\mathbf{a}}) - c_v \cdot \dot{\theta} + D + \rho\, q\,.$$

By assuming that the thermomechanical interactions, the anelastic dissipation and the bulk heat supply are negligible, the equation takes the simpler form

$$\nabla'_o(\mathbf{K}(\nabla\theta)) = -c_v \cdot \dot{\theta}\,.$$

Further, if the thermal conductivity is linear, isotropic and uniform, we get the classical differential law

$$k\,\boldsymbol{\Delta}\,\theta = c_v \cdot \dot{\theta}\,,$$

where $\boldsymbol{\Delta} = \operatorname{div}\operatorname{grad}$ is the laplacian and k is the thermal conductivity.

6.3.1 Integrability condition

The theorem of virtual thermal work provides a powerful integrability condition to be imposed on a square integrable field of *thermal gradients* $\boldsymbol{\theta} \in \mathcal{L}^2(\boldsymbol{\Omega}\,; V)$ to ensure that there exists an admissible temperature field $\theta \in \bar{\theta} + \mathcal{L}(\mathrm{P{\scriptsize AT}}(\boldsymbol{\Omega}))$ with $\bar{\theta} \in \mathrm{T{\scriptsize EMP}}(\boldsymbol{\Omega})$ a prescribed GREEN-regular temperature field, such that

$$\boldsymbol{\theta} = \nabla\theta\,, \qquad \theta \in \bar{\theta} + \mathcal{L}(\mathrm{P{\scriptsize AT}}(\boldsymbol{\Omega}))\,.$$

This integrability condition may be conveniently rephrased by requiring that there exists a conforming temperature field $\theta \in \mathcal{L}(\mathrm{P{\scriptsize AT}}(\boldsymbol{\Omega}))$ such that

$$\boldsymbol{\theta} - \nabla\bar{\theta} = \nabla_{\mathcal{L}}\theta\,, \qquad \theta \in \mathcal{L}(\mathrm{P{\scriptsize AT}}(\boldsymbol{\Omega}))\,.$$

By KORN's inequality, we know that the range $\operatorname{Im}\nabla_{\mathcal{L}}$ is closed in $H(\boldsymbol{\Omega}\,; V)$ and the kernel $Ker\,\nabla_{\mathcal{L}}$ is finite dimensional

The closedness of the linear subspace $\operatorname{Im}\nabla_{\mathcal{L}} \subset H(\boldsymbol{\Omega}\,; V)$ is equivalent to the orthogonality property:

$$\operatorname{Im}\nabla_{\mathcal{L}} = (Ker\nabla^*_{\mathcal{L}})^{\perp}\,.$$

It follows that the integrability condition is equivalent to the property

$$\boldsymbol{\theta} - \nabla\bar{\theta} \in (Ker\nabla^*_{\mathcal{L}})^{\perp}\,,$$

which can be written in variational terms as

$$\int_{\boldsymbol{\Omega}} \mathbf{g}(\delta\mathbf{q}, \boldsymbol{\theta})\, \boldsymbol{\mu} = \int_{\boldsymbol{\Omega}} \mathbf{g}(\delta\mathbf{q}, \nabla\bar{\theta})\, \boldsymbol{\mu}\,, \quad \forall\, \delta\mathbf{q} \in Ker\, \nabla^*_{\mathcal{L}}\,.$$

In boundary value problems, $\delta\mathbf{q} \in Ker\, \nabla^*_{\mathcal{L}}$ implies that $\nabla'_o \delta\mathbf{q} = 0$. Hence by GREEN's formula we get the equality

$$\int_{\boldsymbol{\Omega}} \mathbf{g}(\delta\mathbf{q}, \nabla\bar{\theta})\, \boldsymbol{\mu} \;=\; \int_{\partial \mathrm{PAT}_{q\theta}(\boldsymbol{\Omega})} \mathbf{g}(\delta\mathbf{q}, \mathbf{n})\, \boldsymbol{\Gamma}\bar{\theta}\, \partial\boldsymbol{\mu}\,,$$

and the integrability condition may be written

$$\int_{\boldsymbol{\Omega}} \mathbf{g}(\delta\mathbf{q}, \boldsymbol{\theta})\, \boldsymbol{\mu} = \int_{\partial \mathrm{PAT}_{q\theta}(\boldsymbol{\Omega})} \mathbf{g}(\delta\mathbf{q}, \mathbf{n})\, \boldsymbol{\Gamma}\bar{\theta}\; \partial\boldsymbol{\mu}\,, \quad \forall\, \delta\mathbf{q} \in Ker\, \nabla^*_{\mathcal{L}}\,.$$

A cold-flow $\mathbf{q} \in \mathcal{L}^2(\boldsymbol{\Omega}\,; V)$ belongs to $Ker\, \nabla^*_{\mathcal{L}}$ iff it fulfills the balance equations or the equivalent principle of virtual temperatures in absence of bulk energy-gap field ($p = 0$) and boundary heat supply ($\partial q = 0$):

$$\int_{\partial \mathrm{PAT}_{q\theta}(\boldsymbol{\Omega})} \partial r(\mathbf{q}, \partial q)\, \boldsymbol{\Gamma}\theta\, \partial\boldsymbol{\mu} = \int_{\boldsymbol{\Omega}} \mathbf{g}(\mathbf{q}, \nabla\theta)\, \boldsymbol{\mu}\,, \quad \forall\, \theta \in \mathrm{TEMP}(\boldsymbol{\Omega})\,.$$

6.3.2 Complementary formulation

The variational integrability condition leads to the following complementary formulation of thermal conduction BVP's.

Let us denote by $\mathcal{Q}_{adm}$ the linear manifold of the admissible cold-flows $\mathbf{q} \in \mathcal{L}^2(\boldsymbol{\Omega}\,; V)$, i.e. the ones fulfilling the principle of virtual temperatures:

$$\int_{\boldsymbol{\Omega}} p\,\theta\, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}_{q}(\boldsymbol{\Omega})} \partial q\, \boldsymbol{\Gamma}(\theta)\, \partial\boldsymbol{\mu} = \int_{\boldsymbol{\Omega}} \mathbf{g}(\mathbf{q}, \nabla\theta)\, \boldsymbol{\mu}\,, \quad \forall\, \theta \in \mathcal{L}(\mathrm{PAT}(\boldsymbol{\Omega}))\,,$$

The solution of a thermal conduction BVP in terms of admissible cold-flows can be sought by requiring that the corresponding thermal gradient $\boldsymbol{\theta} = \mathbf{K}^{-1}\mathbf{q}$, provided by the FOURIER's law of thermal conduction, fulfils the integrability condition and hence that the admissible cold-flow $\mathbf{q} \in \mathcal{Q}_{adm} \subset \mathcal{L}^2(\boldsymbol{\Omega}\,; V)$ be solution of the variational problem:

$$\int_{\boldsymbol{\Omega}} \mathbf{g}(\mathbf{K}^{-1}\mathbf{q}, \delta\mathbf{q})\, \boldsymbol{\mu} = \int_{\partial \mathrm{PAT}_{\partial q\theta}(\boldsymbol{\Omega})} \mathbf{g}(\delta\mathbf{q}, \mathbf{n})\, \boldsymbol{\Gamma}\bar{\theta}\; \partial\boldsymbol{\mu}\,, \quad \forall\, \delta\mathbf{q} \in Ker\, \nabla^*_{\mathcal{L}}\,.$$

6.4 Classical balance laws

The analysis illustrated with reference to the energy conservation principle can be applied to discuss classical *balance laws* of the form

$$\partial_{\tau=t} \int_{\boldsymbol{\xi}_{\tau,t}(\mathbf{C})} a_\tau \, \boldsymbol{\mu} = \int_{\mathbf{C}} b_t \, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}_{c_t}(\mathbf{C})} c_t \, \partial \boldsymbol{\mu} \,,$$

where the *control-volume* $\mathbf{C}$ is any open submanifold which, during a time interval I, flows in an ambient manifold $\mathbb{S}$, dragged by a flow $\boldsymbol{\xi}_{\tau,,t} \in \mathrm{C}^1(\mathbb{S}\,;\mathbb{S})$.
The time dependent scalar field $a_t \in \mathcal{L}^2(\mathbb{S}\,;\Re)$ is a spatial density.
The scalar field $b_t \in \mathcal{L}^2(\mathbf{C}\,;\Re)$ is the *volumetric source*.
The scalar field $c_t \in \mathcal{L}^2(\partial \mathrm{PAT}(\mathbf{C})\,;\Re)$ is the *superficial source*.

By the transport theorem, the *balance law* may be written as

$$\begin{aligned} \partial_{\tau=t} \int_{\boldsymbol{\xi}_{\tau,t}(\mathbf{C})} a_\tau \, \boldsymbol{\mu} = \int_{\mathbf{C}} \mathcal{L}_{t,\mathbf{u}} \, (a_t \, \boldsymbol{\mu}) &= \int_{\mathbf{C}} (\mathcal{L}_{t,\mathbf{u}} \, a_t + a_t \, (\operatorname{div} \mathbf{u})) \, \boldsymbol{\mu} \\ &= \int_{\mathbf{C}} \partial_{\tau=t} \; a_\tau + \operatorname{div} (a \, \mathbf{u})) \, \boldsymbol{\mu} \\ &= \int_{\mathbf{C}} b_t \, \boldsymbol{\mu} + \int_{\partial \mathrm{PAT}_{c_t}(\mathbf{C})} c_t \, \partial \boldsymbol{\mu} \,. \end{aligned}$$

Let $\Lambda_{\mathbf{C}}$ be the space of GREEN-regular scalar test fields, defined by

$$\Lambda_{\mathbf{C}} := \{ \, \lambda \in \mathcal{L}^2(\mathbf{C}\,;\Re) \mid \exists \mathrm{PAT}_\lambda(\mathbf{C}) \, : \, \lambda \in H^1(\mathcal{P}\,;\Re) \,, \quad \forall \, \mathcal{P} \in \mathrm{PAT}_\lambda(\mathbf{C}) \, \} \,.$$

The space $\Lambda_{\mathbf{C}}$ is a pre-HILBERT's space with inner product and norm:

$$(\lambda_1, \lambda_2)_{\Lambda_{\mathbf{C}}} := \int_{\mathbf{C}} \lambda_1 \, \lambda_2 \, \boldsymbol{\mu} + \int_{\mathbf{C}} \mathbf{g}(\nabla \lambda_1, \nabla \lambda_2) \, \boldsymbol{\mu} \,,$$

$$\| \lambda \|^2_{\Lambda_{\mathbf{C}}} := \int_{\mathbf{C}} \lambda^2 \, \boldsymbol{\mu} + \int_{\mathbf{C}} \| \nabla \lambda \|^2 \, \boldsymbol{\mu} \,.$$

The *balance law* may then be written in the form

$$\int_{\mathbf{C}} \lambda \, \mathcal{L}_{t,\mathbf{u}} \, (a_t \, \boldsymbol{\mu}) - \int_{\mathbf{C}} b_t \lambda \; \boldsymbol{\mu} - \int_{\partial \mathrm{PAT}_{c_t \lambda}(\mathbf{C})} c_t \lambda \; \partial \boldsymbol{\mu} = 0 \,, \quad \forall \, \lambda \in Ker \, \nabla \,,$$

where $Ker \, \nabla \subseteq \Lambda_{\mathbf{C}}$ is the linear subspace of piecewise constant test fields.

Let $\Lambda(\mathrm{PAT}(\mathbf{C})) = H^1(\mathrm{PAT}(\mathbf{C})\,;\Re)$ be the subspace of test fields, sharing the patchwork $\mathrm{PAT}(\mathbf{C})$ as common regularity support.

The analysis carried out in section 6.2 and the theorem of virtual thermal work may be rephrased in the present context with no changes.

The theorem ensures that there exists a square integrable vector field $\mathbf{q} \in H(\mathbf{C}\,;V)$ which fulfills, for all $\lambda \in \Lambda(\mathrm{PAT}(\mathbf{C}))$, the *variational balance law*:

$$\int_{\mathbf{C}} \lambda\, \mathcal{L}_{t,\mathbf{u}}\,(a_t\, \boldsymbol{\mu}) \;-\; \int_{\mathbf{C}} b_t\, \lambda\, \boldsymbol{\mu} \;-\; \int_{\partial \mathrm{PAT}_{c_t \lambda}(\mathbf{C})} c_t\, \lambda\, \partial\boldsymbol{\mu} \;=\; \int_{\mathbf{C}} \mathbf{g}(\mathbf{q}_t\,, \nabla\lambda)\, \boldsymbol{\mu}\,.$$

The corresponding *local balance equations* are

$$\begin{aligned} -\mathrm{div}\,\mathbf{q} &= \mathcal{L}_{t,\mathbf{u}}\,(a\,\boldsymbol{\mu}) - b = \mathcal{L}_{t,\mathbf{u}}\, a + a\,\mathrm{div}\,\mathbf{u} - b \\ &= \partial_{\tau=t}\, a_\tau + \mathrm{div}\,(a\,\mathbf{u}) - b\,, \qquad \text{bulk source}\,, \\ \mathbf{g}(\mathbf{q}, \mathbf{n}) &= c\,, \qquad \text{boundary flux}\,. \end{aligned}$$

In books on thermodynamics and mechanics (see e.g. [106]) a balance law is usually written in the form:

$$\partial_{\tau=t} \int_{\boldsymbol{\xi}_\tau(\mathbf{C})} a_\tau\, \boldsymbol{\mu} \;=\; \int_{\mathbf{C}} b_t\, \boldsymbol{\mu} \;+\; \int_{\partial \mathrm{PAT}(\mathbf{C})} \mathbf{g}(\mathbf{q}_t\,, \mathbf{n})\, \partial\boldsymbol{\mu}\,,$$

in which the existence of the vector field $\mathbf{q} \in H(\mathbf{C}\,;V)$ is assumed *a priori*. The previous treatment shows instead that its existence is a result of the theory.

More general balance laws, in which boundary conditions are imposed on the scalar test fields can be dealt with in analogy to the treatment developed in section 6.2.

6.5 Mass-flow vector field

As discussed in section 3.3, the mass conservation principle for a travelling control volume $\mathbf{C}$ is equivalent to the mass-balance variational condition

$$\int_{\mathbf{C}} \lambda\, \mathcal{L}_{t,\mathbf{u}}\,(\rho_t\, \boldsymbol{\mu}) - \int_{\partial\mathbf{C}} \lambda\, \mathbf{g}(\rho_t\,(\mathbf{u}-\mathbf{v}), \mathbf{n})\, \partial\boldsymbol{\mu} = 0\,, \quad \forall\, \lambda \in \mathit{Ker}\, \nabla\,,$$

where $\mathbf{v}$ is the velocity of the motion of the body, $\mathbf{u}$ is the velocity of the travelling control-volume and $\mathit{Ker}\,\nabla \subseteq \Lambda_{\mathbf{C}}$ is the linear space of piecewise constant scalar test fields, a proper subspace of the Green-regular scalar test fields in $\mathbf{C}$. By proceeding in analogy with the theorem of virtual thermal work, we may remove the piecewise constancy constraint on the test fields and state the mass-balance variational condition as a *principle of virtual pressures*:

$$\int_{\mathbf{C}} \lambda\, \mathcal{L}_{t,\mathbf{u}}\,(\rho_t\, \boldsymbol{\mu}) + \int_{\partial \mathrm{PAT}(\mathbf{C})} \lambda\, \mathbf{g}(\rho_t\,(\mathbf{v}-\mathbf{u}), \mathbf{n})\, \partial\boldsymbol{\mu} \;=\; \int_{\mathbf{C}} \mathbf{g}(\mathbf{m}_t\,, \nabla\lambda)\, \boldsymbol{\mu}\,,$$

for all $\lambda \in \mathcal{L}(\mathbf{C})$, where $\mathbf{m}_t \in H(\mathbf{C}\,;V)$ is a a square integrable *mass-flow* vector field over $\mathbf{C}$. The motivation for calling *pressures* the test fields will be given by the example of application in the next subsection.

The corresponding local *mass-balance equations* are

$$\begin{aligned} -\mathrm{div}\,\mathbf{m}_t &= \mathcal{L}_{t,\mathbf{u}}\,(\rho_t\,\boldsymbol{\mu}) = \mathcal{L}_{t,\mathbf{u}}\,\rho_t + \rho_t\,\mathrm{div}\,\mathbf{u}\,, \\ &= \partial_{\tau=t}\,\rho_\tau + \mathrm{div}\,(\rho_t\,\mathbf{u})\,, \quad \text{bulk mass source}\,, \\ \mathbf{g}(\mathbf{m}_t, \mathbf{n}) &= \mathbf{g}(\rho_t\,(\mathbf{v}-\mathbf{u}), \mathbf{n})\,, \quad \text{boundary mass flux}\,. \end{aligned}$$

The flux of the mass-flow vector field thru any surface in the trajectory of the body yields the mass crossing the surface per unit time. If the surface is the boundary of a domain, the flux of the mass-flow vector field yields the mass coming into the domain per unit time.

If the control volume is dragged along the trajectory by the motion of the body, we have that $\boldsymbol{\varphi} = \boldsymbol{\xi}$ and $\mathbf{u} = \mathbf{v}$. Hence the local mass-balance equations become

$$\begin{aligned} -\mathrm{div}\,\mathbf{m}_t &= \mathcal{L}_{t,\mathbf{v}}\,(\rho_t\,\boldsymbol{\mu})\,, && \text{bulk mass source}\,, \\ \mathbf{g}(\mathbf{m}_t, \mathbf{n}) &= 0\,, && \text{boundary mass flux}\,. \end{aligned}$$

Since the flux across any closed surface vanishes, the divergence theorem implies that $\mathcal{L}_{t,\mathbf{v}}\,(\rho_t\,\boldsymbol{\mu}) = 0$ and the rate form of the mass conservation principle is recovered.

6.5.1 Flow thru a porous medium

As an application of the analysis developed in the previous section, let us consider a two-phase medium composed by a fluid phase and by a porous solid skeleton in which a fixed control volume is drawn. Let us assume that the fluid has a stationary flow thru the porous skeleton under prescribed boundary conditions on the pressure field.

We will denote by $\mathrm{A{\scriptsize DM}}(\mathbf{C})$ the manifold of admissible pressure fields fulfilling the nonhomogeneous boundary conditions and by $\mathrm{C{\scriptsize ONF}}(\mathbf{C})$ the linear subspace of pressure fields conforming the related homogeneous boundary conditions. The mass conservation principle, stated as principle of virtual pressures, yields the variational *balance law*:

$$\int_{\partial \mathrm{P{\scriptsize AT}}(\mathbf{C})} \delta\lambda\,\mathbf{g}(\rho\,\mathbf{w}, \mathbf{n})\,\partial\boldsymbol{\mu} \;=\; \int_{\mathbf{C}} \mathbf{g}(\mathbf{m}\,, \nabla\delta\lambda)\,\boldsymbol{\mu}\,,$$

for all virtual pressure fields $\delta\lambda \in \mathcal{L}(\mathbf{C})$. Here now $\mathbf{m} \in H(\mathbf{C}\,; V)$ is a square integrable *fluid mass-flow* vector field over $\mathbf{C}$, ρ is the spatial density of the fluid phase and $\mathbf{w}$ is the velocity of the motion of the fluid phase.

We remark that the rate term $\mathcal{L}_{t,\mathbf{u}}\,(\rho_t\,\boldsymbol{\mu})$ vanishes due to the assumptions that the control volume is fixed in the porous skeleton (so that $\mathbf{u} = 0$) and that the flow of the fluid stationary, i.e. the partial time derivatives vanish.

Let us denote by λ_0 a pressure field in the fluid in static equilibrium. By assuming a Darcy-type permeability law

$$\mathbf{m} = \nabla f(\nabla(\lambda - \lambda_0))\,,$$

governed by a convex potential $f \in \mathrm{C}^1(V\,;\Re)$ describing the nonlinear permeability properties of the medium, we get the variational principle for the evaluation of the pressure in the permeating fluid:

$$\int_{\partial\mathrm{P}\scriptstyle{\mathrm{AT}}(\mathbf{C})} \delta\lambda\,\mathbf{g}(\rho\,\mathbf{w},\mathbf{n})\,\partial\boldsymbol{\mu} \;=\; \int_{\mathbf{C}} \mathbf{g}(\nabla f(\nabla(\lambda - \lambda_0))\,, \nabla\delta\lambda)\,\boldsymbol{\mu}\,, \quad \forall\,\delta\lambda \in \mathcal{L}(\mathbf{C})\,.$$

By introducing the functional F defined by

$$F(\lambda) := \int_{\mathbf{C}} f(\nabla(\lambda - \lambda_0))\,\boldsymbol{\mu} - \int_{\partial\mathrm{P}\scriptstyle{\mathrm{AT}}(\mathbf{C})} \lambda\,\mathbf{g}(\rho\,\mathbf{w},\mathbf{n})\,\partial\boldsymbol{\mu}\,,$$

the principle can be written as a stationarity condition at $\lambda \in \mathcal{L}_{adm}(\mathbf{C})$:

$$\langle dF(\lambda), \delta\lambda\rangle = 0\,, \quad \forall\,\delta\lambda \in \mathcal{L}(\mathbf{C})\,,$$

which is analogous to the stationarity condition for the potential energy of an elastic structure as a functional of the displacement field.

Chapter 7

Elements of Functional Analysis

To provide a direct reference to known results of Functional Analysis, we collect here the most important theorems which are referred to in the paper.

The proof of some result are explicitly reported in the simplest context of HILBERT space theory since they are usually formulated and proved in the more general setting of BANACH spaces with deeper arguments.

7.1 Banach's open mapping and closed range theorems

First we recall the statement of BANACH's open mapping and closed range theorems (see [192] for a general proof in FRÉCHET spaces). A proof of the closed range theorem in HILBERT spaces is provided in [159]. We also report some important consequences of the open mapping theorem and their specialization to HILBERT spaces where the projection theorem and the RIESZ representation theorem provide fundamental analytical tools.

- A linear operator $\mathbf{A} : \mathcal{X} \mapsto \mathcal{Y}$ between two HILBERT spaces is *continuous* if the counter-images under $\mathbf{A}$ of open sets in $\mathcal{Y}$ are open sets in $\mathcal{X}$.

Continuity of linear operators is equivalent to boundedness which means that there exists a contant $C > 0$ such that

$$C \, \|\mathbf{x}\|_{\mathcal{X}} \geq \|\mathbf{A}\mathbf{x}\|_{\mathcal{Y}} \,, \quad \forall \, \mathbf{x} \in \mathcal{X} \,.$$

On the basis of Baire-Hausdorff lemma (see [17], theorem II.1) the polish mathematician Stefan Banach proved a number of celebrated results which provide the foundation of modern Functional Analysis.

Indeed most deep results in functional analysis rely upon the following theorem (see [17] theorem II.5).

Theorem 7.1.1 (The open mapping theorem) *Let $\mathcal{X}$ and $\mathcal{Y}$ be* Banach *spaces and $\mathbf{A} \in BL(\mathcal{X}, \mathcal{Y})$ a continuous linear operator which is surjective. Then there exists a constant $c > 0$ such that*

$$\|\mathbf{y}\|_{\mathcal{Y}} < c \Longrightarrow \exists \, \mathbf{x} \in \mathcal{X} \, : \, \|\mathbf{x}\|_{\mathcal{X}} < 1, \; \mathbf{A}\mathbf{x} = \mathbf{y}.$$

The operator $\mathbf{A}$ will then map open sets of $\mathcal{X}$ onto open sets of $\mathcal{Y}$.

As a corollary it can be proved (see [17] corollary II.6) that the inverse of a continuous one-to-one linear map between two Banach spaces also enjoyes the continuity property.

Theorem 7.1.2 (The continuous inverse theorem) *If a continuous linear operator $\mathbf{A} \in BL(\mathcal{X}, \mathcal{Y})$ establishes a one-to-one map between $\mathcal{X}$ and $\mathcal{Y}$ then the inverse operator is linear and continuous.*

In the sequel the symbol $\langle \bullet, \bullet \rangle$ will denote the duality pairing between dual Hilbert spaces. We recall that, given a closed subspace $\mathcal{A}$ of a Banach space $\mathcal{X}$, the factor space $\mathcal{X}/\mathcal{A}$ is a Banach space when endowed with the norm

$$\|\mathbf{x}_{\mathcal{A}}\|_{\mathcal{X}/\mathcal{A}} := \inf\{ \, \|\mathbf{x} - \overline{\mathbf{x}}\|_{\mathcal{X}} \mid \overline{\mathbf{x}} \in \mathcal{A} \, \}$$

where $\mathbf{x}_{\mathcal{A}}$ denotes the equivalence class $\mathbf{x} + \mathcal{A}$. Let $\mathcal{X}$ be a Hilbert space and $\mathbf{\Pi}_{\mathcal{A}}$ be the orthogonal projector on the closed subspace $\mathcal{A} \subseteq \mathcal{X}$. The factor space $\mathcal{X}/\mathcal{A}$ is a Hilbert space for the inner product

$$(\mathbf{x}_{\mathcal{A}}, \mathbf{y}_{\mathcal{A}})_{\mathcal{X}/\mathcal{A}} := (\mathbf{x} - \mathbf{\Pi}_{\mathcal{A}}\mathbf{x}, \mathbf{y} - \mathbf{\Pi}_{\mathcal{A}}\mathbf{y})_{\mathcal{X}} \quad \forall \, \mathbf{x}_{\mathcal{A}}, \mathbf{y}_{\mathcal{A}} \in \mathcal{X}/\mathcal{A} \, ; \quad \mathbf{x} \in \mathbf{x}_{\mathcal{A}}, \; \mathbf{y} \in \mathbf{y}_{\mathcal{A}}$$

and the associated norm can be written as

$$\|\mathbf{x}_{\mathcal{A}}\|_{\mathcal{X}/\mathcal{A}} = \min\{ \, \|\mathbf{x} - \overline{\mathbf{x}}\|_{\mathcal{X}} \mid \overline{\mathbf{x}} \in \mathcal{A} \, \} = \|\mathbf{x} - \mathbf{\Pi}_{\mathcal{A}}\mathbf{x}\|_{\mathcal{X}} .$$

For every element $\mathbf{x} \in \mathbf{x}_{\mathcal{A}}$ we shall set $\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}} := \|\mathbf{x}_{\mathcal{A}}\|_{\mathcal{X}/\mathcal{A}}$.

Given a pair of HILBERT spaces $\{\mathcal{X}, \mathcal{Y}\}$ a bilinear form $\mathbf{a}(\mathbf{x}, \mathbf{y})$ on $\mathcal{X} \times \mathcal{Y}$ is bounded if for a positive constant C the following inequality holds

$$C\,\|\mathbf{x}\|_{\mathcal{X}}\,\|\mathbf{y}\|_{\mathcal{Y}} \geq |\mathbf{a}(\mathbf{x}, \mathbf{y})| \quad \forall\, \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}.$$

Denoting by $\{\mathcal{X}^*, \mathcal{Y}^*\}$ the spaces in duality with $\{\mathcal{X}, \mathcal{Y}\}$, a pair of bounded linear operators $\mathbf{A} \in BL(\mathcal{X}, \mathcal{Y}^*)$ and $\mathbf{A}^* \in BL(\mathcal{Y}, \mathcal{X}^*)$ can be associated with $\mathbf{a}$ by the identity:

$$\mathbf{a}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{A}^*\mathbf{y}, \mathbf{x} \rangle \quad \forall\, \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}.$$

The discussion of the well-posedness of variational formulations is founded upon the following fundamental result due to BANACH. A proof in BANACH spaces can found in [192], [17] and a proof in HILBERT spaces in [159].

Theorem 7.1.3 (The closed range theorem) *Let us consider a pair* $\{\mathcal{X}, \mathcal{Y}\}$ *of* HILBERT *spaces, a bounded bilinear form* $\mathbf{a}(\mathbf{x}, \mathbf{y})$ *on* $\mathcal{X} \times \mathcal{Y}$ *and the bounded linear operators* $\mathbf{A} \in BL(\mathcal{X}, \mathcal{Y}^*)$ *and* $\mathbf{A}^* \in BL(\mathcal{Y}, \mathcal{X}^*)$ *associated with* $\mathbf{a}$*. Then the following properties are equivalent:*

$$\begin{aligned}
&i) && \operatorname{Im}\mathbf{A} \quad \textit{is closed in } \mathcal{Y}^* && \iff \operatorname{Im}\mathbf{A} = (Ker\,\mathbf{A}^*)^{\perp}, \\
&ii) && \operatorname{Im}\mathbf{A}^* \quad \textit{is closed in } \mathcal{X}^* && \iff \operatorname{Im}\mathbf{A}^* = (Ker\,\mathbf{A})^{\perp}, \\
&iii) && \|\mathbf{A}\mathbf{x}\|_{\mathcal{Y}^*} \geq c\,\|\mathbf{x}\|_{\mathcal{X}/Ker\,\mathbf{A}} && \forall\, \mathbf{x} \in \mathcal{X}, \\
&iv) && \|\mathbf{A}^*\mathbf{y}\|_{\mathcal{X}^*} \geq c\,\|\mathbf{y}\|_{\mathcal{Y}/Ker\,\mathbf{A}^*} && \forall\, \mathbf{y} \in \mathcal{Y},
\end{aligned}$$

where $c > 0$ *is a positive constant.*

Remark 7.1.1 *We recall that, by definition*

$$\|\mathbf{A}\mathbf{x}\|_{\mathcal{Y}^*} := \sup_{\mathbf{y} \in \mathcal{Y}} \frac{\mathbf{a}(\mathbf{x}, \mathbf{y})}{\|\mathbf{y}\|_{\mathcal{Y}}}, \qquad \|\mathbf{A}^*\mathbf{y}\|_{\mathcal{X}^*} := \sup_{\mathbf{x} \in \mathcal{X}} \frac{\mathbf{a}(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\|_{\mathcal{X}}}.$$

These expressions can be modified by observing that being

$$\begin{aligned}
\mathbf{a}(\mathbf{x}, \mathbf{y}) &= \langle \mathbf{A}^*\mathbf{y}, \mathbf{x} \rangle = 0 \quad \forall\, \mathbf{x} \in \mathcal{X}, \quad \forall\, \mathbf{y} \in Ker\,\mathbf{A}^*, \\
\mathbf{a}(\mathbf{x}, \mathbf{y}) &= \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = 0 \quad \forall\, \mathbf{y} \in \mathcal{Y}, \quad \forall\, \mathbf{x} \in Ker\,\mathbf{A}
\end{aligned}$$

we have

$$\|\mathbf{A}\mathbf{x}\|_{\mathcal{Y}^*} = \sup_{\mathbf{y}\in\mathcal{Y}} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{y}\|_{\mathcal{Y}}} = \sup_{\mathbf{y}\in\mathcal{Y}} \sup_{\mathbf{y}_o\in Ker\,\mathbf{A}^*} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{y}+\mathbf{y}_o\|_{\mathcal{Y}}} = \sup_{\mathbf{y}\in\mathcal{Y}} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{y}\|_{\mathcal{Y}/Ker\,\mathbf{A}^*}}\,,$$

and

$$\|\mathbf{A}^*\mathbf{y}\|_{\mathcal{X}^*} = \sup_{\mathbf{x}\in\mathcal{X}} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{x}\|_{\mathcal{X}}} = \sup_{\mathbf{x}\in\mathcal{X}} \sup_{\mathbf{x}_o\in Ker\,\mathbf{A}} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{x}+\mathbf{x}_o\|_{\mathcal{X}}} = \sup_{\mathbf{x}\in\mathcal{X}} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{x}\|_{\mathcal{X}/Ker\,\mathbf{A}}}\,.$$

Since the same constant $c>0$ *appears in iii) and iv) of theorem 7.1.3, these inequalities are easily shown [159] to be equivalent to*

$$\inf_{\mathbf{x}\in\mathcal{X}} \sup_{\mathbf{y}\in\mathcal{Y}} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{x}\|_{\mathcal{X}/Ker\,\mathbf{A}}\|\mathbf{y}\|_{\mathcal{Y}/Ker\,\mathbf{A}^*}} = \inf_{\mathbf{y}\in\mathcal{Y}} \sup_{\mathbf{x}\in\mathcal{X}} \frac{\mathbf{a}\,(\mathbf{x},\mathbf{y})}{\|\mathbf{x}\|_{\mathcal{X}/Ker\,\mathbf{A}}\|\mathbf{y}\|_{\mathcal{Y}/Ker\,\mathbf{A}^*}} > 0\,.$$

which will be referred to as the inf-sup conditions.

When the properties in theorem 7.1.3 hold true, we shall say that the bilinear form $\mathbf{a}$ is closed on $\mathcal{X}\times\mathcal{Y}$.

Theorem 7.1.1 implies the following result concerning the sum of two closed subspaces of a Banach space (see [17] theorem II.8)

Lemma 7.1.1 (A representation lemma) *Let* $\mathcal{X}$ *be a* Banach *space and* $\mathcal{A}\subseteq\mathcal{X}$ *and* $\mathcal{B}\subseteq\mathcal{X}$ *closed subspaces such that their sum* $\mathcal{A}+\mathcal{B}$ *is closed. Then there exists a constant* $c>0$ *such that every* $\mathbf{x}\in\mathcal{A}+\mathcal{B}$ *admits a decomposition of the kind* $\mathbf{x}=\mathbf{a}+\mathbf{b}$ *with* $\mathbf{a}\in\mathcal{A}$, $\mathbf{b}\in\mathcal{B}$, $\|\mathbf{a}\|_{\mathcal{X}}\le c\,\|\mathbf{x}\|_{\mathcal{X}}$ *and* $\|\mathbf{b}\|_{\mathcal{X}}\le c\,\|\mathbf{x}\|_{\mathcal{X}}$.

Proof. By endowing the product space $\mathcal{X}\times\mathcal{X}$ with the norm $\|\{\mathbf{x}\,,\mathbf{y}\}\|_{\mathcal{X}\times\mathcal{X}} := \|\mathbf{x}\|_{\mathcal{X}}+\|\mathbf{y}\|_{\mathcal{X}}$ the linear operator $\mathbf{A}\in BL\,(\mathcal{X}\times\mathcal{X},\mathcal{X})$ defined by $\mathbf{A}\{\mathbf{x}\,,\mathbf{y}\} := \mathbf{x}+\mathbf{y}$ is continuous and surjective. Then by theorem 7.1.1 there exists a constant $c>0$ such that every $\mathbf{x}\in\mathcal{A}+\mathcal{B}$ with $\|\mathbf{x}\|_{\mathcal{X}}<c$ can be written as $\mathbf{x}=\mathbf{a}+\mathbf{b}$ with $\mathbf{a}\in\mathcal{A}$, $\mathbf{b}\in\mathcal{B}$ and $\|\mathbf{a}\|_{\mathcal{X}}+\|\mathbf{b}\|_{\mathcal{X}}<1$. Hence by homogeneity we get that $\mathbf{x}\in\mathcal{A}+\mathcal{B}$ admits the decomposition $\mathbf{x}=\mathbf{a}+\mathbf{b}$ with $\mathbf{a}\in\mathcal{A}$, $\mathbf{b}\in\mathcal{B}$ and $\|\mathbf{a}\|_{\mathcal{X}}+\|\mathbf{b}\|_{\mathcal{X}}\le c^{-1}\,\|\mathbf{x}\|_{\mathcal{X}}$. ■

From lemma 7.1.1 we get a useful characterization of the closedness of the sum of two closed subspaces (see [17] corollary II.9 for a proof in Banach spaces).

Theorem 7.1.4 (The finite angle property) *Let $\mathcal{X}$ be a* Hilbert *space and $\mathcal{A} \subseteq \mathcal{X}$ and $\mathcal{B} \subseteq \mathcal{X}$ closed subspaces such that their sum $\mathcal{A} + \mathcal{B}$ is closed. Then there exists a constant $c > 0$ such that*

$$\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq c \left(\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}} + \|\mathbf{x}\|_{\mathcal{X}/\mathcal{B}} \right) \quad \forall \mathbf{x} \in \mathcal{X}.$$

Proof. Let $\mathbf{a} \in \mathcal{A}$ and $\mathbf{b} \in \mathcal{B}$. Then by lemma 7.1.1 there exist $\overline{\mathbf{a}} \in \mathcal{A}$, $\overline{\mathbf{b}} \in \mathcal{B}$ and a constant $k > 0$ such that

$$\mathbf{a} + \mathbf{b} = \overline{\mathbf{a}} + \overline{\mathbf{b}} \quad \text{and} \quad \|\overline{\mathbf{a}}\|_{\mathcal{X}} \leq k \, \|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}} \quad \|\overline{\mathbf{b}}\|_{\mathcal{X}} \leq k \, \|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}}.$$

By observing that $\mathbf{a} - \overline{\mathbf{a}} = \overline{\mathbf{b}} - \mathbf{b} \in \mathcal{A} \cap \mathcal{B}$ we have that $\forall \mathbf{a} \in \mathcal{A}$ and $\forall \mathbf{b} \in \mathcal{B}$

$$\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \|\mathbf{x} - (\mathbf{a} - \overline{\mathbf{a}})\|_{\mathcal{X}} \leq \|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} + \|\overline{\mathbf{a}}\|_{\mathcal{X}} \leq \|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} + k \, \|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}}.$$

We get the result by a further application of the triangle inequality

$$\|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}} \leq \|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} + \|\mathbf{x} - \mathbf{b}\|_{\mathcal{X}},$$

taking the infimum with respect to $\mathbf{a} \in \mathcal{A}$ and $\mathbf{b} \in \mathcal{B}$ and setting $c = k + 1$. ■

Fig. 7.1 provides a geometrical interpretation of proposition 7.1.4.

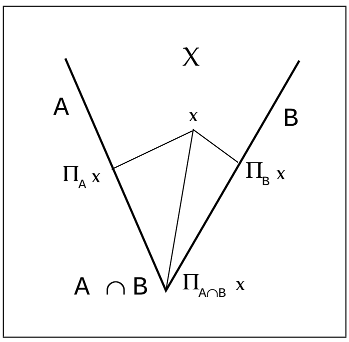

Figure 7.1: A geometrical interpretation of the finite angle property.

$$\|\mathbf{x} - \mathbf{\Pi}_{\mathcal{A}}\mathbf{x}\|_{\mathcal{X}} + \|\mathbf{x} - \mathbf{\Pi}_{\mathcal{B}}\mathbf{x}\|_{\mathcal{X}} \geq c^{-1} \, \|\mathbf{x} - \mathbf{\Pi}_{\mathcal{A}\cap\mathcal{B}}\mathbf{x}\|_{\mathcal{X}} \quad \forall \mathbf{x} \in \mathcal{X}\,.$$

The following lemma provides two basic orthogonality relations.

Lemma 7.1.2 (Orthogonality relations) *Let $\mathcal{A}$ and $\mathcal{B}$ be two subspaces of an* HILBERT *space $\mathcal{X}$, and $\mathcal{A}^\perp$ and $\mathcal{B}^\perp$ their orthogonal complements in the dual* HILBERT *space $\mathcal{X}^*$. Then*

$$i) \quad (\mathcal{A}+\mathcal{B})^\perp = \mathcal{A}^\perp \cap \mathcal{B}^\perp .$$

If $\mathcal{A}$ and $\mathcal{B}$ are closed subspaces we have also that

$$ii) \quad (\mathcal{A}^\perp+\mathcal{B}^\perp)^\perp = \mathcal{A} \cap \mathcal{B}.$$

Proof. The equality $i)$ is evident. To prove $ii)$ we observe that $\mathcal{A}\cap\mathcal{B} \subseteq (\mathcal{A}^\perp+\mathcal{B}^\perp)^\perp$ since $\mathbf{x}\in\mathcal{A}\cap\mathcal{B}$ and $\mathbf{f}\in(\mathcal{A}^\perp+\mathcal{B}^\perp)$ implies $\langle \mathbf{f},\mathbf{x}\rangle = 0$. The converse inclusion follows from $\mathcal{A}^\perp \subseteq \mathcal{A}^\perp+\mathcal{B}^\perp$ so that

$$(\mathcal{A}^\perp+\mathcal{B}^\perp)^\perp \subseteq \mathcal{A}^{\perp\perp} = \mathcal{A}\,.$$

Analogously $(\mathcal{A}^\perp+\mathcal{B}^\perp)^\perp \subseteq \mathcal{B}$ and hence $(\mathcal{A}^\perp+\mathcal{B}^\perp)^\perp \subseteq \mathcal{A}\cap\mathcal{B}$. ■

Remark 7.1.2 *We recall that in a* HILBERT *space we have $\mathcal{A}^{\perp\perp} = \overline{\mathcal{A}}$ where $\overline{\mathcal{A}}$ denotes the closure of $\mathcal{A}$. Given any pair of subspaces $\mathcal{A},\mathcal{B}\subseteq\mathcal{X}$ we have the inclusion $(\mathcal{A}\cap\mathcal{B})^\perp \supseteq \mathcal{A}^\perp+\mathcal{B}^\perp$. If $\mathcal{A}$ and $\mathcal{B}$ are closed, we get an equality if and only if $\mathcal{A}^\perp+\mathcal{B}^\perp$ is a closed subspace of $\mathcal{X}^*$. In fact from property ii) of lemma 7.1.2 we infer that $(\mathcal{A}\cap\mathcal{B})^\perp = (\mathcal{A}^\perp+\mathcal{B}^\perp)^{\perp\perp}$.*

Further, by property i) of lemma 7.1.2, any pair of subspaces $\mathcal{A},\mathcal{B}\subseteq\mathcal{X}$ will meet the relation $(\mathcal{A}+\mathcal{B})^{\perp\perp} = (\mathcal{A}^\perp\cap\mathcal{B}^\perp)^\perp$. Hence the equality $\mathcal{A}+\mathcal{B} = (\mathcal{A}^\perp\cap\mathcal{B}^\perp)^\perp$ holds if and only if $\mathcal{A}+\mathcal{B}$ is closed in $\mathcal{X}$.

A useful criterion for the closedness of the sum of two closed subspaces is provided by the next result [159].

Lemma 7.1.3 (Closedness of the sum of closed subspaces) *Let $\mathcal{A}$ and $\mathcal{B}$ be closed subspaces of a* HILBERT *space $\mathcal{X}$ with one of them finite dimensional. Then the subspace $\mathcal{A}+\mathcal{B}$ is closed.*

We can now prove a deep result (see [17] theorem II.15 for a proof valid in BANACH's spaces).

Theorem 7.1.5 (Closedness of the sum of orthogonal complements) *Let us consider two closed subspaces $\mathcal{A}$ and $\mathcal{B}$ of an* HILBERT *space $\mathcal{X}$, and their orthogonal complements $\mathcal{A}^\perp$ and $\mathcal{B}^\perp$ in the dual* HILBERT *space $\mathcal{X}^*$. Then $\mathcal{A}+\mathcal{B}$ is closed in $\mathcal{X}$ if and only if $\mathcal{A}^\perp+\mathcal{B}^\perp$ is closed in $\mathcal{X}^*$.*

Proof. By virtue of lemma 7.1.2 the following equivalences hold true:

$$i) \quad \mathcal{A}+\mathcal{B} \quad \text{closed} \quad \Longleftrightarrow \quad ii) \quad \mathcal{A}+\mathcal{B} = (\mathcal{A}+\mathcal{B})^{\perp\perp} = (\mathcal{A}^{\perp} \cap \mathcal{B}^{\perp})^{\perp} \,,$$

$$iii) \quad \mathcal{A}^{\perp}+\mathcal{B}^{\perp} \quad \text{closed} \Longleftrightarrow \quad iv) \quad \mathcal{A}^{\perp}+\mathcal{B}^{\perp} = (\mathcal{A}^{\perp}+\mathcal{B}^{\perp})^{\perp\perp} = (\mathcal{A} \cap \mathcal{B})^{\perp} \,.$$

Let us now show that $i) \implies iv)$.

Being $(\mathcal{A} \cap \mathcal{B})^{\perp} = (\mathcal{A}^{\perp}+\mathcal{B}^{\perp})^{\perp\perp} \supseteq \mathcal{A}^{\perp}+\mathcal{B}^{\perp}$ it suffices to prove the converse inclusion $(\mathcal{A} \cap \mathcal{B})^{\perp} \subseteq \mathcal{A}^{\perp}+\mathcal{B}^{\perp}$.

Since $\mathcal{A}+\mathcal{B}$ is closed, lemma 7.1.1 ensures that there exists a constant $c > 0$ such that any $\mathbf{x} \in \mathcal{A}+\mathcal{B}$ admits a decomposition of the kind:

$$\mathbf{x} = \mathbf{a}+\mathbf{b} \quad \text{with} \quad \mathbf{a} \in \mathcal{A} \,, \quad \mathbf{b} \in \mathcal{B} \,,$$

and

$$\|\mathbf{a}\|_{\mathcal{X}} \leq \; c \, \|\mathbf{x}\|_{\mathcal{X}} \,, \quad \|\mathbf{b}\|_{\mathcal{X}} \leq \; c \, \|\mathbf{x}\|_{\mathcal{X}} \,.$$

Now let $\mathbf{f} \in (\mathcal{A} \cap \mathcal{B})^{\perp}$. Then we may define the linear functional $\boldsymbol{\phi}$ on $\mathcal{A}+\mathcal{B}$:

$$\langle \boldsymbol{\phi}, \mathbf{x} \rangle := \langle \mathbf{f}, \mathbf{a} \rangle \quad \forall \, \mathbf{x} \in \mathcal{A}+\mathcal{B} \,,$$

since the definition is independent of the decomposition of $\mathbf{x}$. Further $\boldsymbol{\phi}$ is continuous since

$$|\langle \boldsymbol{\phi}, \mathbf{x} \rangle| = |\langle \mathbf{f}, \mathbf{a} \rangle| \leq \|\mathbf{f}\|_{\mathcal{X}'} \|\mathbf{a}\|_{\mathcal{X}} \leq \; c \, \|\mathbf{f}\|_{\mathcal{X}'} \|\mathbf{x}\|_{\mathcal{X}} \quad \forall \, \mathbf{x} \in \mathcal{A}+\mathcal{B} \,.$$

Let $\mathbf{\Pi}$ be the orthogonal projector on $\mathcal{A}+\mathcal{B}$ in $\mathcal{X}$. The continuous linear functional $\boldsymbol{\varphi} \in \mathcal{X}^*$ defined by

$$\langle \boldsymbol{\varphi}, \mathbf{x} \rangle := \langle \boldsymbol{\phi}, \mathbf{\Pi x} \rangle \,, \quad \forall \, \mathbf{x} \in \mathcal{X} \,,$$

is such that

$$\boldsymbol{\varphi} \in \mathcal{B}^{\perp} \,, \quad \mathbf{f} - \boldsymbol{\varphi} \in \mathcal{A}^{\perp} \,.$$

The implication $iii) \implies ii)$ is proved in an analogous way. ■

I present here some new results which have been discovered by me in the development of the investigation on mixed problems.

First I quote a variant of theorem 7.1.4 providing an inequality which plays a basic role in the analysis carried out in section 8.6. The result is due to the author [158].

Theorem 7.1.6 (A projection property) *Let* $\mathcal{X}$ *be a* HILBERT *space and* $\mathcal{A} \subseteq \mathcal{X}$ *and* $\mathcal{B} \subseteq \mathcal{X}$ *closed subspaces such that their sum* $\mathcal{A} + \mathcal{B}$ *is closed. Let us further denote by* $\mathbf{\Pi}_{\mathcal{A}}$ *and* $\mathbf{\Pi}_{\mathcal{B}}$ *the orthogonal projectors on* $\mathcal{A}$ *and* $\mathcal{B}$ *in* $\mathcal{X}$. *Then there exist a constant* $k > 0$ *such that*

$$\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}} + k\,\|\mathbf{\Pi}_{\mathcal{A}}\mathbf{x}\|_{\mathcal{X}/\mathcal{B}} \quad \forall\, \mathbf{x} \in \mathcal{X},$$
$$\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \|\mathbf{x}\|_{\mathcal{X}/\mathcal{B}} + k\,\|\mathbf{\Pi}_{\mathcal{B}}\mathbf{x}\|_{\mathcal{X}/\mathcal{A}} \quad \forall\, \mathbf{x} \in \mathcal{X}.$$

Proof. The proof of Theorem 7.1.4 shows that

$$\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} + c\,\|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}} \quad \forall\, \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}.$$

Setting $\mathbf{a} = \mathbf{\Pi}_{\mathcal{A}}\mathbf{x}$ and taking the infimum with respect to $\mathbf{b} \in \mathcal{B}$ we get the first inequality. Setting $\mathbf{b} = \mathbf{\Pi}_{\mathcal{B}}\mathbf{x}$ and taking the infimum with respect to $\mathbf{a} \in \mathcal{A}$ we get the second one. ■

A simple geometrical sketch of the previous result is given in Fig.7.2.

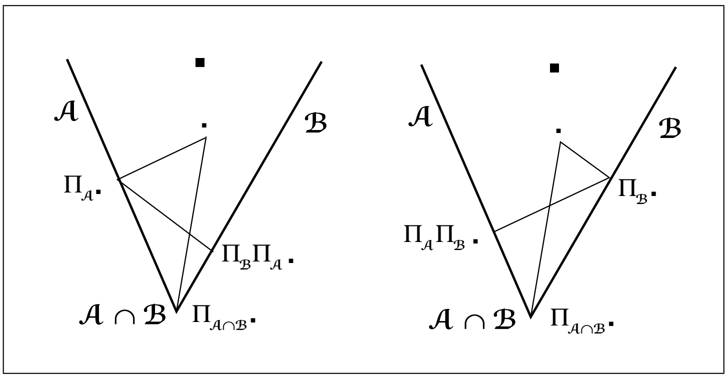

Figure 7.2:

Remark 7.1.3 *For any pair* $\{\mathbf{x}\,,\mathbf{y}\} \in \mathcal{X} \times \mathcal{X}$ *we have*

$$(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)^{1/2} \leq \|\mathbf{x}\| + \|\mathbf{y}\| \leq \sqrt{2}\,(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)^{1/2}$$

and hence the inequalities in propositions 7.1.4 and 7.1 can be rewritten as

$$\|\mathbf{x}\|^2_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \bar{c}\,\big(\|\mathbf{x}\|^2_{\mathcal{X}/\mathcal{A}} + \|\mathbf{x}\|^2_{\mathcal{X}/\mathcal{B}}\big) \qquad \forall\, \mathbf{x} \in \mathcal{X},$$
$$\|\mathbf{x}\|^2_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \bar{k}\,\big(\|\mathbf{x}\|^2_{\mathcal{X}/\mathcal{A}} + \|\mathbf{\Pi}_{\mathcal{A}}\mathbf{x}\|^2_{\mathcal{X}/\mathcal{B}}\big) \quad \forall\, \mathbf{x} \in \mathcal{X},$$
$$\|\mathbf{x}\|^2_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \bar{k}\,\big(\|\mathbf{x}\|^2_{\mathcal{X}/\mathcal{B}} + \|\mathbf{\Pi}_{\mathcal{B}}\mathbf{x}\|^2_{\mathcal{X}/\mathcal{A}}\big) \quad \forall\, \mathbf{x} \in \mathcal{X},$$

with obvious definitions of the constants. These inequalities are the ones directly invoked in our analysis.

From theorem 7.1.5 we derive a useful criterion for the closedness of the image of a product operator.

To this end we premise the following lemma.

Lemma 7.1.4 (An equivalence between closedness properties) *Let $\mathcal{X}$ be a* HILBERT *space and $\mathcal{A}, \mathcal{B}$ subspaces of $\mathcal{X}$ with $\mathcal{B}$ closed. Then $\mathcal{A}+\mathcal{B}$ is closed in $\mathcal{X}$ if and only if the subspace $(\mathcal{A}+\mathcal{B})/\mathcal{B}$ is closed in the factor space $\mathcal{X}/\mathcal{B}$.*

Proof. Let $\mathcal{A}+\mathcal{B}$ be closed in $\mathcal{X}$. Then $\mathcal{A}+\mathcal{B}$ is a HILBERT space for the topology of $\mathcal{X}$ and hence the subspace $(\mathcal{A}+\mathcal{B})/\mathcal{B}$ is closed for the topology of $\mathcal{X}/\mathcal{B}$.

Now let $(\mathcal{A}+\mathcal{B})/\mathcal{B}$ be closed in $\mathcal{X}/\mathcal{B}$. A CAUCHY sequence $\{\mathbf{a}_n + \mathbf{b}_n\}$ with $\mathbf{a}_n \in \mathcal{A}$ and $\mathbf{b}_n \in \mathcal{B}$ will converge to an element $\mathbf{x} \in \mathcal{X}$ and we have to show that $\mathbf{x} \in \mathcal{A}+\mathcal{B}$. First we observe that

$$\|\mathbf{a}_n + \mathbf{b}_n - \mathbf{x}\|_{\mathcal{X}} \geq \inf_{\mathbf{b}\in\mathcal{B}} \|\mathbf{a}_n - \mathbf{x} + \mathbf{b}\|_{\mathcal{X}} = \|\mathbf{a}_n - \mathbf{x}\|_{\mathcal{X}/\mathcal{B}}.$$

Hence by the closedness of $(\mathcal{A}+\mathcal{B})/\mathcal{B}$ the sequence $\{\mathbf{a}_n + \mathcal{B}\} \subset (\mathcal{A}+\mathcal{B})/\mathcal{B}$ will converge to the element $\mathbf{x}+\mathcal{B} \in (\mathcal{A}+\mathcal{B})/\mathcal{B}$. It follows that $\mathbf{x} \in \mathcal{A}+\mathcal{B}$ which was to be proved. ■

We can now state the closedness criterion for the range of the composition of two operators.

Lemma 7.1.5 (Product operators) *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be* HILBERT *spaces and $\mathbf{F} \in BL(\mathcal{X}, \mathcal{Y})$ and $\mathbf{G} \in BL(\mathcal{Y}, \mathcal{Z})$ be continuous linear operators and $\mathbf{F}^* \in BL(\mathcal{Y}^*, \mathcal{X}^*)$ and $\mathbf{G}^* \in BL(\mathcal{Z}^*, \mathcal{Y}^*)$ their duals. Let $\operatorname{Im}\mathbf{F}$ be closed in $\mathcal{Y}$. Then the following equivalence holds*

$$\operatorname{Im}\mathbf{GF} \quad \textit{closed in } \mathcal{Z} \quad \Longleftrightarrow \operatorname{Im}\mathbf{G}^* + Ker\,\mathbf{F}^* \quad \textit{closed in } \mathcal{Y}^*.$$

that is, the image $\operatorname{Im}\mathbf{GF}$ of the product operator $\mathbf{GF} \in BL(\mathcal{X}, \mathcal{Z})$ is closed in $\mathcal{Z}$ if and only if the subspace $\operatorname{Im}\mathbf{G}^* + Ker\,\mathbf{F}^*$ is closed in $\mathcal{Y}^*$.

Proof. Let us consider the operator $\mathbf{G}_o \in BL(\operatorname{Im}\mathbf{F}, \mathcal{Z})$ and its dual ${\mathbf{G}^*}_o \in BL(\mathcal{Z}^*, \mathcal{Y}^*/Ker\,\mathbf{F}^*)$ which are defined by

$$\mathbf{G}_o\mathbf{y} := \mathbf{G}\mathbf{y} \quad \forall\, \mathbf{y} \in \operatorname{Im}\mathbf{F} \qquad {\mathbf{G}^*}_o\mathbf{z}^* := \mathbf{G}^*\mathbf{z}^* + Ker\,\mathbf{F}^* \quad \forall\, \mathbf{z}^* \in \mathcal{Z}.$$

Theorem 7.1.3 shows that $\operatorname{Im}\mathbf{G}_o = \operatorname{Im}\mathbf{GF}$ is closed if and only if $\operatorname{Im}{\mathbf{G}^*}_o = (\operatorname{Im}\mathbf{G}^* + Ker\,\mathbf{F}^*)/Ker\,\mathbf{F}^*$ is closed in $\mathcal{Y}^*/Ker\,\mathbf{F}^*$. By proposition 7.1.4 this property is equivalent to the closedness of $\operatorname{Im}\mathbf{G}^* + Ker\,\mathbf{F}^*$ in $\mathcal{Y}^*$. ■

7.2 Korn's second inequality

The celebrated KORN's second inequality is the milestone along the way that leads to the basic existence results in continuum mechanics and linear elastostatics.

An abstract result by L. TARTAR shows that KORN's inequality implies that the range of the kinematic operator is closed and that its kernel is finite dimensional. A full extension of TARTAR's lemma is provided in this paper and leads to the conclusion that conversely the closedness of the range of the kinematic operator and the finite dimensionality of its kernel are sufficient to ensure the validity of KORN's inequality.

On reading the brilliant proof of KORN's second inequality in the book by G. DUVAUT and J.L. LIONS [39] the author realized that the peculiar form of the sym grad operator plays a basic role in the proof. More specifically he realized that the finite dimensionality of the kernel of sym grad should be a necessary property, although this condition was not appealed to explicitly in the proof.

Some time later the author became aware of a nice result by L. TARTAR concerning an abstract inequality of the KORN's type expressed in term of a bounded linear operator and a compact operator whose kernels have a trivial intersection. TARTAR proved that the inequality implies the finite dimensionality of the kernel and the closedness of the image of the bounded linear operator. The conjecture about the role of the kernel of sym grad in KORN's second inequality was thus confirmed.

At this point it was naturally raised the question whether conversely the finite dimensionality of the kernel of sym grad and the closedness of its image were also sufficient to assess the validity of KORN's second inequality. This converse property requires to complete TARTAR's result with the opposite implication. A full extension of TARTAR's lemma has been provided in the paper [161] and leads to the conclusion that conversely the closedness of the range of the kinematic operator and the finite dimensionality of its kernel are sufficient to ensure the validity of KORN's inequality. The main result contributed here shows that both properties are equivalent to require that a similar inequality be valid for any linear continuous operator.

7.3 Tartar's Lemma

A nice abstract result due to L. TARTAR was reported by F. BREZZI and D. MARINI in [21], lemma 4.1 and quoted by P. G. CIARLET in [29], exer. 3.1.1.

Since TARTAR's lemma plays a basic role in our discussion about KORN's inequality we provide hereafter an explicit proof of this result. Preliminarily we quote that BANACH's open mapping theorem implies the following lemma (see BREZIS [17] th. II.8 and [159], th. 9.1, 9.2).

Theorem 7.3.1 (Bounded decomposition) *Let $\mathcal{X}$ be a* BANACH *space and $\mathcal{A} \subseteq \mathcal{X}$, $\mathcal{B} \subseteq \mathcal{X}$ closed linear subspaces of $\mathcal{X}$ such that their sum $\mathcal{A} + \mathcal{B}$ is closed. Then any $\mathbf{x} \in \mathcal{A} + \mathcal{B}$ admits a decomposition $\mathbf{x} = \mathbf{a} + \mathbf{b}$, with $\mathbf{a} \in \mathcal{A}$, $\mathbf{b} \in \mathcal{B}$, such that*

$$\|\mathbf{x}\|_{\mathcal{X}} \geq c\,\|\mathbf{a}\|_{\mathcal{X}}\,, \quad \|\mathbf{x}\|_{\mathcal{X}} \geq c\,\|\mathbf{b}\|_{\mathcal{X}}\,,$$

where $c > 0$.

If $\mathcal{X} = \mathcal{A} + \mathcal{B}$ and $\mathcal{A} \cap \mathcal{B} = \{\mathbf{o}\}$, the closed subspaces $\mathcal{A}$ and $\mathcal{B}$ are topological supplements in $\mathcal{X}$ and the projectors $\mathbf{P}_{\mathcal{A}}\,\mathbf{x} = \mathbf{a}$ and $\mathbf{P}_{\mathcal{B}}\,\mathbf{x} = \mathbf{b}$ are well defined linear bounded operators from $\mathcal{X}$ to $\mathcal{X}$.

A decomposition $\mathcal{X} = \mathcal{A} \dotplus \mathcal{B}$ of $\mathcal{X}$ into the direct sum of two topological supplementary subspaces $\mathcal{A}$ and $\mathcal{B}$ certainly exists if either $\mathcal{X}$ is a HILBERT space or at least one of them, say $\mathcal{A}$, is finite dimensional.

In the former case $\mathcal{B}$ is simply the orthogonal complement of $\mathcal{A}$ in $\mathcal{X}$. In the latter case we can take as $\mathcal{B}$ the annihilator in $\mathcal{X}$ of a subspace of $\mathcal{X}^*$ generated by fixing a basis in $\mathcal{A}$, taking the dual basis in $\mathcal{A}^*$ and extending its functionals to $\mathcal{X}^*$ (by the HAHN-BANACH theorem).

From Theorem 7.3.1, being $\mathbf{P}_{\mathcal{A}}\,\mathbf{a} = \mathbf{a} \quad \forall\,\mathbf{a} \in \mathcal{A}$, we infer that

$$\|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} \geq c\,\|(\mathbf{x} - \mathbf{a}) - \mathbf{P}_{\mathcal{A}}\,(\mathbf{x} - \mathbf{a})\|_{\mathcal{X}} = c\,\|\mathbf{x} - \mathbf{P}_{\mathcal{A}}\,\mathbf{x}\|_{\mathcal{X}}\,, \quad \forall\,\mathbf{a} \in \mathcal{A}\,, \quad \forall\,\mathbf{x} \in \mathcal{X}\,,$$

which is equivalent to $\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}} \geq c\,\|\mathbf{x} - \mathbf{P}_{\mathcal{A}}\,\mathbf{x}\|_{\mathcal{X}} \quad \forall\,\mathbf{x} \in \mathcal{X}$. Hence we have that

$$\|\mathbf{x} - \mathbf{P}_{\mathcal{A}}\,\mathbf{x}\|_{\mathcal{X}} \geq \|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}} \geq c\,\|\mathbf{x} - \mathbf{P}_{\mathcal{A}}\,\mathbf{x}\|_{\mathcal{X}}\,, \quad \forall\,\mathbf{x} \in \mathcal{X}\,.$$

Theorem 7.3.2 (Tartar's Lemma) *Let H be a reflexive* BANACH *space, E, F be normed linear spaces and $\mathbf{A} \in BL(H, E)$ a bounded linear operator. If there exists a bounded linear operator $\mathbf{L}_o \in BL(H, F)$ such that*

$$\begin{cases} i) & \mathbf{L}_o \in BL(H, F) \quad \text{is compact}\,, \\ ii) & \|\mathbf{A}\mathbf{u}\|_E + \|\mathbf{L}_o\mathbf{u}\|_F \geq \alpha\,\|\mathbf{u}\|_H \quad \forall\,\mathbf{u} \in H\,, \end{cases}$$

then we have that

$$\begin{cases} a) & \dim(Ker\,\mathbf{A}) < +\infty\,, \\ b) & \|\mathbf{A}\mathbf{u}\|_E \geq c_{\mathbf{A}}\,\|\mathbf{u}\|_{H/Ker\,\mathbf{A}} \quad \forall\,\mathbf{u} \in H\,. \end{cases}$$

Proof. Let's prove that the closed linear subspace $Ker\,\mathbf{A} \subset H$ is finite dimensional. We first note that $ii)$ implies that

$$\|\mathbf{L}_o \mathbf{u}\|_F \geq \alpha\, \|\mathbf{u}\|_H \quad \forall\, \mathbf{u} \in Ker\,\mathbf{A}\,.$$

On the other hand, denoting by $\xrightarrow{w}$ the weak convergence in H, the compactness property $i)$ implies that

$$\left.\begin{array}{l} \{\mathbf{u}_n\} \subset Ker\,\mathbf{A}\,, \\ \mathbf{u}_n \xrightarrow{w} \mathbf{u}_\infty \quad \text{in} \quad H\,, \end{array}\right\} \Longrightarrow \|\mathbf{L}_o(\mathbf{u}_n - \mathbf{u}_\infty)\|_F \to 0 \Longrightarrow \|\mathbf{u}_n - \mathbf{u}_\infty\|_H \to 0\,.$$

We may then conclude that every weakly convergent sequence in $Ker\,\mathbf{A}$ is strongly convergent. Hence, by the reflexivity of $\mathcal{H}$ ([17] III.2, remark 4) we must have $\dim(Ker\,\mathbf{A}) < \infty$ and $a)$ is proved.

Then $Ker\,\mathbf{A}$ admits a topological supplement $\mathcal{S}$ and we can consider the bounded linear operator $\mathbf{P}_{\mathbf{A}} \in BL(H,H)$ which is the projector on $Ker\,\mathbf{A}$ subordinated to the decomposition $H = Ker\,\mathbf{A} \dot{+} \mathcal{S}$.

Let us now suppose that $b)$ is false.

There would exists a sequence $\{\mathbf{u}_n\} \subset H$ such that $\|\mathbf{A}\mathbf{u}_n\|_E \to 0$ and $\|\mathbf{u}_n\|_{H/\mathrm{Ker}\,\mathbf{A}} = 1$. By the inequality $\|\mathbf{u}_n\|_{H/\mathrm{Ker}\,\mathbf{A}} \geq c\,\|\mathbf{u}_n - \mathbf{P}_{\mathcal{A}}\,\mathbf{u}_n\|_H$ the sequence $\mathbf{u}_n - \mathbf{P}_{\mathbf{A}}\mathbf{u}_n$ is bounded in H. Hence the compactness of the operator $\mathbf{L}_o \in BL(H,F)$ ensures that we can extract from the sequence $\mathbf{L}_o(\mathbf{u}_n - \mathbf{P}_{\mathbf{A}}\mathbf{u}_n)$ a CAUCHY subsequence $\mathbf{L}_o(\mathbf{u}_k - \mathbf{P}_{\mathbf{A}}\mathbf{u}_k)$ in F.

The sequence $\mathbf{A}\mathbf{u}_k$ is convergent in E by assumption and hence we infer from $ii)$ that $\mathbf{u}_k - \mathbf{P}_{\mathbf{A}}\mathbf{u}_k$ is a CAUCHY sequence which by the completeness of H converges to an element $\mathbf{u}_\infty \in H$. Since $\mathbf{A}\mathbf{u}_k$ converges to zero in E the boundedness of $\mathbf{A} \in BL(H,E)$ ensures that $\mathbf{u}_\infty \in Ker\,\mathbf{A}$ so that also $\mathbf{P}_{\mathbf{A}}\mathbf{u}_k + \mathbf{u}_\infty \in Ker\,\mathbf{A}$. Finally from $ii)$ we get that

$$\alpha\,\|\mathbf{u}_k\|_{H/\mathrm{Ker}\,\mathbf{A}} \leq \|\mathbf{A}\mathbf{u}_k\|_E + \|\mathbf{L}_o(\mathbf{u}_k - \mathbf{P}_{\mathbf{A}}\mathbf{u}_k - \mathbf{u}_\infty)\|_F \to 0\,,$$

and this is absurd since $\|\mathbf{u}_k\|_{H/\mathrm{Ker}\,\mathbf{A}} = 1$. ■

Remark 7.3.1 TARTAR*'s lemma is quoted in [29] referring to [21] for the proof of the statement. Although in [21] and [29] the space H was assumed to be a (non reflexive)* BANACH *space, property $a)$ cannot be inferred in this general context. A well-known counterexample is provided by the space l^1 of absolutely convergent real sequences. In fact* SHUR'S *theorem states that in this infinite dimensional* BANACH *space every weakly convergent sequence is also strongly convergent (see [192] V.1 theorem 5 and [17] III.2, remark 4).*

We also note that the proof of property $b)$ *, as developed in [21], requires the existence of a weakly convergent subsequence of a bounded sequence and hence, by the* EBERLEIN-SHMULYAN *theorem (see [192]), the* BANACH *space* H *should be reflexive. The proof of property* $b)$ *proposed here is instead based on a completeness argument which does not require the reflexivity of the* BANACH *space* H *(private communication by prof.* RENATO FIORENZA*).*

7.4 Inverse Lemma

Let us now face the question whether TARTAR's lemma can be completed by assessing the converse implication. A positive answer needs an existence result. We have in fact to prove that properties $a)$ and $b)$ in TARTAR's lemma imply the existence of a compact operator $\mathbf{L}_o \in BL(H,F)$ fulfilling property $ii)$.

Firstly we observe that $ii)$ implies that $Ker\,\mathbf{A} \cap Ker\,\mathbf{L}_o = \{\mathbf{o}\}$. Our strategy consists in relaxing the requests on $\mathbf{L}_o$ by considering at its place any operator $\mathbf{L} \in BL(H,F)$. We then try to establish the inequality

$$\|\mathbf{Au}\|_E + \|\mathbf{Lu}\|_F \geq \alpha_{\mathbf{L}}\, \|\mathbf{u}\|_{H/(\mathrm{Ker}\,\mathbf{A} \cap \mathrm{Ker}\,\mathbf{L})} \quad \forall\, \mathbf{u} \in H$$

for any $\mathbf{L} \in BL(H,F)$. Once this goal has been achieved we can choose $\mathbf{L}$ to be compact and such that $Ker\,\mathbf{A} \cap Ker\,\mathbf{L} = \{\mathbf{o}\}$. We need some preliminary results. From Theorem 7.3.1 we infer the next proposition.

Theorem 7.4.1 (Distance inequalities) *Let* $\mathcal{X}$ *be a* BANACH *space and* $\mathcal{A} \subseteq \mathcal{X}$ *,* $\mathcal{B} \subseteq \mathcal{X}$ *closed linear subspaces of* $\mathcal{X}$ *such that their sum* $\mathcal{A} + \mathcal{B}$ *is closed. Then, setting* $k = c^{-1} > 0$ *we have*

$$i) \quad \|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} + k\,\|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}}\,, \quad \mathbf{x} \in \mathcal{X}\,, \quad \forall\, \{\mathbf{a}\,,\mathbf{b}\} \in \mathcal{A} \times \mathcal{B}\,.$$

If $\mathcal{A}$ *admits a topological supplement* $\mathcal{S}$ *so that* $\mathcal{X} = \mathcal{A} \dotplus \mathcal{S}$ *then we infer that*

$$ii) \quad \|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \|\mathbf{x} - \mathbf{P}_{\mathcal{A}}\,\mathbf{x}\|_{\mathcal{X}} + k\,\|\mathbf{P}_{\mathcal{A}}\,\mathbf{x}\|_{\mathcal{X}/\mathcal{B}}\,, \quad \mathbf{x} \in \mathcal{X}\,.$$

where $\mathbf{P}_{\mathcal{A}}$ *is the projector on* $\mathcal{A}$ *subordinated to the direct sum decomposition of* $\mathcal{X}$ *.*

Proof. Theorem 7.3.1 ensures that for every $\mathbf{x} \in \mathcal{X}$, $\mathbf{a} \in \mathcal{A}$, $\mathbf{b} \in \mathcal{B}$ there exists a $\boldsymbol{\rho} \in \mathcal{A} \cap \mathcal{B}$ such that $\|\mathbf{a} + \boldsymbol{\rho}\|_{\mathcal{X}} \leq k\,\|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}}$. Hence we infer $i)$:

$$\|\mathbf{x}\|_{\mathcal{X}/\mathcal{A}\cap\mathcal{B}} \leq \|\mathbf{x} + \boldsymbol{\rho}\|_{\mathcal{X}} \leq \|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} + \|\mathbf{a} + \boldsymbol{\rho}\|_{\mathcal{X}} \leq \|\mathbf{x} - \mathbf{a}\|_{\mathcal{X}} + k\,\|\mathbf{a} + \mathbf{b}\|_{\mathcal{X}}.$$

Setting $\mathbf{a} = \mathbf{P}_{\mathcal{A}}\,\mathbf{x}$ and taking the infimum with respect to $\mathbf{b} \in \mathcal{B}$ we get the inequality $ii)$. ■

The following two lemmas yield the tools for the main result.

Theorem 7.4.2 (Projection inequality) *Let* H *be a* BANACH *space and* E *,* F *be linear normed spaces. Let moreover* $\mathbf{A} \in BL(H, E)$ *e* $\mathbf{L} \in BL(H, F)$ *be linear bounded operators such that*

$$\begin{cases} i) \quad \|\mathbf{A}\mathbf{u}\|_E \geq c_{\mathbf{A}} \|\mathbf{u}\|_{H/Ker\,\mathbf{A}}, \quad \forall\, \mathbf{u} \in H, \\ ii) \quad \|\mathbf{L}\mathbf{u}\|_F \geq c_{\mathbf{L}} \|\mathbf{u}\|_{H/Ker\mathbf{L}}, \quad \forall\, \mathbf{u} \in Ker\,\mathbf{A}. \end{cases}$$

Let moreover $Ker\,\mathbf{A}$ *admit a topological supplement* $\mathcal{S}$ *so that* $H = Ker\,\mathbf{A} \dot{+} \mathcal{S}$ *. Then we have*

$$a) \quad \|\mathbf{A}\mathbf{u}\|_E + \|\mathbf{L}\,\mathbf{u}\|_F \geq \alpha \|\mathbf{P}_{\mathbf{A}}\mathbf{u}\|_{H/Ker\mathbf{L}}, \quad \forall\, \mathbf{u} \in H.$$

where $\mathbf{P}_{\mathbf{A}} \in BL(H, H)$ *is the projector on* $Ker\,\mathbf{A}$ *subordinated to the decomposition* $H = Ker\,\mathbf{A} \dot{+} \mathcal{S}$ *.*

Proof. If $a)$ would be false we could find a sequence $\{\mathbf{u}_n\} \subset H$ such that

$$\|\mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_{H/\mathrm{Ker}\,\mathbf{L}} = 1, \quad \|\mathbf{A}\mathbf{u}_n\|_E \to 0, \quad \|\mathbf{L}\mathbf{u}_n\|_F \to 0.$$

Since $\|\mathbf{u}\|_{H/\mathrm{Ker}\,\mathbf{A}} \geq c \|\mathbf{u} - \mathbf{P}_{\mathcal{A}}\,\mathbf{u}\|_H \quad \forall\, \mathbf{u} \in H$ we infer from $i)$ that

$$\|\mathbf{A}\mathbf{u}_n\|_E \to 0 \implies \|\mathbf{u}_n - \mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_H \to 0.$$

Moreover we have

$$\begin{cases} \|\mathbf{L}\|\, \|\mathbf{u}_n - \mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_H \geq \|\mathbf{L}(\mathbf{u}_n - \mathbf{P}_{\mathbf{A}}\mathbf{u}_n)\|_F, \\ \|\mathbf{L}\mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_F \leq \|\mathbf{L}(\mathbf{u}_n - \mathbf{P}_{\mathbf{A}}\mathbf{u}_n)\|_F + \|\mathbf{L}\mathbf{u}_n\|_F. \end{cases}$$

Hence $\|\mathbf{L}\mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_F \to 0$ and from $ii)$ we get

$$\|\mathbf{L}\mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_F \geq c_{\mathbf{L}} \|\mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_{H/\mathrm{Ker}\,\mathbf{L}} \implies \|\mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_{H/\mathrm{Ker}\,\mathbf{L}} \to 0,$$

which is absurd since $\|\mathbf{P}_{\mathbf{A}}\mathbf{u}_n\|_{H/\mathrm{Ker}\,\mathbf{L}} = 1$. ■

Theorem 7.4.3 (Abstract inequality) *Let* H *be a* BANACH *space and* E *,* F *be linear normed spaces. Let moreover* $\mathbf{A} \in BL(H, E)$ *e* $\mathbf{L} \in BL(H, F)$ *be linear bounded operators such that*

$$\begin{cases} i) \quad \|\mathbf{A}\mathbf{u}\|_E \geq c_{\mathbf{A}} \|\mathbf{u}\|_{H/Ker\,\mathbf{A}}, \quad \forall\, \mathbf{u} \in H, \\ ii) \quad \|\mathbf{L}\mathbf{u}\|_F \geq c_{\mathbf{L}} \|\mathbf{u}\|_{H/Ker\mathbf{L}}, \quad \forall\, \mathbf{u} \in Ker\,\mathbf{A}, \\ iii) \quad Ker\,\mathbf{A} + Ker\,\mathbf{L} \quad \text{closed in} \quad H. \end{cases}$$

Let moreover $Ker\,\mathbf{A}$ *admit a topological supplement* $\mathcal{S}$ *so that* $H = Ker\,\mathbf{A} \dot{+} \mathcal{S}$ *. Then we have*

$$c) \quad \|\mathbf{A}\mathbf{u}\|_E + \|\mathbf{L}\mathbf{u}\|_F \geq \alpha\,\|\mathbf{u}\|_{H/(Ker\,\mathbf{A} \cap Ker\,\mathbf{L})} .$$

Proof. Summing up the inequalities $a)$ and $i)$ in Theorem 7.4.2 we get

$$\|\mathbf{A}\mathbf{u}\|_E + \|\mathbf{L}\,\mathbf{u}\|_F \geq \alpha_o\,(\,\|\mathbf{u}\|_{H/\mathrm{Ker}\,\mathbf{A}} + \|\mathbf{P_A}\mathbf{u}\|_{H/\mathrm{Ker}\,\mathbf{L}}\,)\,, \quad \forall\,\mathbf{u} \in H\,.$$

Moreover, by assumption $iii)$, Theorem 7.4.1 implies that

$$\|\mathbf{u} - \mathbf{P_A}\mathbf{u}\|_H + k\,\|\mathbf{P_A}\mathbf{u}\|_{H/\mathrm{Ker}\,\mathbf{L}} \geq c\,\|\mathbf{u}\|_{H/\mathrm{Ker}\,\mathbf{A} \cap \mathrm{Ker}\,\mathbf{L}}\,, \quad \forall\,\mathbf{u} \in H\,.$$

Recalling that $\|\mathbf{u}\|_{H/\mathrm{Ker}\,\mathbf{A}} \geq c\,\|\mathbf{u} - \mathbf{P}_{\mathcal{A}}\,\mathbf{u}\|_H \quad \forall\,\mathbf{u} \in H$ we get the result. ■

The next lemma yields the crucial result for our analysis.

Theorem 7.4.4 (Inverse lemma) *Let* H *be a* Banach *space and* E *,* F *be linear normed spaces. Let moreover* $\mathbf{A} \in BL\,(H, E)$ *be a linear bounded operator such that*

$$\begin{cases} a) \quad \dim Ker\,\mathbf{A} < +\infty\,, \\ b) \quad \|\mathbf{A}\mathbf{u}\|_E \geq c_{\mathbf{A}}\,\|\mathbf{u}\|_{H/Ker\,\mathbf{A}}\,, \quad \forall\,\mathbf{u} \in H\,. \end{cases}$$

Then for any $\mathbf{L} \in BL\,(H, F)$ *we have*

$$i) \quad \|\mathbf{A}\mathbf{u}\|_E + \|\mathbf{L}\mathbf{u}\|_F \geq \alpha\,\|\mathbf{u}\|_{H/(Ker\,\mathbf{A} \cap Ker\,\mathbf{L})}\,, \quad \forall\,\mathbf{u} \in H\,.$$

Proof. It suffices to observe that any finite dimensional subspace admits a topological supplement in H and that condition $a)$ implies the validity of $ii)$ and $iii)$ of the **abstract inequality** for any $\mathbf{L} \in BL\,(H, F)$. ■

Now we recall that any continuous projection operator on a finite dimensional subspace is compact.

It follow that if $\dim Ker\,\mathbf{A} < +\infty$ there exists at least a compact operator $\mathbf{L}_o \in BL\,(H, F)$ such that $Ker\,\mathbf{A} \cap Ker\,\mathbf{L}_o = \{\mathbf{o}\}$. Indeed we can set $\mathbf{L}_o = \mathbf{P_A} \in BL\,(H, H)$, the projection operator on the finite dimensional subspace $Ker\,\mathbf{A} \subset H$ defined by a direct sum decomposition $H = (Ker\,\mathbf{A}) \dot{+} \mathcal{S}$ with $\mathcal{S}$ topological supplement of $Ker\,\mathbf{A}$.

We can now provide a full extension of Tartar's lemma by including the converse implication and the equivalence to a new property.

Theorem 7.4.5 (Equivalent inequalities) *Let* H *be a reflexive* BANACH *space,* E , F *be normed linear spaces and* $\mathbf{A} \in BL(H,E)$ *a bounded linear operator. Then the following propositions are equivalent:*

$$P_1) \quad \begin{cases} \dim Ker\,\mathbf{A} < +\infty\,, \\ \|\mathbf{A}\mathbf{u}\|_E \geq c_{\mathbf{A}}\, \|\mathbf{u}\|_{H/Ker\,\mathbf{A}}\,, \quad \forall\, \mathbf{u} \in H\,, \end{cases}$$

$$P_2) \quad \begin{cases} \textit{There exists } \mathbf{L}_o \in BL(H,F) \textit{ compact} \\ \textit{such that } Ker\,\mathbf{A} \cap Ker\,\mathbf{L}_o = \{\mathbf{o}\} \textit{ and} \\ \|\mathbf{A}\mathbf{u}\|_E + \|\mathbf{L}_o\mathbf{u}\|_F \geq \alpha\, \|\mathbf{u}\|_H\,, \quad \forall\, \mathbf{u} \in H\,, \end{cases}$$

$$P_3) \quad \begin{cases} \dim Ker\,\mathbf{A} < +\infty\,, \\ \|\mathbf{A}\mathbf{u}\|_E + \|\mathbf{L}\mathbf{u}\|_F \geq \alpha\, \|\mathbf{u}\|_{H/(Ker\,\mathbf{A} \cap Ker\,\mathbf{L})}\,, \quad \forall\, \mathbf{u} \in H\,, \forall\, \mathbf{L} \in BL(H,F)\,, \end{cases}$$

Proof. $P_3 \implies P_1$ setting $\mathbf{L} = \mathbf{O}$. $P_3 \implies P_2$ setting $\mathbf{L} = \mathbf{L}_o = \mathbf{P}_{\mathbf{A}}$. $P_1 \implies P_3$ by Lemma 7.4.4. Finally $P_2 \implies P_1$ by TARTAR's lemma which is the one requiring the reflexivity of the BANACH space H. ■

7.5 Korn's Inequality

In continuum mechanics the fundamental theorems concerning the variational formulation of equilibrium and compatibility are founded on the property that the kinematic operator has a closed range and a finite dimensional kernel. The abstract framework is the following. A structural model is defined on a regular bounded domain $\mathbf{\Omega}$ of an euclidean space and is governed by a kinematic operator $\mathbf{B}$ which is the regular part of a distributional differential operator $\mathbb{B} : \mathcal{V}(\mathbf{\Omega}) \mapsto \mathbb{D}'(\mathbf{\Omega})$ of order m acting on kinematic fields $\mathbf{u} \in \mathcal{V}(\mathbf{\Omega})$ which are square integrable on $\mathbf{\Omega}$ and such that the corresponding distributional linearized strain field $\mathbb{B}\mathbf{u} \in \mathbb{D}'(\mathbf{\Omega})$ is square integrable on a finite subdivision $\mathcal{T}_{\mathbf{u}}(\mathbf{\Omega})$ of $\mathbf{\Omega}$. The kinematic space $\mathcal{V}(\mathbf{\Omega})$ is a pre-HILBERT space when endowed with the topology induced by the norm

$$\|\mathbf{u}\|^2_{\mathcal{V}(\mathbf{\Omega})} = \|\mathbf{u}\|^2_{H(\mathbf{\Omega})} + \|\mathbf{B}\mathbf{u}\|^2_{\mathcal{H}(\mathbf{\Omega})}\,,$$

where $H(\mathbf{\Omega})$ and $\mathcal{H}(\mathbf{\Omega})$ are the spaces of kinematic and linearized strain fields which are square integrable on $\mathbf{\Omega}$ [160]. The conforming kinematisms $\mathbf{u} \in \mathcal{L}(\mathbf{\Omega})$ belong to a closed linear subspace $\mathcal{L}(\mathbf{\Omega}) \subset H^m(\mathcal{T}(\mathbf{\Omega})) \subset \mathcal{V}(\mathbf{\Omega})$ of the SOBOLEV space $H^m(\mathcal{T}(\mathbf{\Omega}))$, where $\mathcal{T}(\mathbf{\Omega})$ is a given finite subdivision of $\mathbf{\Omega}$. Thus $\mathcal{L}(\mathbf{\Omega}) \subset H^m(\mathcal{T}(\mathbf{\Omega}))$ is an HILBERT space and the operator

$\mathbf{B}_{\mathcal{L}} \in BL(\mathcal{L}(\mathbf{\Omega}), \mathcal{H}(\mathbf{\Omega}))$ defining the linearized regular strain $\mathbf{Bu} \in \mathcal{H}(\mathbf{\Omega})$ associated with the conforming kinematic field $\mathbf{u} \in \mathcal{L}(\mathbf{\Omega})$ is linear and continuous. The kinematic operator $\mathbf{B} \in BL(\mathcal{V}(\mathbf{\Omega}), \mathcal{H}(\mathbf{\Omega}))$ is assumed to be regular in the sense that for any $\mathcal{L}(\mathbf{\Omega}) \subset \mathcal{V}(\mathbf{\Omega})$ the following conditions are met [160]

$$\begin{cases} \dim Ker\, \mathbf{B}_{\mathcal{L}} < +\infty\,, \\ \|\mathbf{Bu}\|_{\mathcal{H}(\mathbf{\Omega})} \geq c_{\mathbf{B}}\, \|\mathbf{u}\|_{\mathcal{L}(\mathbf{\Omega})/\mathrm{Ker}\, \mathbf{B}_{\mathcal{L}}}\,, \quad \forall\, \mathbf{u} \in \mathcal{L}(\mathbf{\Omega}) \iff \mathrm{Im}\, \mathbf{B}_{\mathcal{L}} \text{ closed in } \mathcal{H}(\mathbf{\Omega})\,. \end{cases}$$

The requirement that the property must hold for any $\mathcal{L}(\mathbf{\Omega}) \subset \mathcal{V}(\mathbf{\Omega})$ is motivated by the observation that in applications it is fundamental to assess that the basic existence results hold for any choice of the kinematic contraints. The regularity of $\mathbf{B} \in BL(\mathcal{V}(\mathbf{\Omega}), \mathcal{H}(\mathbf{\Omega}))$ is the basic tool for the proof of the theorem of virtual powers which ensures the existence of a stress field in equilibrium with an equilibrated system of active forces.

Theorem 7.5.1 (Theorem of Virtual Powers) *Let* $\mathbf{f} \in \mathcal{L}^*(\mathbf{\Omega})$ *be a system of active forces. Then*

$$\mathbf{f} \in (Ker\, \mathbf{B}_{\mathcal{L}})^{\perp} \implies \exists\, \boldsymbol{\sigma} \in \mathcal{H}(\mathbf{\Omega}) \,:\, \langle \mathbf{f}, \mathbf{v} \rangle = (\!(\, \boldsymbol{\sigma}\,, \mathbf{Bv}\,)\!)\,, \quad \forall\, \mathbf{v} \in \mathcal{L}(\mathbf{\Omega})\,.$$

Proof. Let $\mathbf{B}'_{\mathcal{L}} \in BL(\mathcal{H}(\mathbf{\Omega}), \mathcal{L}^*(\mathbf{\Omega}))$ be the equilibrium operator dual to $\mathbf{B}_{\mathcal{L}}$. By BANACH's closed range theorem we have that $\mathbf{f} \in (Ker\, \mathbf{B}_{\mathcal{L}})^{\perp} = \mathrm{Im}\, \mathbf{B}'_{\mathcal{L}}$ and the duality relation yields the result. ■

A linearized strain field $\boldsymbol{\varepsilon} \in \mathcal{H}(\mathbf{\Omega})$ is kinematically compatible if there exists a conforming kinematic field $\mathbf{u} \in \mathcal{L}(\mathbf{\Omega})$ such that $\boldsymbol{\varepsilon} = \mathbf{Bu}$. Self-equilibrated stress fields are the elements of $\mathcal{H}(\mathbf{\Omega})$ which belong to the kernel of the equilibrium operator $\mathbf{B}'_{\mathcal{L}} \in BL(\mathcal{H}(\mathbf{\Omega}), \mathcal{L}^*(\mathbf{\Omega}))$. The regularity of $\mathbf{B} \in BL(\mathcal{L}(\mathbf{\Omega}), \mathcal{H}(\mathbf{\Omega}))$ provides the following variational condition.

Theorem 7.5.2 (Kinematical compatibility)

$$(\!(\, \boldsymbol{\sigma}\,, \boldsymbol{\varepsilon}\,)\!) = 0 \quad \forall\, \boldsymbol{\sigma} \in Ker\, \mathbf{B}'_{\mathcal{L}} \implies \exists\, \mathbf{u} \in \mathcal{L}(\mathbf{\Omega}) \,:\, \boldsymbol{\varepsilon} = \mathbf{Bu}\,.$$

Proof. By BANACH's closed range theorem $\mathrm{Im}\, \mathbf{B}_{\mathcal{L}} = (Ker\, \mathbf{B}'_{\mathcal{L}})^{\perp}$. ■

The regularity of the kinematic operator $\mathbf{B} \in BL(\mathcal{V}(\mathbf{\Omega}), \mathcal{H}(\mathbf{\Omega}))$ is then a fundamental property to be assessed in a structural model. Our analysis shows that a necessary and sufficient condition is the validity of an inequality of the KORN's type

$$\|\mathbf{Bu}\|_{\mathcal{H}(\mathbf{\Omega})} + \|\mathbf{u}\|_{H(\mathbf{\Omega})} \geq \alpha\, \|\mathbf{u}\|_{H^m(\mathbf{\Omega})}\,, \quad \forall\, \mathbf{u} \in H^m(\mathbf{\Omega})\,,$$

Note that, by RELLICH selection principle [56], the canonical immersion from $H^m(\mathbf{\Omega})$ into $H(\mathbf{\Omega}) = \mathcal{L}^2(\mathbf{\Omega})$ is compact. If KORN's inequality holds for any $\mathbf{u} \in H^m(\mathbf{\Omega})$ it will hold also for any $\mathbf{u} \in H^m(\mathcal{T}(\mathbf{\Omega}))$ and then a fortiori for any $\mathbf{u} \in \mathcal{L}(\mathbf{\Omega})$.

With reference to the three-dimensional continuous model we remark that KORN's first inequality can be easily derived from KORN's second inequality by appealing to Lemma 7.4.4.

In fact denoting by $H^{1/2}(\partial\mathbf{\Omega})^3$, the space of traces of fields in $H^1(\mathbf{\Omega})^3$ on the boundary $\partial\mathbf{\Omega}$ of $\mathbf{\Omega}$ and taking $\mathbf{L}$ to be the boundary trace operator $\mathbf{\Gamma} \in BL\,(H^1(\mathbf{\Omega})^3, H^{1/2}(\partial\mathbf{\Omega})^3)$ we get

$$\|\mathbf{Bu}\|_{\mathcal{H}(\mathbf{\Omega})} + \|\mathbf{\Gamma u}\|_{H^{1/2}(\partial\mathbf{\Omega})^3} \geq \alpha\, \|\mathbf{u}\|_{H^1(\mathbf{\Omega})^3} \quad \forall\, \mathbf{u} \in H^1(\mathbf{\Omega})^3\,,$$

and hence

$$\|\mathbf{Bu}\|_{\mathcal{H}(\mathbf{\Omega})} \geq \alpha\, \|\mathbf{u}\|_{H^1(\mathbf{\Omega})^3} \quad \forall\, \mathbf{u} \in H^1(\mathbf{\Omega})^3 \cap Ker\,\mathbf{\Gamma} = H^1_0(\mathbf{\Omega})^3\,,$$

which is KORN's first inequality. The original form of the second inequality as stated by KORN was in fact

$$\|\mathrm{sym}\,\mathrm{grad}\,\mathbf{u}\|_{\mathcal{L}^2(\mathbf{\Omega})} \geq \alpha\, \|\mathbf{u}\|_{H^1(\mathbf{\Omega})} \quad \forall\, \mathbf{u} \in H^1(\mathbf{\Omega}) \,:\, \int_{\mathbf{\Omega}} \mathrm{emi}\,\mathrm{grad}\,\mathbf{u}\; d\mu = \mathbf{O}\,.$$

By the **inverse lemma** also this original form can be recovered simply by setting

$$\mathbf{L} \in BL\,(H^1(\mathbf{\Omega})^3, \Re^6)\,, \quad \mathbf{Lu} := \int_{\mathbf{\Omega}} \mathrm{emi}\,\mathrm{grad}\,\mathbf{u}\; d\mu\,.$$

We thus get the inequality

$$\|\mathrm{sym}\,\mathrm{grad}\,\mathbf{u}\|_{\mathcal{L}^2(\mathbf{\Omega})} + \left\| \int_{\mathbf{\Omega}} \mathrm{emi}\,\mathrm{grad}\,\mathbf{u}\; d\mu \right\| \geq \alpha\, \|\mathbf{u}\|_{H^1(\mathbf{\Omega})} \quad \forall\, \mathbf{u} \in H^1(\mathbf{\Omega})\,.$$

which immediately implies KORN's original inequality.

The proof of the converse implication is more involved and can be found in G. FICHERA's article [56], remark on page 384. A more detailed version of the proof is provided in [159], Lemma 7.11.

From Lemma 7.4.4 we can also infer POINCARÉ inequality.

Let $\mathbf{\Omega}$ be an open bounded connected set in $\Re^d$ with a regular boundary. Denoting by $\mathbf{p}$ a d-multi-index and by $|\mathbf{p}|$ the sum of its components we set:

- $\mathbf{A} \in BL(H^m(\mathbf{\Omega}), \mathcal{L}^2(\mathbf{\Omega})^k)$ continuous linear operator $\mathbf{Au} = \{D^{\mathbf{p}}\mathbf{u}\}$, with $k = \text{card}\{\mathbf{p} \in \mathcal{N}^d \,:\, |\mathbf{p}| = m\}$ and $|\mathbf{p}| = m$,
- $\mathbf{L}_o \in BL(H^m(\mathbf{\Omega}), H^{m-1}(\mathbf{\Omega}))$ compact identity map $\mathbf{L}_o\mathbf{u} = \mathbf{u}$,
- $\mathbf{L} \in BL(H^m(\mathbf{\Omega}), \mathcal{L}^2(\mathbf{\Omega})^r)$ continuous linear operator defined by

$$\mathbf{Lu} = \left\{ \frac{1}{\sqrt{\text{meas}\,\mathbf{\Omega}}} \int_{\mathbf{\Omega}} D^{\mathbf{p}}\mathbf{u}(\mathbf{x})\, d\mu \right\}; \quad 0 \le |\mathbf{p}| \le m-1,$$

  with $r = \text{card}\{\mathbf{p} \in \mathcal{N}^d \,:\, |\mathbf{p}| < m\}$.

We set $H = H^m(\mathbf{\Omega})$, $E = \mathcal{L}^2(\mathbf{\Omega})^k$, $E_o = H^{m-1}(\mathbf{\Omega})$, $F = \mathcal{L}^2(\mathbf{\Omega})^r$, so that

$$\mathbf{A} \in BL(H, E), \quad \mathbf{L}_o \in BL(H, E_o), \quad \mathbf{L} \in BL(H, F).$$

Then property P_2 of Theorem 7.4.5 is fulfilled since

$$\begin{cases} \|\mathbf{Au}\|_E^2 + \|\mathbf{L}_o\mathbf{u}\|_{E_o}^2 = \|\mathbf{u}\|_H^2, \\ \mathbf{L}_o \in BL(H, E_o) \quad \text{is compact}. \end{cases}$$

We remark that $Ker\,\mathbf{A} = P_{m-1}(\mathbf{\Omega})$ is the finite dimensional linear subspace of polynomials of total degree not greater than $m-1$ so that $\dim P_{m-1}(\mathbf{\Omega}) = (m-1+d)!/(d!\,(m-1)!)$. Moreover we have that

$$Ker\,\mathbf{A} \cap Ker\,\mathbf{L} = \{\mathbf{o}\},$$

and hence property P_3 of Theorem 7.4.5 yields

$$\|\mathbf{Au}\|_E + \|\mathbf{Lu}\|_F \ge \alpha\, \|\mathbf{u}\|_H \quad \forall\, \mathbf{u} \in H,$$

or explicitly, for all $\mathbf{u} \in H^m(\mathbf{\Omega})$:

$$\sum_{|\mathbf{p}|=m} \int_{\mathbf{\Omega}} |D^{\mathbf{p}}\mathbf{u}(\mathbf{x})|^2\, d\mu + \sum_{|\mathbf{p}|<m} \left| \int_{\mathbf{\Omega}} D^{\mathbf{p}}\mathbf{u}(\mathbf{x})\, d\mu \right|^2 \ge \alpha\, \|\mathbf{u}\|^2_{H^m(\mathbf{\Omega})},$$

which is Poincaré inequality.

Remark 7.5.1 *While proof-reading the paper [161] the author became aware of a result, quoted by* Roger Temam *in [182], section I.1, which is a special case of the inverse lemma. This result was not explicitly proved in [182]* *and was resorted to in deriving a proof of* KORN*'s inequality from the property that the distributional operator* grad $\in BL\,(\mathcal{L}^2(\mathbf{\Omega})^n, H^{-1}(\mathbf{\Omega})^{n\times n})$ *has a closed range and a one-dimensional kernel consisting of the constant fields on* $\mathbf{\Omega}$ *(see [159] for an explicit proof). This property is in turn a direct consequence of a fundamental inequality due to* J. NECAS *[126].*

Chapter 8

Linear elastostatics

This chapter is devoted to the theoretical analysis of the elastic equilibrium problem under the assumption of a linear elastic behavior. Existence and uniqueness of the solution are discussed in the context of HILBERT space theory. New results, concerning the closedness of the product of two linear operators and a projection property equivalent to the closedness of the sum of two closed subspaces, are contributed. A set of necessary and sufficient conditions for the well-posedness of an elastic problem with a singular elastic compliance provides the most general result of this kind in linear elasticity. Sufficient criteria for the well-posedness of elastic problems in structural mechanics including the presence of supporting elastic beds are contributed and applications are exemplified.

8.1 Introduction

Mixed formulations in elasticity, in which both the stress and the kinematic fields are taken as basic unknowns of the problem, are motivated either by singularities of the constitutive operators or by computational requirements.

The pioneering contributions by I. BABUŠKA [11] and F. BREZZI [18] have provided mixed formulations leading to saddle-point problems with a sound mathematical foundation. A comprehensive presentation of the state of art can be found in chapter II of [20] where existence and uniqueness results and *a priori* error estimates are contributed.

The present paper is devoted to the abstract analysis of linear elasticity problems in which the elastic compliance is allowed tohave a non trivial kernel, so that the elastic strains are subject to a linear constraint. Problems involving such constraints have been recently analysed in [5], [105] and critically reviewed in [156]. Our aim is to provide criteria for the assessment of the well-posedness property for this class of problems. Well-posedness corresponds to the engineering expectation that a (possibly non-unique) solution of a problem must exist under suitable variational conditions of admissibility on the data.

An elastic model capable to encompass all the usual engineering applications must include a possibly singular elastic compliance and external elastic constraints characterized by a non-coercive stiffness operator.

The treatment of such general kind of models is out of the range of applicability of the results that can be found in treatises on the foundation of elasticity (see e.g. [56], [39]). New necessary and sufficient conditions for the existence of a solution and applicable criteria for their fulfilment are thus needed.

Banach's fundamental theorems in Functional Analysis and basic elements of the theory of Hilbert spaces are the essential background for the investigation [192], [17]. A review of the essential notions and propositions can be found in e.g. in [19] and [159].

To provide a self-consistent presentation, I have devoted chapter 7 to a brief exposition of classical results of functional analysis referred to in other chapters.

The proof of some new results concerning closedness properties is contributed in a preliminary section. They are the inequality which characterizes the closedness of the sum of two closed subspaces and with a criterion for the closedness of the image of the product of two operators.

An abstract treatment of linear problems governed by symmetric bilinear forms yields a reference framework for the subsequent analysis. The characteristic properties of structural models are then illustrated and the problems of equilibrium and of kinematic compatibility are discussed.

The mixed formulation of an elastic structural problem with a singular behaviour of the constitutive operator and of the external elastic constraints is then discussed. The analysis is based on the split of the stress field into its elastically effective and ineffective parts. By expressing the effective part in terms of the strain field an equivalent problem in terms of the kinematic field and of the ineffective stress field is obtained. The discussion of this problem is illuminating and reveals which condition must be fulfilled for its equivalence to a reduced problem whose sole unknown is the kinematic field.

This is a classical symmetric one-field problem in which trial and test fields belong to the same space. The necessary and sufficient conditions for well-posedness of the reduced problem are discussed in detail and applicable criteria for their fulfilment are contributed.

The well-posedness of the more challenging situation in which the external elastic energy is not semielliptic is then discussed. This extension is motivated by the analysis of elastic structures resting on elastic beds. The treatment starts with the observation that in the applications the external elastic energy can be assumed to be semielliptic with respect to rigid kinematisms and is based on an original result named the elastic bed inequality.

It is shown that the condition ensuring the equivalence of the mixed problem to a reduced one and the well-posedness criteria of the reduced problem are always met for simple structural models, defined to be those in which the subspaces of rigid displacements and of self-stresses are finite dimensional.

This result provides a theoretical basis to engineers' confidence to get a solution of structural assemblies composed by one dimensional elements such as bars and beams with possibly singular elastic compliances and resting on elastic beds.

The discussion of two or three-dimensional structural models with singular elastic compliance is by far more difficult and the answer to well-posedness is generally negative due to the infinite dimensionality of the subspace of self-stresses. The condition which fails to be met is the one ensuring the equivalence between the mixed problem and the corresponding reduced one. Actually, a singularity of the elastic compliance imposes a constraint on the strain fields. The compatibility requirement induces a corresponding constraint on the kinematic fields and hence reactive forces are originated.

The equivalence above requires the existence of elastically ineffective stresses in equilibrium with the reactive forces. The trouble arises from the fact that only very special singularities of the elastic compliance ensure the existence of such stress fields. This difficulty explains why the discussion of mixed problems is by far more challenging than the discussion of one-field problems.

8.2 Symmetric linear problems

In view of its application to the theory of linear elastic problems we discuss here an abstract symmetric linear problem in a HILBERT space.

Let $\mathbf{a} \in BL(\mathcal{X}^2 ; \Re)$ be a continuous symmetric bilinear form on the product space $\mathcal{X}^2 = \mathcal{X} \times \mathcal{X}$ and $\mathbf{A} \in BL(\mathcal{X}, \mathcal{X}^*)$ the associated symmetric continuous operator, so that

$$\mathbf{a}(\mathbf{x}, \mathbf{y}) = \mathbf{a}(\mathbf{y}, \mathbf{x}) = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle \quad \forall \, \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

Given a closed subspace $\mathcal{L}$ of $\mathcal{X}$ and a functional $\ell \in \mathcal{X}^*$, we consider the linear problem

$$\mathbb{P})\quad \mathbf{a}\,(\mathbf{x},\mathbf{y}) = \ell(\mathbf{y}) \quad \mathbf{x} \in \mathcal{L} \quad \forall\, \mathbf{y} \in \mathcal{L}.$$

The duality between $\mathcal{X}$ and $\mathcal{X}^*$ induces a duality between $\mathcal{L} \subseteq \mathcal{X}$ and the quotient space $\mathcal{X}^*/\mathcal{L}^\perp$ by setting for any $\overline{\mathbf{x}} \in \mathcal{X}^*/\mathcal{L}^\perp$

$$\langle \overline{\mathbf{x}}, \mathbf{y} \rangle := \langle \mathbf{x}, \mathbf{y} \rangle \quad \forall\, \mathbf{y} \in \mathcal{L} \quad \forall\, \mathbf{x} \in \overline{\mathbf{x}}\,.$$

It is then convenient to provide an alternative formulation of the problem in terms of a reduced operator $\mathbf{A}_o \in BL\,(\mathcal{L}, \mathcal{X}^*/\mathcal{L}^\perp)$ and of a reduced functional $\ell_o \in \mathcal{X}^*/\mathcal{L}^\perp$ defined by

$$\mathbf{A}_o\mathbf{x} := \mathbf{A}\mathbf{x} + \mathcal{L}^\perp \quad \forall\, \mathbf{x} \in \mathcal{L}\,; \quad \ell_o := \ell + \mathcal{L}^\perp.$$

We have

$$\mathbf{a}\,(\mathbf{x},\mathbf{y}) = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{A}_o\mathbf{x}, \mathbf{y} \rangle \quad \forall\, \mathbf{x}, \mathbf{y} \in \mathcal{L}.$$

and problem $\mathbb{P}$ can now be rewritten as

$$\mathbb{P})\quad \mathbf{A}_o\mathbf{x} = \ell_o \quad \mathbf{x} \in \mathcal{L}.$$

Definition 8.2.1 (Well-posedness) *The symmetric problem $\mathbb{P}$ is said to be well-posed if it admits a unique solution for any data $\ell_o \in (Ker\,\mathbf{A}_o)^\perp$.*

BANACH's closed range theorem, recalled in chapter 7 as theorem 7.1.3, shows that the well-posedness of problem $\mathbb{P}$ is equivalent to the closedness of $\operatorname{Im} \mathbf{A}_o$ in $\mathcal{X}^*/\mathcal{L}^\perp$. The basic properties of well-posed symmetric linear problems are reported hereafter.

Theorem 8.2.1 (Existence and uniqueness properties) *The solution set of a well-posed symmetric problem $\mathbb{P}$ meets the following alternative:*
i) If $Ker\,\mathbf{A}_o \neq \{\mathbf{o}\}$ the solution set is a non-empty linear manifold parallel to $Ker\,\mathbf{A}_o$ for any admissible data $\ell_o \in (Ker\,\mathbf{A}_o)^\perp$,
ii) If $Ker\,\mathbf{A}_o = \{\mathbf{o}\}$ the solution is unique for any data $\ell_o \in \mathcal{X}^*/\mathcal{L}^\perp$.

We notice that the range and the kernel of the reduced operator are given by

$$\operatorname{Im} \mathbf{A}_o = (\mathbf{A}\mathcal{L} + \mathcal{L}^\perp)/\mathcal{L}^\perp$$

$$Ker\,\mathbf{A}_o = (\mathbf{A}^{-1}\mathcal{L}^\perp) \cap \mathcal{L} = (\mathbf{A}\mathcal{L})^\perp \cap \mathcal{L}$$

The closedness of $\operatorname{Im} \mathbf{A}_o$ can be expressed by stating that the bilinear form $\mathbf{a}$ is closed on $\mathcal{L} \times \mathcal{L}$ and is equivalently expressed by the conditions

$$
\begin{aligned}
&i) \quad \|\mathbf{A}_o \mathbf{x}\|_{\mathcal{X}^*/\mathcal{L}^\perp} \geq c_{\mathbf{a}} \, \|\mathbf{x}\|_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o} \quad \forall \mathbf{x} \in \mathcal{L} \\
&ii) \quad \sup_{\mathbf{y} \in \mathcal{L}} \frac{\mathbf{a}(\mathbf{x}, \mathbf{y})}{\|\mathbf{y}\|_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o}} \geq c_{\mathbf{a}} \, \|\mathbf{x}\|_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o} \quad \forall \mathbf{x} \in \mathcal{L} \\
&iii) \quad \inf_{\mathbf{x} \in \mathcal{L}} \sup_{\mathbf{y} \in \mathcal{L}} \frac{\mathbf{a}(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\|_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o} \|\mathbf{y}\|_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o}} \geq c_{\mathbf{a}} > 0 \\
&iv) \quad \mathbf{A}\mathcal{L} + \mathcal{L}^\perp \quad \text{is closed in} \quad \mathcal{X}^* .
\end{aligned}
$$

Property *iv*) is a direct consequence of lemma 7.1.4.

It is important to provide an expression of the kernel of the reduced operator in terms of the kernel of the symmetric bilinear form $\mathbf{a} \in BL(\mathcal{X}^2 ; \Re)$ defined by

$$
Ker\, \mathbf{a} = Ker\, \mathbf{A} := \{ \mathbf{x} \in \mathcal{X} \mid \mathbf{a}(\mathbf{x}, \mathbf{y}) = 0 \quad \forall \mathbf{y} \in \mathcal{X} \} .
$$

Although in general we have only that

$$
Ker\, \mathbf{A}_o = (\mathbf{A}\mathcal{L})^\perp \cap \mathcal{L} \supseteq Ker\, \mathbf{a} \cap \mathcal{L} ,
$$

the next result provides a sufficient condition to get an equality in the expression above.

Theorem 8.2.2 (A formula for the kernel) *If the symmetric bilinear form* $\mathbf{a}$ *is positive on the whole space* $\mathcal{X}$

$$
\mathbf{a}(\mathbf{x}, \mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \mathcal{X}
$$

then we have that

$$
Ker\, \mathbf{A}_o = (\mathbf{A}\mathcal{L})^\perp \cap \mathcal{L} = Ker\, \mathbf{a} \cap \mathcal{L} .
$$

Proof. We first observe that

$$
\begin{aligned}
\mathbf{x} \in (\mathbf{A}\mathcal{L})^\perp \cap \mathcal{L} &\Longleftrightarrow \mathbf{a}(\mathbf{x}, \mathbf{y}) = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{A}\mathbf{y}, \mathbf{x} \rangle = 0 , \quad \mathbf{x} \in \mathcal{L} \quad \forall \mathbf{y} \in \mathcal{L} \\
&\Longrightarrow \mathbf{a}(\mathbf{x}, \mathbf{x}) = 0 , \quad \mathbf{x} \in \mathcal{L} .
\end{aligned}
$$

By the positivity of $\mathbf{a} \in BL(\mathcal{X}^2\,;\Re)$, the zero value is an absolute minimum of $\mathbf{a}$ in $\mathcal{X}$ so that any directional derivative will vanish at a minimum point. Hence we have

$$\mathbf{a}(\mathbf{x},\mathbf{x}) = 0 \quad \mathbf{x} \in \mathcal{L} \Longrightarrow \mathbf{a}(\mathbf{x},\mathbf{y}) = 0\,, \quad \mathbf{x} \in \mathcal{L} \quad \forall\, \mathbf{y} \in \mathcal{X}$$
$$\Longleftrightarrow \mathbf{x} \in Ker\,\mathbf{a} \cap \mathcal{L}\,,$$

and the proposition is proved. ■

The next result provides a criterion for the closedness of $\mathbf{a}$ on $\mathcal{L} \times \mathcal{L}$.

Theorem 8.2.3 (A sufficient closedness condition) *The inequality:*

$$\mathbf{a}(\mathbf{x},\mathbf{x}) \geq c_{\mathbf{a}}\,\|\mathbf{x}\|^2_{\mathcal{X}/Ker\,\mathbf{A}_o} \quad c_{\mathbf{a}} > 0 \qquad \forall\,\mathbf{x} \in \mathcal{L}$$

implies the closedness of $\mathbf{a}$ *on* $\mathcal{L} \times \mathcal{L}$.

Proof. It suffices to observe that the inequality

$$\inf_{\mathbf{x}\in\mathcal{L}}\ \sup_{\mathbf{y}\in\mathcal{L}} \frac{\mathbf{a}(\mathbf{x},\mathbf{y})}{\|\mathbf{x}\|_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o}\|\mathbf{y}\|_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o}} \geq \inf_{\mathbf{x}\in\mathcal{L}} \frac{\mathbf{a}(\mathbf{x},\mathbf{x})}{\|\mathbf{x}\|^2_{\mathcal{X}/\mathrm{Ker}\,\mathbf{A}_o}} \geq c_{\mathbf{a}} > 0\,,$$

provides the result. ■

By theorems 8.2.2 and 8.2.3 we get the result which will be directly referred to in the discussion of elastic problems.

Theorem 8.2.4 (Semi-ellipticity) *Let the bilinear form* $\mathbf{a} \in BL(\mathcal{X}^2\,;\Re)$ *be symmetric and positive on the whole space* $\mathcal{X}$. *Then the property of semi-ellipticity of* $\mathbf{a}$ *on* $\mathcal{L}$*:*

$$\mathbf{a}(\mathbf{x},\mathbf{x}) \geq c_{\mathbf{a}}\,\|\mathbf{x}\|^2_{\mathcal{X}/(Ker\,\mathbf{a}\cap\mathcal{L})} \quad \forall\,\mathbf{x} \in \mathcal{L}$$

implies the closedness of $\mathbf{a}$ *on* $\mathcal{L} \times \mathcal{L}$.

8.3 Linear structural problems

The formal framework for the analysis of linear structural models is provided by two pairs of dual Hilbert spaces:

- the kinematic space $\mathcal{V}$ and the force space $\mathcal{F}$,

- the strain space $\mathcal{D}$ and the stress space $\mathcal{S}$,

and a pair of dual operators:

- the kinematic operator $\mathbf{B} \in BL(\mathcal{V}, \mathcal{D})$,
- the equilibrium operator $\mathbf{B}' \in BL(\mathcal{S}, \mathcal{F})$.

Remark 8.3.1 *In applications stresses and strains are defined to be square integrable fields. Accordingly we shall identify the stress space $\mathcal{S}$ and the strain space $\mathcal{D}$ with a pivot* Hilbert *space. The inner product in $\mathcal{D} = \mathcal{S}$ will be denoted by $(\!(\cdot\,,\cdot)\!)$ and the duality pairing between $\mathcal{V}$ and $\mathcal{F}$ by $\langle \cdot, \cdot \rangle$.*

The kinematic and the equilibrium operators are the dual counterparts of a fundamental bilinear form $\mathbf{b}$ which describes the geometry of the model:

$$\mathbf{b}(\mathbf{v}, \boldsymbol{\sigma}) := (\!(\,\boldsymbol{\sigma}\,, \mathbf{B}\mathbf{v}\,)\!) = \langle \mathbf{B}'\boldsymbol{\sigma}, \mathbf{v} \rangle \quad \forall \boldsymbol{\sigma} \in \mathcal{S}, \mathbf{v} \in \mathcal{V}.$$

As we shall see, the well-posedness of the structural model requires the closedness of the fundamental form $\mathbf{b}$ on $\mathcal{S} \times \mathcal{V}$ which is expressed by the inf-sup condition [165]

$$\inf_{\boldsymbol{\sigma} \in \mathcal{S}} \sup_{\mathbf{v} \in \mathcal{V}} \frac{\mathbf{b}(\mathbf{v}, \boldsymbol{\sigma})}{\|\boldsymbol{\sigma}\|_{\mathcal{S}/\mathrm{Ker}\,\mathbf{B}'} \|\mathbf{v}\|_{\mathcal{V}/\mathrm{Ker}\,\mathbf{B}}} = \inf_{\mathbf{v} \in \mathcal{V}} \sup_{\boldsymbol{\sigma} \in \mathcal{S}} \frac{\mathbf{b}(\mathbf{v}, \boldsymbol{\sigma})}{\|\boldsymbol{\sigma}\|_{\mathcal{S}/\mathrm{Ker}\,\mathbf{B}'} \|\mathbf{v}\|_{\mathcal{V}/\mathrm{Ker}\,\mathbf{B}}} > 0.$$

This means that the kinematic and the equilibrium operator have closed ranges and can be expressed by stating any one of the equivalent inequalities

$$\|\mathbf{B}\mathbf{v}\|_{\mathcal{D}} \geq c_{\mathbf{b}} \|\mathbf{v}\|_{\mathcal{V}/\mathrm{Ker}\,\mathbf{B}} \quad \forall \mathbf{v} \in \mathcal{V} \iff \|\mathbf{B}'\boldsymbol{\sigma}\|_{\mathcal{F}} \geq c_{\mathbf{b}} \|\boldsymbol{\sigma}\|_{\mathcal{S}/\mathrm{Ker}\,\mathbf{B}'} \quad \forall \boldsymbol{\sigma} \in \mathcal{S}$$

where $c_{\mathbf{b}}$ is a positive constant.

8.3.1 Linear constraints

Rigid bilateral constraints acting on the structure are modeled by considering a closed subspace $\mathcal{L} \subseteq \mathcal{V}$ of conforming kinematisms.

The duality between $\mathcal{V}$ and $\mathcal{F}$ induces a duality pairing between the closed subspace $\mathcal{L}$ and the quotient space $\mathcal{F}/\mathcal{L}^{\perp}$ by setting

$$\langle \bar{\mathbf{f}}, \mathbf{v} \rangle := \langle \mathbf{f}, \mathbf{v} \rangle \quad \forall \mathbf{v} \in \mathcal{L} \quad \forall \mathbf{f} \in \bar{\mathbf{f}} \in \mathcal{F}/\mathcal{L}^{\perp}.$$

It is convenient to introduce the following pair of reduced dual operators:

- the reduced kinematic operator $\mathbf{B}_{\mathcal{L}} \in BL(\mathcal{L}, \mathcal{D})$, defined as the restriction of $\mathbf{B}$ to $\mathcal{L}$,
- the reduced equilibrium operator $\mathbf{B}'_{\mathcal{L}} \in BL(\mathcal{S}, \mathcal{F}/\mathcal{L}^{\perp})$, defined by the position $\mathbf{B}'_{\mathcal{L}}\boldsymbol{\sigma} := \mathbf{B}'\boldsymbol{\sigma} + \mathcal{L}^{\perp}$.

The kernels and the images of the reduced operators are given by

$$Ker\,\mathbf{B}_{\mathcal{L}} = Ker\,\mathbf{B} \cap \mathcal{L}\,; \quad Ker\,\mathbf{B}'_{\mathcal{L}} = (\mathbf{B}')^{-1}\mathcal{L}^{\perp} = (\mathbf{B}\mathcal{L})^{\perp}$$

$$\operatorname{Im}\mathbf{B}_{\mathcal{L}} = \mathbf{B}\mathcal{L}\,; \quad \operatorname{Im}\mathbf{B}'_{\mathcal{L}} = (\operatorname{Im}\mathbf{B}' + \mathcal{L}^{\perp})/\mathcal{L}^{\perp}$$

and we denote by

- $\mathcal{L}_{\mathbf{R}} := Ker\,\mathbf{B} \cap \mathcal{L}$ the subspace of conforming rigid kinematisms and by
- $\mathcal{S}_{\text{SELF}} := (\mathbf{B}\mathcal{L})^{\perp}$ the subspace of self-equilibrated stresses (self-stresses).

A variational theory of structural models with linear external constraints requires that the fundamental form $\mathbf{b}$ is closed on $\mathcal{S} \times \mathcal{L}$. As shown below, this property is in fact necessary and sufficient to express in variational form the problems of equilibrium and of kinematic compatibility.

We recall that by BANACH's closed range theorem 7.1.3, the closedness of $\mathbf{b}$ on $\mathcal{S} \times \mathcal{L}$ can be stated in the equivalent forms:

- orthogonality conditions:

$$\operatorname{Im}\mathbf{B}_{\mathcal{L}} = (Ker\,\mathbf{B}'_{\mathcal{L}})^{\perp}, \qquad \operatorname{Im}\mathbf{B}'_{\mathcal{L}} = (Ker\,\mathbf{B}_{\mathcal{L}})^{\perp},$$

- inequality conditions:

$$\|\mathbf{B}\mathbf{u}\|_{\mathcal{D}} \geq c_{\mathbf{b}}\,\|\mathbf{u}\|_{\mathcal{V}/(\operatorname{Ker}\mathbf{B}\cap\mathcal{L})} \quad \forall\,\mathbf{u} \in \mathcal{L}, \quad c_{\mathbf{b}} > 0,$$

$$\|\mathbf{B}'_{\mathcal{L}}\boldsymbol{\sigma}\|_{\mathcal{F}/\mathcal{L}^{\perp}} \geq c_{\mathbf{b}}\,\|\boldsymbol{\sigma}\|_{\mathcal{S}/(\mathbf{B}\mathcal{L})^{\perp}} \quad \forall\,\boldsymbol{\sigma} \in \mathcal{S}, \quad c_{\mathbf{b}} > 0,$$

- inf-sup conditions:

$$\inf_{\boldsymbol{\sigma}\in\mathcal{S}}\sup_{\mathbf{v}\in\mathcal{L}} \frac{\mathbf{b}(\mathbf{v},\boldsymbol{\sigma})}{\|\boldsymbol{\sigma}\|_{\mathcal{S}/\operatorname{Ker}\mathbf{B}'_{\mathcal{L}}}\|\mathbf{v}\|_{\mathcal{V}/\operatorname{Ker}\mathbf{B}}} = \inf_{\mathbf{v}\in\mathcal{L}}\sup_{\boldsymbol{\sigma}\in\mathcal{S}} \frac{\mathbf{b}(\mathbf{v},\boldsymbol{\sigma})}{\|\boldsymbol{\sigma}\|_{\mathcal{S}/\operatorname{Ker}\mathbf{B}'_{\mathcal{L}}}\|\mathbf{v}\|_{\mathcal{V}/\operatorname{Ker}\mathbf{B}}} > 0.$$

The closedness of $\mathbf{b}$ on $\mathcal{S}\times\mathcal{L}$ can be also expressed by requiring the closedness of the sum of two subspaces, as shown hereafter.

Theorem 8.3.1 (Equivalent closedness properties) *Let $\mathcal{L}$ be a closed subspace of $\mathcal{V}$. Then we have*

$$\mathbf{B}\mathcal{L} \quad \textit{closed in } \mathcal{D} \iff \operatorname{Im}\mathbf{B}' + \mathcal{L}^{\perp} \quad \textit{closed in } \mathcal{F}.$$

If in addition $\operatorname{Im}\mathbf{B}$ *is closed in* $\mathcal{D}$ *the closedness properties above are equivalent to the closedness of* $\mathit{Ker}\,\mathbf{B} + \mathcal{L}$ *in* $\mathcal{V}$.

Proof. The first result follows directly from the expressions of $\operatorname{Im}\mathbf{B}_{\mathcal{L}}$ and $\operatorname{Im}\mathbf{B}'_{\mathcal{L}}$ by recalling theorem 7.1.3 and lemma 7.1.4. The last statement is a simple consequence of theorem 7.1.5. ■

We may then state the main results.

Theorem 8.3.2 (Equilibrium) *Let $\ell \in \mathcal{F}$ be an external force and $\ell_o = \ell + \mathcal{L}^{\perp} \in \mathcal{F}/\mathcal{L}^{\perp}$ the corresponding load on a constrained structural model. The property that* $\mathbf{B}\mathcal{L}$ *is closed in* $\mathcal{D}$ *is necessary and sufficient to ensure that the equilibrium problem*

$$\mathbf{B}'_{\mathcal{L}}\boldsymbol{\sigma} = \ell_o \quad \boldsymbol{\sigma} \in \mathcal{S} \iff \mathbf{B}'\boldsymbol{\sigma} = \ell + \mathbf{r} \quad \boldsymbol{\sigma} \in \mathcal{S}, \mathbf{r} \in \mathcal{L}^{\perp}$$

admits a solution for every load satisfying the consistency condition

$$\ell_o \in (\mathit{Ker}\,\mathbf{B}_{\mathcal{L}})^{\perp} \iff \ell \in (\mathit{Ker}\,\mathbf{B} \cap \mathcal{L})^{\perp}$$

or in variational form $\quad \langle \ell, \mathbf{v} \rangle = 0 \quad \forall\, \mathbf{v} \in \mathcal{L}_{\mathbf{R}} = \mathit{Ker}\,\mathbf{B} \cap \mathcal{L}$.

The degeneracy condition $\mathcal{S}_{\text{SELF}} = \{\mathbf{o}\}$ *is necessary and sufficient for the solution to be unique.*

Theorem 8.3.3 (Compatibility) *A kinematic pair* $\{\boldsymbol{\varepsilon}\,, \mathbf{w}\}$ *with* $\boldsymbol{\varepsilon} \in \mathcal{D}$ *and* $\mathbf{w} \in \mathcal{V}$ *is said to be compatible with the constraints if there exists a conforming kinematic field* $\mathbf{v} \in \mathcal{L}$ *such that*

$$\mathbf{B}\mathbf{v} = \boldsymbol{\varepsilon} - \mathbf{B}\mathbf{w}.$$

The property that $\mathbf{B}\mathcal{L}$ *is closed in* $\mathcal{D}$ *is necessary and sufficient to ensure that the compatibility problem admits solution for every kinematic pair satisfying the consistency condition*

$$\boldsymbol{\varepsilon} - \mathbf{B}\mathbf{w} \in (\mathit{Ker}\,\mathbf{B}'_{\mathcal{L}})^{\perp} = (\mathcal{S}_{\text{SELF}})^{\perp}$$

or in variational form

$$(\!(\, \boldsymbol{\sigma}\,, \boldsymbol{\varepsilon} \,)\!) = (\!(\, \boldsymbol{\sigma}\,, \mathbf{B}\mathbf{w} \,)\!) \quad \forall\, \boldsymbol{\sigma} \in \mathcal{S}_{\text{SELF}}.$$

The degeneracy of the subspace $\mathcal{L}_{\mathbf{R}}$ *of rigid conforming kinematisms is necessary and sufficient in order that the solution be unique.*

8.3.2 Elastic structures

A linearly elastic structure is characterized by a symmetric elastic operator $\mathbf{E} \in BL(\mathcal{D}, \mathcal{S})$ which is $\mathcal{D}$-elliptic:

$$(\!(\, \mathbf{E}\boldsymbol{\varepsilon} \,,\, \boldsymbol{\varepsilon} \,)\!) \geq c_{\mathbf{e}} \, \|\boldsymbol{\varepsilon}\|^2_{\mathcal{D}} \quad c_{\mathbf{e}} > 0 \quad \forall\, \boldsymbol{\varepsilon} \in \mathcal{D}.$$

The elastic strain energy in terms of kinematisms is provided by one-half the quadratic form associated with the positive symmetric bilinear form:

$$\mathbf{a}\,(\mathbf{u}, \mathbf{v}) := (\!(\, \mathbf{EBu} \,,\, \mathbf{Bv} \,)\!) \quad \forall\, \mathbf{u}, \mathbf{v} \in \mathcal{V}.$$

which is called the bilinear form of elastic strain energy.

The elastostatic problem for a constrained structural model consists in evaluating a conforming kinematism $\mathbf{u} \in \mathcal{L}$ such that the corresponding stress field $\boldsymbol{\sigma} = \mathbf{EBu}$ is in equilibrium with the prescribed load $\ell_o = \ell + \mathcal{L}^{\perp} \in \mathcal{F}/\mathcal{L}^{\perp}$.

In terms of elastic strain energy the problem is written as

$$\mathbf{a}\,(\mathbf{u}, \mathbf{v}) = \ell(\mathbf{v}) \qquad \mathbf{u} \in \mathcal{L} \quad \forall\, \mathbf{v} \in \mathcal{L}$$

and is well-posed if and only if $\mathbf{a}$ is closed on $\mathcal{L} \times \mathcal{L}$.

The elastic stiffness of the structure $\mathbf{A} = \mathbf{B}'\mathbf{EB} \in BL(\mathcal{V}, \mathcal{F})$ is the symmetric bounded linear operator associated with $\mathbf{a}$ according to the formula

$$\langle\, \mathbf{Au}, \mathbf{v} \,\rangle = \mathbf{a}\,(\mathbf{u}, \mathbf{v}) \quad \forall\, \mathbf{v} \in \mathcal{V}.$$

A direct verification of the closure property of $\mathbf{a}$ on $\mathcal{L} \times \mathcal{L}$ is often not possible in applications and hence it is natural to look for simpler sufficient conditions.

A key result is provided by the following

Proposition 8.3.1 (Closedness of the elastic operator) *The closedness of* $\mathbf{B}\mathcal{L}$ *and the* $\mathcal{D}$*-ellipticity of the elastic operator* $\mathbf{E}$ *imply the closedness of the bilinear form* $\mathbf{a}$ *on* $\mathcal{L} \times \mathcal{L}$*.*

Proof. From the inequalities

$$\langle\, \mathbf{E}\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \,\rangle \geq c_{\mathbf{e}} \, \|\boldsymbol{\varepsilon}\|^2_{\mathcal{D}} \quad \forall\, \boldsymbol{\varepsilon} \in \mathcal{D}$$

$$\|\mathbf{Bu}\|_{\mathcal{D}} \geq c_{\mathbf{b}} \, \|\mathbf{u}\|_{\mathcal{V}/(\mathrm{Ker}\,\mathbf{B} \cap \mathcal{L})} \quad \forall\, \mathbf{u} \in \mathcal{L}\,.$$

it follows that

$$\mathbf{a}\,(\mathbf{u}, \mathbf{u}) \geq c_{\mathbf{a}} \, \|\mathbf{u}\|^2_{\mathcal{V}/(\mathrm{Ker}\,\mathbf{B} \cap \mathcal{L})} \quad \forall\, \mathbf{u} \in \mathcal{L}$$

where $c_{\mathbf{a}} = c_{\mathbf{e}}\, c^2_{\mathbf{b}}$.

The strict positivity of $\mathbf{E}$ ensures that $Ker\,\mathbf{a} = Ker\,\mathbf{B}$ so that the inequality above can be written as

$$\mathbf{a}\,(\mathbf{u},\mathbf{u}) \geq c_{\mathbf{a}}\,\|\mathbf{u}\|^2_{\mathcal{V}/(\mathrm{Ker}\,\mathbf{a}\cap\mathcal{L})} \quad \forall\,\mathbf{u}\in\mathcal{L}\,,$$

which by proposition 8.2 implies the closedness of $\mathbf{a}$ on $\mathcal{L}\times\mathcal{L}$. ∎

In applications the $\mathcal{D}$-ellipticity of the elastic operator $\mathbf{E}$ is easily checked so that the real task is to verify the closedness of $\mathbf{B}\mathcal{L}$.

Proposition 8.3.2 (A closedness criterion) *Let* $\mathrm{Im}\,\mathbf{B}$ *be closed in* $\mathcal{D}$*. Then the subspace* $\mathbf{B}\mathcal{L}$ *is closed in* $\mathcal{D}$ *if the subspace* $Ker\,\mathbf{B}$ *can be written as the sum of a finite dimensional subspace and of a subspace included in* $\mathcal{L}$

$$Ker\,\mathbf{B} = \mathcal{N} + \mathcal{L}_o\,, \quad \dim\mathcal{N} < +\infty\,, \quad \mathcal{L}_o \subseteq \mathcal{L}\,.$$

Proof. By theorem 8.3.1 we have to verify the closedness of the subspace $Ker\,\mathbf{B} + \mathcal{L}$ in $\mathcal{V}$. The assumption ensures that $Ker\,\mathbf{B} + \mathcal{L} = \mathcal{N} + \mathcal{L}$ with $\dim\mathcal{N} < +\infty$ and hence setting $\mathcal{A} = \mathcal{L}$ and $\mathcal{B} = \mathcal{N}$ in Lemma 7.1.3 we get the result. ∎

Remark 8.3.2 *In most engineering applications the kernel of the kinematic operator* $\mathbf{B}$ *is finite dimensional so that the condition in Proposition 8.3.2 is trivially fulfilled. A relevant exception is provided by the models of cable or membrane structures in which the subspace* $Ker\,\mathbf{B}$ *of rigid kinematic fields is not finite dimensional. The condition in Proposition 8.3.2 is however still met.*

8.4 Mixed formulations

A more challenging problem concerns the elastic equilibrium of a structural model with a partially rigid constitutive behaviour and subject to external elastic constraints.

Rigid bilateral constraints, which have already been analysed, will not be explicitly considered to simplify the presentation. Anyway, they can be taken into account by substituting the kinematic operator $\mathbf{B} \in BL\,(\mathcal{V},\mathcal{D})$ with the reduced operator $\mathbf{B}_{\mathcal{L}} \in BL\,(\mathcal{L},\mathcal{D})$.

The analytical properties of the general model of elastic structure under investigation are described hereafter.

- The internal elastic compliance of the structure is a *continuous*, *symmetric positive* and *closed* bilinear form $\mathbf{c} \in BL(\mathcal{S} \times \mathcal{S}\,;\Re)$:

$$
\begin{array}{lll}
i) & \|\mathbf{c}\|_{\mathcal{S}}\|\boldsymbol{\sigma}\|_{\mathcal{S}}\|\boldsymbol{\tau}\|_{\mathcal{S}} \geq |\mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\tau})| & \forall\, \boldsymbol{\sigma},\boldsymbol{\tau} \in \mathcal{S} \\
ii) & \mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \mathbf{c}(\boldsymbol{\tau},\boldsymbol{\sigma}) & \forall\, \boldsymbol{\sigma},\boldsymbol{\tau} \in \mathcal{S} \\
iii) & \mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\sigma}) \geq 0 & \forall\, \boldsymbol{\sigma} \in \mathcal{S} \\
iv) & \inf\limits_{\boldsymbol{\tau}\in\mathcal{S}} \sup\limits_{\boldsymbol{\sigma}\in\mathcal{S}} \dfrac{\mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\tau})}{\|\boldsymbol{\sigma}\|_{\mathcal{S}/\mathrm{Ker}\,\mathbf{C}}\|\boldsymbol{\tau}\|_{\mathcal{S}/\mathrm{Ker}\,\mathbf{C}}} > 0 &
\end{array}
$$

  The elastic compliance operator $\mathbf{C} \in BL(\mathcal{S},\mathcal{D})$ is defined by

$$
(\!(\, \mathbf{C}\boldsymbol{\sigma} \,,\, \boldsymbol{\tau} \,)\!) := \mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\tau}) \quad \forall\, \boldsymbol{\sigma},\boldsymbol{\tau} \in \mathcal{S},
$$

  and $\mathrm{Im}\,\mathbf{C}$ is closed in $\mathcal{D}$ by virtue of *iv*). The elements of the kernel of $\mathbf{C}$ are the elastically ineffective stress fields.

- The *external elastic stiffness* of the structure is expressed by a *continuous symmetric* and *positive* bilinear form $\mathbf{k} \in BL(\mathcal{V} \times \mathcal{V}\,;\Re)$:

$$
\begin{array}{lll}
i) & \|\mathbf{k}\|_{\mathcal{V}}\|\mathbf{u}\|_{\mathcal{V}}\|\mathbf{v}\|_{\mathcal{V}} \geq |\mathbf{k}(\mathbf{u},\mathbf{v})| & \forall\, \mathbf{u},\mathbf{v} \in \mathcal{V}\,, \\
ii) & \mathbf{k}(\mathbf{u},\mathbf{v}) = \mathbf{k}(\mathbf{v},\mathbf{u}) & \forall\, \mathbf{u},\mathbf{v} \in \mathcal{V}\,, \\
iii) & \mathbf{k}(\mathbf{u},\mathbf{u}) \geq 0 & \forall\, \mathbf{u} \in \mathcal{V}\,.
\end{array}
$$

The external elastic stiffness operator $\mathbf{K} \in BL(\mathcal{V},\mathcal{F})$ is defined by

$$
\langle \mathbf{K}\mathbf{u}, \mathbf{v} \rangle := \mathbf{k}(\mathbf{u},\mathbf{v}) \quad \forall\, \mathbf{u},\mathbf{v} \in \mathcal{V}.
$$

The elements of the kernel of $\mathbf{K}$ are kinematic fields which do not involve reactions of the external elastic constraints.

We emphasize that the bilinear form $\mathbf{k}$ is not assumed to be closed on $\mathcal{V} \times \mathcal{V}$. As we shall see this is important in applications and makes the static and the kinematic equations of the mixed formulation play different roles.

The mixed elastostatic problem is formulated in operator form as

$$
\mathbb{M})\quad \begin{cases} \mathbf{K}\mathbf{u} + \mathbf{B}'\boldsymbol{\sigma} = \mathbf{f} \\ \mathbf{B}\mathbf{u} - \mathbf{C}\boldsymbol{\sigma} = \boldsymbol{\delta} \end{cases} \quad \text{or} \quad \mathbf{S}\begin{vmatrix} \mathbf{u} \\ \boldsymbol{\sigma} \end{vmatrix} = \begin{vmatrix} \mathbf{K} & \mathbf{B}' \\ \mathbf{B} & -\mathbf{C} \end{vmatrix}\begin{vmatrix} \mathbf{u} \\ \boldsymbol{\sigma} \end{vmatrix} = \begin{vmatrix} \mathbf{f} \\ \boldsymbol{\delta} \end{vmatrix}
$$

where $\mathbf{S} \in BL(\mathcal{V} \times \mathcal{S}, \mathcal{F} \times \mathcal{D})$ is called the structural operator.

Equation $\mathbb{M}_1$ expresses the equilibrium condition in which

- $\mathbf{f} \in \mathcal{F}$ is the assigned load,
- $-\mathbf{K}\mathbf{u} \in \mathcal{F}$ is the reaction of the external elastic constraints,
- $\mathbf{B}'\boldsymbol{\sigma} \in \mathcal{F}$ is the total external force.

Equation $\mathbb{M}_2$ expresses the kinematic compatibility condition in which

- $\boldsymbol{\delta} \in \mathcal{D}$ is an imposed distorsion,
- $\mathbf{C}\boldsymbol{\sigma} \in \mathcal{D}$ is the elastic strain,
- $\mathbf{B}\mathbf{u} \in \mathcal{D}$ is the total strain field.

Imposed distorsions are often considered in engineering applications e.g. to simulate the effect of temperature fields in the structures.

The variational form of the mixed elastostatic problem is given by

$$\mathbb{M}) \quad \begin{cases} \mathbf{k}(\mathbf{u},\mathbf{v}) + \mathbf{b}(\mathbf{v},\boldsymbol{\sigma}) = \langle \mathbf{f},\mathbf{v} \rangle & \mathbf{u} \in \mathcal{V}, \quad \forall\, \mathbf{v} \in \mathcal{V}, \\ \mathbf{b}(\mathbf{u},\boldsymbol{\tau}) - \mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\tau}) = \langle \boldsymbol{\delta},\boldsymbol{\tau} \rangle & \boldsymbol{\sigma} \in \mathcal{S}, \quad \forall\, \boldsymbol{\tau} \in \mathcal{S}. \end{cases}$$

Problems of this kind have been longly analysed in the literature (see e.g. the references in [10], [136], [148]) following the pioneering works by I. Babuška [11] and F. Brezzi [18]. A comprehensive presentation of the state of the art can be found in the book [20] by F. Brezzi and M. Fortin on Mixed and Hybrid FEM formulations.

The approach proposed here is directly related with the original existence and uniqueness theorem by Brezzi [18].

His analysis was concerned with a mixed problem $\mathbb{M}$ in which the form $\mathbf{c}$ was taken to be zero and neither the simmetry nor the positivity of the form $\mathbf{k}$ were assumed.

A more general case in which a positive and symmetric form $\mathbf{c}$ is included has been recently addressed in [20], theorem II.1.2, by adopting a perturbation technique. A sufficient condition for the existence of a solution of the mixed problem is provided in [20] under a special assumption concerning the bilinear form $\mathbf{c}$ of elastic compliance.

However many engineering models of elastic structures fall outside the range of the existing results.

The analysis which we develop here is intended to provide a well-posedness result capable to encompass the usual engineering models in elasticity.

We preliminarily quote a result concerning the kernel of the structural operator $\mathbf{S} \in BL(\mathcal{V} \times \mathcal{S}, \mathcal{F} \times \mathcal{D})$.

Proposition 8.4.1 (Representation of the kernel) *Let the bilinear forms* $\mathbf{c} \in BL(\mathcal{S} \times \mathcal{S}\,;\Re)$ *and* $\mathbf{k} \in BL(\mathcal{V} \times \mathcal{V}\,;\Re)$ *be symmetric and positive. Then the kernel of the structural operator* $\mathbf{S} \in BL(\mathcal{V} \times \mathcal{S}, \mathcal{F} \times \mathcal{D})$ *is given by*

$$Ker\,\mathbf{S} = \begin{vmatrix} Ker\,\mathbf{B} \cap Ker\,\mathbf{K} \\ Ker\,\mathbf{B}' \cap Ker\,\mathbf{C} \end{vmatrix}.$$

Proof. A pair $\{\mathbf{u}\,,\boldsymbol{\sigma}\}$ belongs to $Ker\,\mathbf{S}$ if and only if

$$\begin{cases} \mathbf{k}(\mathbf{u},\mathbf{v}) + \mathbf{b}(\mathbf{v},\boldsymbol{\sigma}) = 0 & \forall\, \mathbf{v} \in \mathcal{V}, \\ \mathbf{b}(\mathbf{u},\boldsymbol{\tau}) - \mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\tau}) = 0 & \forall\, \boldsymbol{\tau} \in \mathcal{S}, \end{cases} \iff \begin{cases} \mathbf{K}\mathbf{u} + \mathbf{B}'\boldsymbol{\sigma} = 0 \\ \mathbf{B}\mathbf{u} - \mathbf{C}\,\boldsymbol{\sigma} = 0, \end{cases}$$

which imply that

$$\begin{cases} \mathbf{k}(\mathbf{u},\mathbf{u}) + \mathbf{b}(\mathbf{u},\boldsymbol{\sigma}) = 0\,, \\ \mathbf{b}(\mathbf{u},\boldsymbol{\sigma}) - \mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\sigma}) = 0\,. \end{cases}$$

Subtracting we get $\mathbf{k}(\mathbf{u},\mathbf{u}) + \mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\sigma}) = 0$ and the positivity of $\mathbf{k}$ and $\mathbf{c}$ implies that $\mathbf{k}(\mathbf{u},\mathbf{u}) = 0$ and $\mathbf{c}(\boldsymbol{\sigma},\boldsymbol{\sigma}) = 0$. Hence, being $\mathbf{u} \in \mathcal{V}$ and $\boldsymbol{\sigma} \in \mathcal{S}$ absolute minimum points of $\mathbf{k}$ and $\mathbf{c}$, their derivatives must vanish there. By the symmetry of $\mathbf{k}$ and $\mathbf{c}$ these conditions are expressed by $\mathbf{K}\mathbf{u} = \mathbf{o}$ and $\mathbf{C}\boldsymbol{\sigma} = \mathbf{o}$. Substituting in the expression of the kernel we infer that $\mathbf{B}\mathbf{u} = \mathbf{o}$ and $\mathbf{B}'\boldsymbol{\sigma} = \mathbf{o}$. ■

If a solution $\{\mathbf{u}\,,\boldsymbol{\sigma}\} \in \mathcal{V} \times \mathcal{S}$ to problem $\mathbb{M}$ exists, the data $\{\mathbf{f}\,,\boldsymbol{\delta}\} \in \mathcal{F} \times \mathcal{D}$ must necessarily meet the following variational conditions of admissibility

$$\mathbf{f} \in (Ker\,\mathbf{B} \cap Ker\,\mathbf{K})^{\perp}, \qquad \boldsymbol{\delta} \in (Ker\,\mathbf{B}' \cap Ker\,\mathbf{C})^{\perp}$$

which express the orthogonality of $\{\mathbf{f}\,,\boldsymbol{\delta}\}$ to the kernel of the structural operator.

The engineers' confidence in finding solutions to elasticity problems is based upon the implicit assumption of well-posedness of the problem, a condition explicitly stated hereafter by recalling Definition 8.2.1.

Definition 8.4.1 (Well-posedness of the mixed problem) *The mixed problem* $\mathbb{M}$ *is well-posed if the structural operator* $\mathbf{S} \in BL(\mathcal{V}\times\mathcal{S}, \mathcal{F}\times\mathcal{D})$ *has a closed range. The variational conditions of admissibility on the data* $\{\mathbf{f}\,,\boldsymbol{\delta}\} \in (Ker\,\mathbf{S})^{\perp}$ *are then also sufficient to ensure the existence of a solution, unique to within fields of the kernel* $Ker\,\mathbf{S}$ *of the structural operator.*

Remark 8.4.1 *The well-posedness of the mixed problem* $\mathbb{M}$ *requires the validity of the orthogonality relations*

$$\operatorname{Im}\mathbf{B}' + \operatorname{Im}\mathbf{K} = (\mathit{Ker}\,\mathbf{B} \cap \mathit{Ker}\,\mathbf{K})^{\perp}\,, \quad \operatorname{Im}\mathbf{B} + \operatorname{Im}\mathbf{C} = (\mathit{Ker}\,\mathbf{B}' \cap \mathit{Ker}\,\mathbf{C})^{\perp}\,.$$

By remark 7.1.2 the equalities above hold if and only if the sum of the two subspaces on the left hand sides is closed.

8.4.1 Solution strategy

Our aim is to provide a necessary and sufficient condition for the well-posedness of the mixed problem $\mathbb{M}$.

Planning the attack, we first try to transform the mixed problem $\mathbb{M}$ into a problem involving only kinematic fields.

To this end we must modify condition $\mathbb{M}_2$ of kinematic compatibility by inverting the elastic law to get an expression of the stress field $\boldsymbol{\sigma} \in \mathcal{S}$ in terms of the strain associated with the kinematic field $\mathbf{u} \in \mathcal{V}$. Since the internal elastic compliance operator $\mathbf{C} \in BL(\mathcal{S}, \mathcal{D})$ is singular, we have to pick up its non-singular part.

Due to the symmetry of $\mathbf{C}$ and the closedness of $\operatorname{Im}\mathbf{C}$, the subspace $\mathit{Ker}\,\mathbf{C}$ of elastically ineffective stresses and the subspace $\operatorname{Im}\mathbf{C}$ of elastic strains fulfil the orthogonality conditions

$$\mathit{Ker}\,\mathbf{C} = (\operatorname{Im}\mathbf{C})^{\perp} \quad \text{and} \quad \operatorname{Im}\mathbf{C} = (\mathit{Ker}\,\mathbf{C})^{\perp}\,.$$

Recalling remark 8.3.1 the spaces $\mathcal{D}$ and $\mathcal{S}$ can be identified without loss in generality. We can then perform the direct sum decomposition of the stress-strain space into complementary orthogonal subspaces

$$\mathcal{D} = \mathcal{S} = \operatorname{Im}\mathbf{C} \oplus \mathit{Ker}\,\mathbf{C}.$$

The reduced compliance operator $\mathbf{C}_o \in BL(\operatorname{Im}\mathbf{C}, \operatorname{Im}\mathbf{C})$, defined by

$$\mathbf{C}_o\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\sigma} \quad \forall\, \boldsymbol{\sigma} \in \operatorname{Im}\mathbf{C} \subseteq \mathcal{S},$$

is positive definite and the operator $\mathbf{C}$ can be partitioned as follows:

$$\begin{vmatrix} \mathbf{C}_o & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{vmatrix} \begin{vmatrix} \boldsymbol{\sigma}^* \\ \boldsymbol{\sigma}_o \end{vmatrix} \quad \textit{with} \quad \begin{cases} \boldsymbol{\sigma}^* \in \operatorname{Im}\mathbf{C} \\ \boldsymbol{\sigma}_o \in \mathit{Ker}\,\mathbf{C}\,. \end{cases}$$

We also define in $\mathcal{S} = \mathcal{D}$ the symmetric orthogonal projector $\mathbf{P} = \mathbf{P}'$ onto the subspace $Ker\,\mathbf{C}$ of elastically ineffective stresses so that

$$\operatorname{Im}\mathbf{P} = Ker\,\mathbf{C}, \quad Ker\,\mathbf{P} = \operatorname{Im}\mathbf{C}.$$

The kernel of the product operator $\mathbf{PB} \in BL(\mathcal{V}, Ker\,\mathbf{C})$ is defined by

$$Ker\,\mathbf{PB} = \{\,\mathbf{u} \in \mathcal{V} \mid \mathbf{Bu} \in \operatorname{Im}\mathbf{C}\,\}$$

and its elements are the kinematic fields which generate elastic strain fields.

Remark 8.4.2 *According to Remark 8.4.1, the closedness of* $\operatorname{Im}\mathbf{B} + \operatorname{Im}\mathbf{C} = \operatorname{Im}\mathbf{B} + Ker\,\mathbf{P}$ *is a necessary condition for the well-posedness of the mixed problem. Further, by Lemma 7.1.5, this assumption is also equivalent to the closedness of* $\operatorname{Im}\mathbf{B}'\mathbf{P}'$ *in* $\mathcal{F}$ *and hence, by the closed range theorem 7.1.3, to the closedness of* $\operatorname{Im}\mathbf{PB}$*.*

Let us then assume that $\operatorname{Im}\mathbf{PB}$ is closed in $\mathcal{D}$ so that for any $\boldsymbol{\delta} \in (Ker\,\mathbf{B}' \cap Ker\,\mathbf{C})^{\perp}$ we can perform the decomposition

$$\boldsymbol{\delta} = \boldsymbol{\delta}_o + \boldsymbol{\delta}^* \quad \text{with} \quad \boldsymbol{\delta}_o \in \operatorname{Im}\mathbf{B} \quad \text{and} \quad \boldsymbol{\delta}^* \in \operatorname{Im}\mathbf{C}.$$

Choosing $\mathbf{u}_o \in \mathcal{V}$ such that $\mathbf{Bu}_o = \boldsymbol{\delta}_o$ the compatibility equation $\mathbb{M}_2$ can be rewritten as

$$\mathbf{Bu}^* = \mathbf{C}_o\boldsymbol{\sigma}^* + \boldsymbol{\delta}^*.$$

Denoting by $\mathbf{E}$ the inverse of $\mathbf{C}_o$ we can also write

$$\boldsymbol{\sigma}^* = \mathbf{E}(\mathbf{Bu}^* - \boldsymbol{\delta}^*).$$

Substituting into the equilibrium equation $\mathbb{M}_1$ we get the following problem in the unknown fields $\mathbf{u}^* \in Ker\,\mathbf{PB}$ and $\boldsymbol{\sigma}_o \in Ker\,\mathbf{C}$

$$\mathbb{P}) \quad (\mathbf{K} + \mathbf{B}'\mathbf{EB})\mathbf{u}^* + \mathbf{B}'\boldsymbol{\sigma}_o = \mathbf{f} - \mathbf{Ku}_o + \mathbf{B}'\mathbf{E}\boldsymbol{\delta}^*.$$

Let us now define the bilinear form of the elastic energy

$$\mathbf{a}(\mathbf{u}^*, \mathbf{v}) := \mathbf{k}(\mathbf{u}^*, \mathbf{v}) + ((\,\mathbf{EBu}^*\,, \mathbf{Bv}\,)) \quad \forall\,\mathbf{u}^* \in Ker\,\mathbf{PB} \quad \forall\,\mathbf{v} \in \mathcal{V}$$

and the effective load

$$\langle \ell, \mathbf{v}\rangle := \langle \mathbf{f}, \mathbf{v}\rangle - \mathbf{k}(\mathbf{u}_o, \mathbf{v}) + ((\,\mathbf{E}\boldsymbol{\delta}^*\,, \mathbf{Bv}\,)) \quad \forall\,\mathbf{v} \in \mathcal{V}.$$

The stiffness operator $\mathbf{A} = \mathbf{K} + \mathbf{B}'\mathbf{EB}$ is defined by the identity

$$\langle \mathbf{Au}^*, \mathbf{v}\rangle = \mathbf{a}(\mathbf{u}^*, \mathbf{v}) \quad \forall\,\mathbf{u}^* \in Ker\,\mathbf{PB} \quad \forall\,\mathbf{v} \in \mathcal{V}.$$

The discussion above is summarized in the next statement.

Proposition 8.4.2 (First equivalence property) *The closedness of* $\mathrm{Im}\,\mathbf{PB}$ *ensures that for any given* $\boldsymbol{\delta} \in (Ker\,\mathbf{B}' \cap Ker\,\mathbf{C})^{\perp}$ *the mixed problem*

$$\mathbb{M})\quad \begin{cases} \mathbf{k}\,(\mathbf{u},\mathbf{v}) + \mathbf{b}\,(\mathbf{v},\boldsymbol{\sigma}) = \langle\, \mathbf{f},\mathbf{v}\,\rangle & \mathbf{u} \in \mathcal{V}, \quad \forall\, \mathbf{v} \in \mathcal{V}, \\ \mathbf{b}\,(\mathbf{u},\boldsymbol{\tau}) - \mathbf{c}\,(\boldsymbol{\sigma},\boldsymbol{\tau}) = \langle\, \boldsymbol{\delta},\boldsymbol{\tau}\,\rangle & \boldsymbol{\sigma} \in \mathcal{S}, \quad \forall\, \boldsymbol{\tau} \in \mathcal{S} \end{cases}$$

in the unknown fields $\mathbf{u} \in \mathcal{V}$ *and* $\boldsymbol{\sigma} \in \mathcal{S}$ *is equivalent to the variational problem*

$$\mathbb{P})\quad \mathbf{a}\,(\mathbf{u}^*,\mathbf{v}) + (\!(\, \boldsymbol{\sigma}_o\,,\, \mathbf{Bv}\,)\!) = \langle\, \ell,\mathbf{v}\,\rangle \quad \forall\, \mathbf{v} \in \mathcal{V}$$

in the unknown fields $\mathbf{u}^* \in Ker\,\mathbf{PB}$ *and* $\boldsymbol{\sigma}_o \in Ker\,\mathbf{C}$ *provided that the pair* $\{\mathbf{u}_o\,,\boldsymbol{\delta}\}^* \in \mathcal{V} \times \mathrm{Im}\,\mathbf{C}$ *is such that* $\boldsymbol{\delta} = \mathbf{B}\mathbf{u}_o + \boldsymbol{\delta}^*$.

The discussion of problem $\mathbb{P}$ is based on its equivalence to a classical one-field problem which is formulated by restricting the test fields $\mathbf{v} \in \mathcal{V}$ to range in the subspace $Ker\,\mathbf{PB} \subseteq \mathcal{V}$.

Proposition 8.4.3 (Second equivalence property) *The closedness of* $\mathrm{Im}\,\mathbf{PB}$ *ensures that the variational problem*

$$\mathbb{P})\quad \mathbf{a}\,(\mathbf{u}^*,\mathbf{v}) + (\!(\, \boldsymbol{\sigma}_o\,,\, \mathbf{Bv}\,)\!) = \langle\, \ell,\mathbf{v}\,\rangle \quad \forall\, \mathbf{v} \in \mathcal{V}$$

in the unknown fields $\mathbf{u}^* \in Ker\,\mathbf{PB}$ *and* $\boldsymbol{\sigma}_o \in Ker\,\mathbf{C}$ *is equivalent to the reduced problem*

$$\mathbb{P}^*)\quad \mathbf{a}\,(\mathbf{u}^*,\mathbf{v}^*) = \langle\, \ell,\mathbf{v}\,\rangle^* \quad \forall\, \mathbf{v}^* \in Ker\,\mathbf{PB}$$

in the unknown field $\mathbf{u}^* \in Ker\,\mathbf{PB}$.

Proof. Clearly if $\{\mathbf{u}^*,\boldsymbol{\sigma}_o\} \in Ker\,\mathbf{PB} \times Ker\,\mathbf{C}$ is a solution of problem $\mathbb{P}$ then $\mathbf{u}^*$ will be solution of problem $\mathbb{P}^*$. In fact we have that $(\!(\, \boldsymbol{\sigma}_o\,,\, \mathbf{Bv}^*\,)\!) = 0$ for all $\mathbf{v}^* \in Ker\,\mathbf{PB}$ since $\boldsymbol{\sigma}_o \in Ker\,\mathbf{C}$ and $\mathbf{Bv}^* \in Ker\,\mathbf{P} = \mathrm{Im}\,\mathbf{C} = (Ker\,\mathbf{C})^{\perp}$.

Conversely if $\mathbf{u}^* \in Ker\,\mathbf{PB}$ is solution of problem $\mathbb{P}^*$ the reactive force $\mathbf{r} \in \mathcal{F}$ defined by

$$\langle\, \mathbf{r},\mathbf{v}\,\rangle := \mathbf{a}\,(\mathbf{u}^*,\mathbf{v}) - \langle\, \ell,\mathbf{v}\,\rangle \quad \forall\, \mathbf{v} \in \mathcal{V}$$

will belong to $(Ker\,\mathbf{PB})^{\perp}$. The assumption $\mathrm{Im}\,\mathbf{PB}$ closed ensures that $\mathrm{Im}\,\mathbf{B}'\mathbf{P}' = (Ker\,\mathbf{PB})^{\perp}$ and hence for any $\mathbf{r} \in (Ker\,\mathbf{PB})^{\perp}$ we can find a $\boldsymbol{\sigma}_o \in \mathrm{Im}\,\mathbf{P}' = Ker\,\mathbf{C}$ such that $\mathbf{B}'\boldsymbol{\sigma}_o = \mathbf{r}$. Then $\langle\, \mathbf{r},\mathbf{v}\,\rangle = (\!(\, \boldsymbol{\sigma}_o\,,\, \mathbf{Bv}\,)\!)$ for all $\mathbf{v} \in \mathcal{V}$ and the pair $\{\mathbf{u}^*,\boldsymbol{\sigma}_o\}$ is solution of problem $\mathbb{P}$. The field $\boldsymbol{\sigma}_o$ is unique to within elements of the subspace $Ker\,\mathbf{B}' \cap Ker\,\mathbf{C}$ of elastically ineffective self-stresses. ■

Remark 8.4.3 *It is worth noting that the expression of the effective load* ℓ *depends upon* $\boldsymbol{\delta}^*$ *and the field* $\mathbf{u}_o$ *which in turn is determined by* $\boldsymbol{\delta}_o$ *only to within an additional rigid field.*

Further the additive decomposition of admissible distorsions $\boldsymbol{\delta}$ *into the sum* $\boldsymbol{\delta}_o + \boldsymbol{\delta}^*$ *is unique only to within elements of* $\operatorname{Im}\mathbf{B} \cap \operatorname{Im}\mathbf{C}$.

Anyway it can be easily shown that the solution $\mathbf{u} = \mathbf{u}_o + \mathbf{u}^*$ *and* $\boldsymbol{\sigma} = \boldsymbol{\sigma}_o + \boldsymbol{\sigma}^*$ *of the mixed problem* $\mathbb{M}$ *remains unaffected by this indeterminacy of* ℓ.

Let us now discuss the well-posedness of the reduced problem $\mathbb{P}^*$.

8.4.2 The reduced structural model

Problem $\mathbb{P}^*$ is the variational formulation of the elastostatic problem for a structural model subject to the rigid bilateral constraints defined by the subspace $Ker\,\mathbf{PB} \subseteq \mathcal{V}$ of conforming kinematic fields. It is formally equivalent to the symmetric linear problems discussed in section 8.2.

Preliminarily we remark that by proposition 7.1.2 the continuity of the elastic stiffness $\mathbf{E} = \mathbf{C}_o{}^{-1}$ is ensured by the continuity of $\mathbf{C}$ and the closedness of $\operatorname{Im}\mathbf{C}$. The continuity of $\mathbf{E} \in BL(Ker\,\mathbf{P}, Ker\,\mathbf{P})$ implies the continuity of $\mathbf{A} = \mathbf{K} + \mathbf{B}'\mathbf{E}\mathbf{B}$ so that $\mathbf{A} \in BL(Ker\,\mathbf{PB}, \mathcal{F})$.

The bilinear form $\mathbf{a}$ is then continuous on $Ker\,\mathbf{PB} \times \mathcal{V}$ and hence *a fortiori* on $Ker\,\mathbf{PB} \times Ker\,\mathbf{PB}$.

We then consider the canonical surjection $\boldsymbol{\Pi} \in BL(\mathcal{F}, \mathcal{F}/(Ker\,\mathbf{PB})^{\perp})$ and define

- the reduced elastic stiffness $\mathbf{A}_o := \boldsymbol{\Pi}\mathbf{A} \in BL(Ker\,\mathbf{PB}, \mathcal{F}/(Ker\,\mathbf{PB})^{\perp})$
- the reduced effective load $\ell_o := \boldsymbol{\Pi}\ell \in \mathcal{F}/(Ker\,\mathbf{PB})^{\perp}$

or explicitly

$$\mathbf{A}_o\mathbf{u}^* := \mathbf{A}\mathbf{u}^* + (Ker\,\mathbf{PB})^{\perp} \quad \forall\,\mathbf{u}^* \in Ker\,\mathbf{PB} \quad \textit{and} \quad \ell_o := \ell + (Ker\,\mathbf{PB})^{\perp}.$$

The following result is a direct consequence of the discussion carried out in section 8.2.

Proposition 8.4.4 (Well-posedness of the reduced problem) *The symmetric linear problem*

$$\mathbb{P}^*) \quad \mathbf{A}_o\mathbf{u}^* = \ell_o \quad \mathbf{u}^* \in Ker\,\mathbf{PB}\,.$$

is well-posed if and only if $\operatorname{Im} \mathbf{A}_o$ *is closed in* $\mathcal{F}/(Ker\,\mathbf{PB})^{\perp}$*. This closure property is equivalent to the closedness of the symmetric form* $\mathbf{a}$ *on* $Ker\,\mathbf{PB} \times Ker\,\mathbf{PB}$ *and is expressed by the inf-sup condition*

$$\inf_{\mathbf{u}^* \in Ker\,\mathbf{PB}} \sup_{\mathbf{v}^* \in Ker\,\mathbf{PB}} \frac{\mathbf{a}(\mathbf{u}^*, \mathbf{v}^*)}{\|\mathbf{u}^*\|_{\mathcal{V}/Ker\,\mathbf{A}_o} \|\mathbf{v}^*\|_{\mathcal{V}/Ker\,\mathbf{A}_o}} > 0\,.$$

The existence of a solution is thus garanteed if and only if $\ell_o \in (Ker\,\mathbf{A}_o)^{\perp}$ *and the solution is unique to within elements of* $Ker\,\mathbf{A}_o$*.*

The positivity of the elastic compliance $\mathbf{C}$ in $\mathcal{S}$ implies that the elastic stiffness $\mathbf{E} = \mathbf{C}_o^{-1}$ is positive definite on $\operatorname{Im} \mathbf{C}$. On this basis the next result provides an important formula for $Ker\,\mathbf{A}_o$.

Proposition 8.4.5 (Kernel of the reduced stiffness) *Let the forms* $\mathbf{c}$ *and* $\mathbf{k}$ *be symmetric and positive. The kernel of the reduced stiffness operator* $\mathbf{A}_o$ *is then given by*

$$Ker\,\mathbf{A}_o = Ker\,\mathbf{B} \cap Ker\,\mathbf{K}$$

Proof. By definition the elements of $Ker\,\mathbf{A}_o$ are the kinematic fields $\mathbf{u}^* \in Ker\,\mathbf{PB}$ which meet the variational condition

$$\mathbf{k}(\mathbf{u}^*, \mathbf{v}^*) + (\!(\, \mathbf{EBu}^*\,, \mathbf{Bv}^* \,)\!) = 0 \quad \forall\, \mathbf{v}^* \in Ker\,\mathbf{PB}.$$

Setting $\mathbf{v}^* = \mathbf{u}^* \in Ker\,\mathbf{PB}$ we get

$$\mathbf{k}(\mathbf{u}^*, \mathbf{u}^*) + (\!(\, \mathbf{EBu}^*\,, \mathbf{Bu}^* \,)\!) = 0.$$

Both terms, being non negative, must vanish. Hence by the positive definiteness of $\mathbf{E}$ on $\operatorname{Im} \mathbf{C}$ we have that $\mathbf{u}^* \in Ker\,\mathbf{B}$.

By the positivity of $\mathbf{k}$ in $\mathcal{V}$ and the condition $\mathbf{k}(\mathbf{u}^*, \mathbf{u}^*) = 0$ we infer that the field $\mathbf{u}^* \in Ker\,\mathbf{PB}$ is an absolute minimum point of $\mathbf{k}$ in $\mathcal{V}$. Taking the directional derivative along an arbitrary direction $\mathbf{v} \in \mathcal{V}$ by the symmetry of $\mathbf{k}$ we get

$$\mathbf{k}(\mathbf{u}^*, \mathbf{v}) = 0 \quad \forall\, \mathbf{v} \in \mathcal{V} \iff \mathbf{Ku}^* = \mathbf{o} \iff \mathbf{u}^* \in Ker\,\mathbf{K}$$

and the result is proved. ■

By the representation formula of $Ker\,\mathbf{A}_o$ provided in the previous proposition the admissibility condition on the data of problem $\mathbb{P}^*$ can be written

$$\ell_o \in (Ker\,\mathbf{B} \cap Ker\,\mathbf{K})^{\perp}.$$

Now for any pair $\{\mathbf{u}_o\,, \boldsymbol{\delta}^*\} \in \mathcal{V} \times \mathrm{Im}\,\mathbf{C}$ we have

$$(\!(\, \mathbf{E}\boldsymbol{\delta}*\,,\, \mathbf{Bv} \,)\!) - \mathbf{k}\,(\mathbf{u}_o, \mathbf{v}) = 0 \quad \forall\, \mathbf{v} \in Ker\,\mathbf{B} \cap Ker\,\mathbf{K}.$$

The admissibility condition on ℓ_o amounts then to the orthogonality requirement

$$\mathbf{f} \in (Ker\,\mathbf{B} \cap Ker\,\mathbf{K})^{\perp}.$$

On the other hand, when the pair $\{\mathbf{u}_o\,, \boldsymbol{\delta}^*\}$ ranges in $\mathcal{V} \times \mathrm{Im}\,\mathbf{C}$, the corresponding distorsion $\boldsymbol{\delta} = \mathbf{B}\mathbf{u}_o + \boldsymbol{\delta}^*$ will range over the whole subspace $\mathrm{Im}\,\mathbf{B} + \mathrm{Im}\,\mathbf{C}$ and this subspace, by the assumed closedness of $\mathrm{Im}\,\mathbf{PB}$, coincides with $(Ker\,\mathbf{B}' \cap Ker\,\mathbf{C})^{\perp}$.

In conclusion the admissibility condition

$$\mathbf{f} \in (Ker\,\mathbf{B} \cap Ker\,\mathbf{K})^{\perp}, \qquad \{\mathbf{u}_o\,, \boldsymbol{\delta}^*\} \in \mathcal{V} \times \mathrm{Im}\,\mathbf{C},$$

for the data of problem $\mathbb{P}^*$ coincides with the admissibility condition

$$\mathbf{f} \in (Ker\,\mathbf{B} \cap Ker\,\mathbf{K})^{\perp}, \qquad \boldsymbol{\delta} \in (Ker\,\mathbf{B}' \cap Ker\,\mathbf{C})^{\perp},$$

for the corresponding data of the mixed problem $\mathbb{M}$.

The previous results are summarized in the following theorem.

Proposition 8.4.6 (Well-posedness conditions for the mixed problem) *Let the continuous bilinear form* $\mathbf{k}$ *be positive and symmetric on* $\mathcal{V} \times \mathcal{V}$ *and the continuous bilinear form* $\mathbf{c}$ *be positive, symmetric and closed on* $\mathcal{S} \times \mathcal{S}$. *The mixed elastostatic problem* $\mathbb{M}$ *is well-posed if and only if the following two conditions are fulfilled:*

a_1*) The image of* $\mathbf{PB}$ *is closed in* $\mathcal{D}$*, that is,* $\mathrm{Im}\,\mathbf{B} + \mathrm{Im}\,\mathbf{C}$ *is closed in* $\mathcal{D}$*, i.e.*

$$\inf_{\mathbf{u} \in \mathcal{V}} \sup_{\boldsymbol{\sigma} \in \mathcal{S}} \frac{(\!(\, \boldsymbol{\sigma}\,,\, \mathbf{PBu} \,)\!)}{\|\boldsymbol{\sigma}\|_{\mathcal{S}/(Ker\,\mathbf{B}'\mathbf{P}')} \|\mathbf{u}\|_{\mathcal{V}/(Ker\,\mathbf{PB})}} =$$

$$\inf_{\boldsymbol{\sigma} \in \mathcal{S}} \sup_{\mathbf{u} \in \mathcal{V}} \frac{(\!(\, \boldsymbol{\sigma}\,,\, \mathbf{PBu} \,)\!)}{\|\boldsymbol{\sigma}\|_{\mathcal{S}/(Ker\,\mathbf{B}'\mathbf{P}')} \|\mathbf{u}\|_{\mathcal{V}/(Ker\,\mathbf{PB})}} > 0,$$

a_2) the bilinear form of the elastic energy is closed on $Ker\,\mathbf{PB} \times Ker\,\mathbf{PB}$*, i.e.*

$$\inf_{\mathbf{u}^* \in Ker\,\mathbf{PB}} \sup_{\mathbf{v}^* \in Ker\,\mathbf{PB}} \frac{\mathbf{k}\,(\mathbf{u}^*, \mathbf{v}^*) + (\!(\, \mathbf{EBu}^*\,, \mathbf{Bv}^*\,)\!)}{\|\mathbf{u}^*\|_{\mathcal{V}/(Ker\,\mathbf{B} \cap Ker\,\mathbf{K})} \|\mathbf{v}^*\|_{\mathcal{V}/(Ker\,\mathbf{B} \cap Ker\,\mathbf{K})}} > 0.$$

In other terms conditions a_1) and a_2) are equivalent to state that the structural operator $\mathbf{S} \in BL\,(\mathcal{V} \times \mathcal{S}, \mathcal{F} \times \mathcal{D})$ has a closed range so that the orthogonality condition $\operatorname{Im}\mathbf{S} = (Ker\,\mathbf{S})^{\perp}$ holds.

Applicable sufficient criteria for the fulfilment of the conditions a_1) and a_2) will be discussed in the next section.

8.5 Sufficient criteria

Proposition 8.4.6 provides a set of two necessary and sufficient conditions for the well-posedness of a general elastic problem. More precisely condition a_1) states the equivalence of the mixed problem

$$\mathbb{M})\quad \begin{cases} \mathbf{k}\,(\mathbf{u}, \mathbf{v}) + \mathbf{b}\,(\mathbf{v}, \boldsymbol{\sigma}) = \langle \mathbf{f}, \mathbf{v} \rangle & \mathbf{u} \in \mathcal{V}, \quad \forall\, \mathbf{v} \in \mathcal{V}, \\ \mathbf{b}\,(\mathbf{u}, \boldsymbol{\tau}) - \mathbf{c}\,(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \langle \boldsymbol{\delta}, \boldsymbol{\tau} \rangle & \boldsymbol{\sigma} \in \mathcal{S}, \quad \forall\, \boldsymbol{\tau} \in \mathcal{S}. \end{cases}$$

to the reduced problem $\mathbb{P}^*$ and condition a_2) provides the well-posedness of problem $\mathbb{P}^*$.

Let us now discuss these two conditions in detail.

8.5.1 Discussion of condition a_1

By remark 8.4.2 the condition a_1) can be stated in the equivalent forms

- the subspace $\operatorname{Im}\mathbf{PB}$ is closed in $\mathcal{D}$,
- the subspace $\operatorname{Im}\mathbf{B}'\mathbf{P}'$ is closed in $\mathcal{F}$,
- the sum $Ker\,\mathbf{P} + \operatorname{Im}\mathbf{B} = \operatorname{Im}\mathbf{C} + \operatorname{Im}\mathbf{B}$ is closed in $\mathcal{D}$.

Condition a_1) is trivially fulfilled by the structural models belonging to one of the two extreme cathegories:

i) the elastic compliance is not singular, so that $Ker\,\mathbf{C} = \mathbf{o}\{\}$ and $\mathbf{P} = \mathbf{O}$,

ii) the elastic compliance is null, so that $Ker\,\mathbf{C} = \mathcal{S}$ and $\mathbf{P} = \mathbf{I}$. Case *i*) corresponds to classical elasticity problems in which every stress field is elastically effective.

Case *ii*) corresponds to the opposite situation in which every stress field is elastically ineffective. The statics of a rigid structure resting on elastic supports is described by an elastic problem of the this kind whose mixed formulation is

$$\mathbb{F}) \quad \begin{cases} \mathbf{k}\,(\mathbf{u},\mathbf{v}) + \mathbf{b}\,(\mathbf{v},\boldsymbol{\sigma}) = \langle \mathbf{f},\mathbf{v} \rangle & \mathbf{u} \in \mathcal{V}, \quad \forall\, \mathbf{v} \in \mathcal{V}, \\ \mathbf{b}\,(\mathbf{u},\boldsymbol{\tau}) = \langle \boldsymbol{\delta},\boldsymbol{\tau} \rangle & \boldsymbol{\sigma} \in \mathcal{S}, \quad \forall\, \boldsymbol{\tau} \in \mathcal{S}. \end{cases}$$

This is exactly the saddle point problem first analysed by BREZZI in [18].

The existence and uniqueness proof contributed in [18] addressed the more general case in which the bilinear form $\mathbf{k}$ in problem $\mathbb{F}$ was neither positive nor symmetric.

A discussion of the general mixed problem

$$\mathbb{G}) \quad \begin{cases} \mathbf{k}\,(\mathbf{u},\mathbf{v}) + \mathbf{h}\,(\mathbf{v},\boldsymbol{\sigma}) = \langle \mathbf{f},\mathbf{v} \rangle & \mathbf{u} \in \mathcal{U}, \quad \forall\, \mathbf{v} \in \mathcal{V}, \\ \mathbf{b}\,(\mathbf{u},\boldsymbol{\tau}) - \mathbf{c}\,(\boldsymbol{\sigma},\boldsymbol{\tau}) = \langle \boldsymbol{\delta},\boldsymbol{\tau} \rangle & \boldsymbol{\sigma} \in \Sigma, \quad \forall\, \boldsymbol{\tau} \in \mathcal{S}. \end{cases}$$

in which the bilinear forms $\mathbf{k}$ and $\mathbf{c}$ are neither positive nor symmetric, is carried out by ROMANO et al. in [157]. The results contributed in [157] include as special cases the existence and uniqueness theorem by BREZZI and its extensions due to NICOLAIDES [128] and BERNARDI et al. [14] in which the bilinear form $\mathbf{c}$ was absent.

Remark 8.5.1 *The analysis performed in the previous section addressed the general case of an elastic mixed problem* $\mathbb{M}$ *with a possibly non-degenerate kernel of the structural operators* $\mathbf{S}$*. Structural problems in which the kernel of* $\mathbf{S}$ *in non-degenerate are usually dealt with in the engineering applications. An example is provided by elastic problems in which rigid kinematic fields not involving reactions of the elastic supports are admitted by the constraints.*

To deal with the presence of a non-degenerate kernel, the symmetry of the governing operator $\mathbf{S}$ *and the positivity of the elastic operators* $\mathbf{K}$ *and* $\mathbf{C}$ *seem however to be unavoidable assumptions. They play in fact an essential role in deriving the representation formulas for the kernels provided in section 8.4.1 and Proposition 8.4.4.*

Remark 8.5.2 *It is worth noting that, for two- or three-dimensional non rigid structural models with a singular elastic compliance, condition* a_1) *is difficult to be checked and is far from being verified as a rule.*

A relevant exception is provided by the incompressibility constraint of Stokes *problem ([88], [182]). We emphasize that a singularity of the elastic compliance* **C** *is equivalent to the imposition of constraints on the strain fields. Strain constraints in continua have been recently discussed by* Antman *and* Marlow *in [5], [105] and critically reviewed by* Romano *et al. in [156].*

8.5.2 Discussion of condition a_2

Under the assumption that the bilinear form $\mathbf{k}$ is $Ker\,\mathbf{PB}$-semielliptic, and hence closed on $Ker\,\mathbf{PB} \times Ker\,\mathbf{PB}$, the next result yields a sufficient criterion for the fulfilment of condition a_2).

Proposition 8.5.1 *Condition a_2) is satisfied if the following properties hold*

$$i)\quad \mathbf{k}(\mathbf{u},\mathbf{u}) \geq c_{\mathbf{k}}\, \|\mathbf{u}\|^2_{\mathcal{V}/Ker\,\mathbf{K}} \qquad c_{\mathbf{k}} > 0 \quad \forall\, \mathbf{u} \in Ker\,\mathbf{PB}$$

$$ii)\quad ((\, \mathbf{EBu}\,,\, \mathbf{Bu}\,)) \geq c\, \|\mathbf{u}\|^2_{\mathcal{V}/Ker\,\mathbf{B}} \qquad c > 0 \quad \forall\, \mathbf{u} \in Ker\,\mathbf{PB}$$

$$iii)\quad Ker\,\mathbf{B} + Ker\,\mathbf{K} \ \textit{closed.}$$

Proof. By Theorem 7.1.4 and Remark 7.1.3 property (iii) is equivalent to

$$\|\mathbf{u}\|^2_{\mathcal{V}/\mathrm{Ker}\,\mathbf{K}} + \|\mathbf{u}\|^2_{\mathcal{V}/\mathrm{Ker}\,\mathbf{B}} \geq \alpha\, \|\mathbf{u}\|^2_{\mathcal{V}/(\mathrm{Ker}\,\mathbf{K} \cap \mathrm{Ker}\,\mathbf{B})} \quad \forall\, \mathbf{u} \in \mathcal{V}$$

so that, summing up (i) and (ii), we get

$$\mathbf{k}(\mathbf{u},\mathbf{u}) + ((\, \mathbf{EBu}\,,\, \mathbf{Bu}\,)) \geq c_{\mathbf{a}}\, \|\mathbf{u}\|^2_{\mathcal{V}/(\mathrm{Ker}\,\mathbf{K} \cap \mathrm{Ker}\,\mathbf{B})} \quad \forall\, \mathbf{u} \in Ker\,\mathbf{PB}\,,$$

with a suitably positive constant $c_{\mathbf{a}}$. This implies the closedness condition a_2). ■

- Condition i) is fulfilled in structural problems with discrete external elastic constraints. In fact when only a finite number of external elastic constraints are imposed, the subspace $\mathrm{Im}\,\mathbf{K}$ is finite dimensional and the constant $c_{\mathbf{k}}$ is provided by the smallest positive eigenvalue of the symmetric positive matrix associated with the restriction of the bilinear form $\mathbf{k}$ to $\mathcal{V}/Ker\,\mathbf{K} \times \mathcal{V}/Ker\,\mathbf{K}$. An example is provided by an elastic plate resting on a finite number of elastic supports, as shown in Fig.8.1
- Condition ii) follows from a standard ellipticity property of internal elasticity:
$$((\, \mathbf{C}\boldsymbol{\sigma}\,,\, \boldsymbol{\sigma}\,)) \geq c_{\boldsymbol{\sigma}}\, \|\boldsymbol{\sigma}\|^2_{\mathcal{S}/\mathrm{Ker}\,\mathbf{C}} \quad \forall\, \boldsymbol{\sigma} \in \mathcal{S},$$

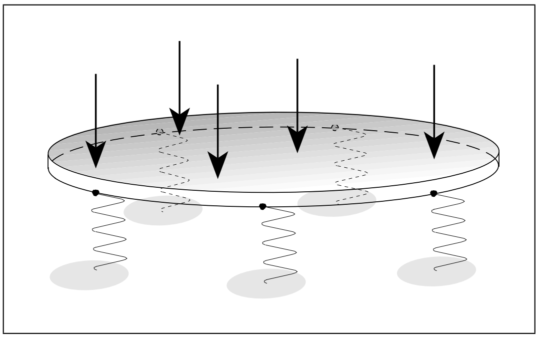

Figure 8.1: Elastic plate on a finite number of elastic supports $\dim \operatorname{Im} \mathbf{K} < +\infty$.

equivalent to

$$(\!(\, \mathbf{E}\boldsymbol{\varepsilon} \, , \, \boldsymbol{\varepsilon} \,)\!) \geq \, c_{\mathbf{e}} \, \|\boldsymbol{\varepsilon}\|^2_{\mathcal{D}} \quad \forall \, \boldsymbol{\varepsilon} \in \mathit{Ker} \, \mathbf{P} = \operatorname{Im} \mathbf{C},$$

and from the closedness of the fundamental form $\mathbf{b}\,(\mathbf{u}, \boldsymbol{\sigma})$

$$\|\mathbf{B}\mathbf{u}\|_{\mathcal{D}} \geq \, c_{\mathbf{b}} \, \|\mathbf{u}\|_{\mathcal{V}/\mathrm{Ker}\,\mathbf{B}} \quad \forall \, \mathbf{u} \in \mathcal{V}.$$

The positive constant in (ii) is given by $c = \, c_{\mathbf{e}} \, c^2_{\mathbf{b}}$.

- Condition $iii)$ is a consequence of the finite dimensionality of $\mathit{Ker}\,\mathbf{B}$ in most structural models. More generally it follows from the closedness condition in proposition 8.3.2.

8.6 Elastic beds

Let us finally consider the general problem of the elastic equilibrium of a structural model in which

- the constitutive behaviour is partially rigid,
- the external elastic constraints include the presence of elastic beds so that $\operatorname{Im}\mathbf{K}$ is not finite dimensional in $\mathcal{F}$.

An example is provided by an elastic plate resting on an elastic bed, as in fig. 8.2. Such a model is commonly adopted in engineering applications to simulate a foundation interacting with a supporting soil.

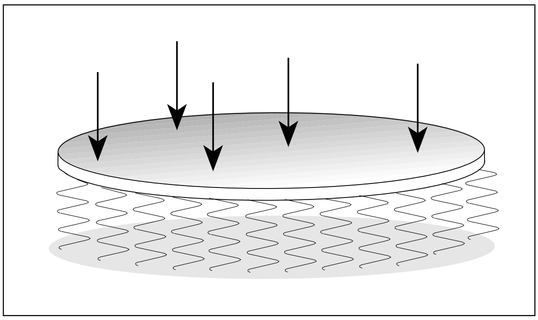

Figure 8.2: Elastic plate resting on an elastic bed.

The difficulty connected with this kind of problems lies in the fact that the bilinear form of the external elastic energy is not semi-elliptic on $\mathcal{V} \times \mathcal{V}$ as required by condition $i)$ of Proposition 8.5.1.

To enlight the problem let us consider the model of an elastic beam resting on an elastic bed of springs (Winkler soil model). The flexural elastic energy of the beam is provided by one-half the integral of the squared second derivative of the transverse displacement. On the other hand, the elastic energy stored into the elastic springs is equal to one-half the integral of the squared transverse displacement. The kinematic space $\mathcal{V}$ is defined to be the SOBOLEV space $\mathcal{H}^2$ to ensure a finite value of the elastic energy. Considering a rapidly varying elastic curve of the beam, as depicted in fig. 8.3, we get an extremely high value of the elastic energy in the beam and a negligible energy in the elastic bed.

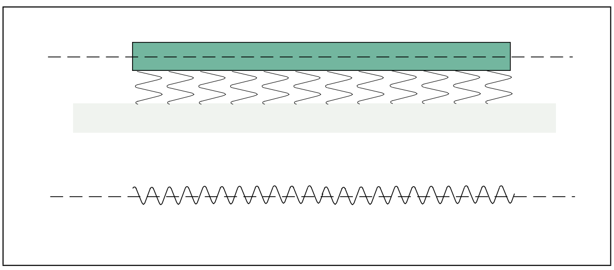

Figure 8.3: Large elastic energy with small displacements.

The discussion above leads to the conclusion that the semi-ellipticity condition on the bilinear form $\mathbf{k}$ of elastic constraints energy must be relaxed.

A by far less stringent requirement is the property that $\mathbf{k}$ is positive semi-definite on $Ker\,\mathbf{PB} \times Ker\,\mathbf{PB}$ and semi-elliptic only on $Ker\,\mathbf{B} \times Ker\,\mathbf{B}$, that is with respect to rigid kinematic fields, according to the inequalities:

$$i) \quad \mathbf{k}(\mathbf{u},\mathbf{u}) \geq 0\,, \qquad \forall\,\mathbf{u} \in Ker\,\mathbf{PB}\,,$$

$$ii) \quad \mathbf{k}(\mathbf{u},\mathbf{u}) \geq\; c_{\mathbf{k}}\, \|\mathbf{u}\|^2_{\mathcal{V}/\mathrm{Ker}\,\mathbf{K}}\,, \quad c_{\mathbf{k}} > 0\,, \quad \forall\,\mathbf{u} \in Ker\,\mathbf{B}\,.$$

We remark that rigid kinematic fields cannot undergo very sauvage oscillations. In the case of the simple beam of 8.3 they are in fact affine functions. In general, when $Ker\,\mathbf{B}$ is finite dimensional, property $ii)$ above is a consequence of property $i)$ since $c_{\mathbf{k}} > 0$ is the smallest positive eigenvalue of a non-null symmetric and positive matrix. We have now to prove that these less stringent assumptions on $\mathbf{k}$ are sufficient to ensure the fulfilment of condition a_2.

To this end we provide a preliminary result.

Proposition 8.6.1 (The elastic bed inequality) *The assumptions*

$$i) \quad \mathbf{k}(\mathbf{u},\mathbf{u}) \geq 0\,, \quad \forall\,\mathbf{u} \in Ker\,\mathbf{PB}\,,$$

$$ii) \quad \mathbf{k}(\mathbf{u},\mathbf{u}) \geq\; c_{\mathbf{k}}\, \|\mathbf{u}\|^2_{\mathcal{V}/Ker\,\mathbf{K}}\,, \quad \forall\,\mathbf{u} \in Ker\,\mathbf{B}\,,$$

$$iii) \quad (\!(\, \mathbf{EBu}\,,\, \mathbf{Bu}\,)\!) \geq\; c_{\mathbf{e}}\, c^2_{\mathbf{b}}\, \|\mathbf{u}\|^2_{\mathcal{V}/Ker\,\mathbf{B}}\,, \quad \forall\,\mathbf{u} \in Ker\,\mathbf{PB}\,,$$

ensure the validity of the inequality

$$\mathbf{k}(\mathbf{u},\mathbf{u}) + (\!(\, \mathbf{EBu}\,,\, \mathbf{Bu}\,)\!) \geq\; c_{\boldsymbol{\pi}}\, \|\mathbf{\Pi u}\|^2_{\mathcal{V}/Ker\,\mathbf{K}} \quad c_{\boldsymbol{\pi}} > 0\,, \quad \forall\,\mathbf{u} \in Ker\,\mathbf{PB}\,,$$

where $\mathbf{\Pi}$ *denotes the orthogonal projector on* $Ker\,\mathbf{B}$ *in* $\mathcal{V}$.

Proof. We proceed *per absurdum* by assuming that the inequality is false. Then, prescribing that $\|\mathbf{\Pi u}\|_{\mathcal{V}/\mathrm{Ker}\,\mathbf{K}} = 1$, the infimum of the first member would be zero. By taking a minimizing sequence $\{\mathbf{u}_n\}$ we have

$$\lim_{n\to\infty} \mathbf{k}(\mathbf{u}_n,\mathbf{u}_n) + (\!(\, \mathbf{EBu}_n\,,\, \mathbf{Bu}_n\,)\!) = 0\,.$$

By (i) both terms of the sum are non-negative and then vanish at the limit. Hence from (iii) we get

$$\lim_{n\to\infty} (\!(\, \mathbf{EBu}_n\,,\, \mathbf{Bu}_n\,)\!) = 0 \implies \lim_{n\to\infty} \|\mathbf{u}_n - \mathbf{\Pi u}_n\|_{\mathcal{V}} = 0\,,$$

and by the continuity of $\mathbf{k}$ and assumption (ii)

$$\lim_{n\to\infty} \mathbf{k}(\mathbf{u}_n,\mathbf{u}_n) = 0 \implies \lim_{n\to\infty} \mathbf{k}(\mathbf{\Pi u}_n,\mathbf{\Pi u}_n) = 0$$

$$\implies \lim_{n\to\infty} \|\mathbf{\Pi u}_n\|_{\mathcal{V}/\mathrm{Ker}\,\mathbf{K}} = 0\,,$$

contrary to the assumption that $\|\mathbf{\Pi u}\|_{\mathcal{V}/\mathrm{Ker}\,\mathbf{K}} = 1$. ■

An applicable criterion for the validity of condition a_2) is now at hand.

Proposition 8.6.2 *Condition a_2) is satisfied if the following properties hold*

$$\begin{aligned} &i) \quad \mathbf{k}(\mathbf{u},\mathbf{u}) \geq 0\,, \quad \forall\,\mathbf{u} \in \mathit{Ker}\,\mathbf{PB}, \\ &ii) \quad \mathbf{k}(\mathbf{u},\mathbf{u}) \geq c_{\mathbf{k}}\,\|\mathbf{u}\|^2_{\mathcal{V}/\mathit{Ker}\,\mathbf{K}}\,, \quad \forall\,\mathbf{u} \in \mathit{Ker}\,\mathbf{B}\,, \\ &iii) \quad (\!(\,\mathbf{EBu}\,,\,\mathbf{Bu}\,)\!) \geq c_{\mathbf{e}}\,c^2_{\mathbf{b}}\,\|\mathbf{u}\|^2_{\mathcal{V}/\mathit{Ker}\,\mathbf{B}}\,, \quad \forall\,\mathbf{u} \in \mathit{Ker}\,\mathbf{PB}\,, \\ &iv) \quad \mathit{Ker}\,\mathbf{B} + \mathit{Ker}\,\mathbf{K} \text{ is closed}\,. \end{aligned}$$

Proof. Theorem 7.1 and Remark 7.1.3 ensure that property *iv*) imply the existence of a constant $\alpha > 0$ such that

$$\|\mathbf{\Pi u}\|^2_{\mathcal{V}/\mathrm{Ker}\,\mathbf{K}} + \|\mathbf{u}\|^2_{\mathcal{V}/\mathrm{Ker}\,\mathbf{B}} \geq \alpha\,\|\mathbf{u}\|^2_{\mathcal{V}/(\mathrm{Ker}\,\mathbf{K}\cap\mathrm{Ker}\,\mathbf{B})}\,, \quad \forall\,\mathbf{u} \in \mathcal{V}\,,$$

Condition a_2) then follows by adding inequality (*iii*) and the one proved in proposition 8.6.1. ■

8.6.1 A well-posedness criterion

Conditions $i), ii), iii)$ of proposition 8.6.2 are always fulfilled by elastic structural models.

Further, by remark 7.1.2, the closedness of $\mathrm{Im}\,\mathbf{B}$ ensures that condition a_1) can be equivalently stated by requiring the closedness of $\mathit{Ker}\,\mathbf{C} + \mathit{Ker}\,\mathbf{B}'$.

Then, to get a well posed mixed problem, what we really have to check is the fulfilment of the two properties concerning the kernels of the elastic operators, as stated in the next proposition.

Proposition 8.6.3 (Well-posedness criterion) *Let* $\mathrm{Im}\,\mathbf{B}$ *be closed in* $\mathcal{D}$ *and conditions* $i), ii), iii)$ *of proposition 8.6.2 be fulfilled. Then the closedness properties:*

- $\mathit{Ker}\,\mathbf{B}' + \mathit{Ker}\,\mathbf{C}$ *is closed in* $\mathcal{S}$,
- $\mathit{Ker}\,\mathbf{B} + \mathit{Ker}\,\mathbf{K}$ *is closed in* $\mathcal{V}$,

ensure that the mixed elastostatic problem $\mathbb{M}$ *is well posed.*

By virtue of proposition 7.1.3 a relevant situation in which condition a_1) and b) are fulfilled is provided by the following family of structural models.

Definition 8.6.1 (Simple structures) *A structural model is said to be simple if the subspaces $Ker\,\mathbf{B}$ of rigid kinematisms and $Ker\,\mathbf{B}'$ of self-equilibrated stress fields are finite dimensional.*

All one-dimensional engineering structural models composed by beam and bar elements belong to this class and hence the related elastic problems are always well-posed. A simple frame composed of two beams which are axially undeformable and flexurally elastic is depicted hereafter. The stress fields are pairs of diagrams of bending moments and axial forces.

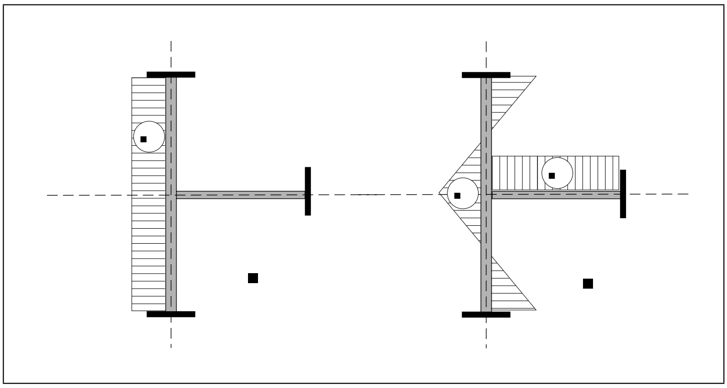

Figure 8.4: Flexurally elastic and axially rigid beams

Fig. 8.4 a) shows the diagram of axial forces in the vertical beam which corresponds to a self-equilibrated and elastically ineffective stress field. It cannot be evaluated by solving the elastic problem.

Since this stress field generates the whole subspace $Ker\,\mathbf{B}' \cap Ker\,\mathbf{C}$ the imposed distorsions $\boldsymbol{\delta}$ must satisfy the related orthogonality condition which requires that the mean elongation of the vertical beam must vanish. Fig. 8.4 b) shows a diagram of axial forces and bending moments which is self-equilibrated but elastically effective.

A beam on elastic supports is sketched in fig. 8.5 and 8.6 to show examples of kinematic fields which respectively belong to $Ker\,\mathbf{K}$ and to $Ker\,\mathbf{B} \cap Ker\,\mathbf{K}$.

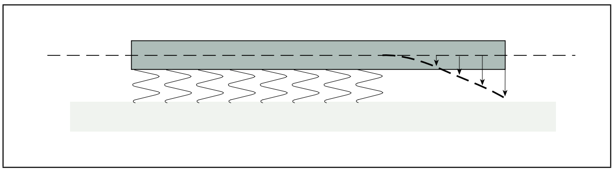

Figure 8.5: $\mathbf{u} \in Ker\,\mathbf{K}$

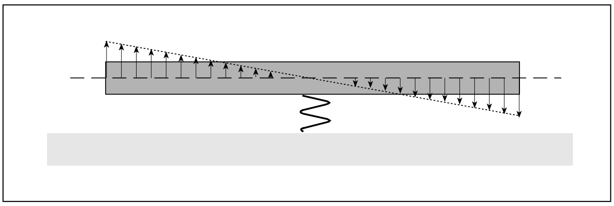

Figure 8.6: $\mathbf{u} \in Ker\,\mathbf{B} \cap Ker\,\mathbf{K}$

8.7 Conclusions

The analytical properties of mixed formulations in elasticity have been investigated with an approach which provides a clear mechanical interpretation of the properties of the model and of the conditions for its well-posedness.

Necessary and sufficient conditions for the existence of a solution have been proved and effective criteria for their application have been contributed. In particular we have shown that all familiar one-dimensional engineering models of structural assemblies composed of bars and beams fulfill the well-posedness property in the presence of any singularity of the elastic compliance.

The case of two- or three-dimensional structural models drastically changes the scenary due to the infinite dimensionality of the subspace of self-stresses so that well-posedness of the mixed problem will be almost never fulfilled when the elastic compliance is singular. As relevant exceptions we quote structural models with either fully elastic or perfectly rigid behaviour. The problem is strictly connected with the discussion of constrained structural models in which a linear constraint is imposed on the strain fields. By means of simple counterexamples [105], [156] it can be shown that there is little hope to get well-posedness of a mixed problem when the elastic compliance is singular. In this respect Stokes problem concerning the incompressible viscous flow of fluids, for which well-posedness is fulfilled, must be considered as an exception.

Although the analysis has beeen carried out with explicit reference to elastostatic problems, we observe that the results can as well be applied to the discussion of a number of interesting problems in mathematical physics modeled by analogous mixed formulations. A variant of the proposed approach can also be applied to the discussion of problems in which linear constraints are imposed on the stress field.

Chapter 9

Approximate models

The actual demand of designing structures with more and more complex geometrical shapes and constitutive behaviors, have led to a deep study of computational methods based on discretizations of a continuous model. The main issues under investigation are the interpolation properties related to discretization criteria and error estimates concerned with the evaluation of the gap between the discrete solution and the continuous one. In this chapter essential aspects of discretization and error estimate methods are illustrated.

9.1 Discrete mixed models

From a mathematical point of view, the formulation of a discrete structural model associated with a given linear continuous model, consists in imposing that the dispacement, stress and deformation fields belong to finite dimensional linear subspaces of the linear state-spaces. Discretization may then be interpreted as a linear constraint imposed on the state-variables by providing explicit representations of the linear subspaces of conforming state variables. To describe the procedure in detail, let us consider a structural model $\mathcal{M}(\mathbf{\Omega}, \mathcal{L}, \mathbf{B})$ and the interpolating subspaces

$$\begin{aligned} \mathcal{L}_h &\subset \mathcal{L}\,, \quad \text{discrete displacements}\,, \\ \mathcal{S}_h &\subset \mathcal{H}\,, \quad \text{discrete stresses}\,, \end{aligned}$$

defining an associated *discrete model* $\mathcal{M}(\mathbf{\Omega}, \mathcal{L}, , \mathbf{B}, \mathcal{L}_h, \mathcal{S}_h)$.

Definition 9.1.1 (Discrete active force systems) *A discrete active force system is a functional of the* BANACH *space* $\mathcal{F}_h$ *, dual of the* BANACH *space* $\mathcal{L}_h \subset \mathcal{L}$ *of discrete displacements, according to the topology induced by the* BANACH *space* $\mathcal{L}$ *.*

Definition 9.1.2 (Discrete reactive force systems) *A discrete reactive force system is a functional of the* BANACH *space* $\mathcal{R}_h = \overset{\circ}{\mathcal{L}_h} \subset \mathcal{F}_h$ *where* $\overset{\circ}{\mathcal{L}_h}$ *is the annihilator defined by*

$$\overset{\circ}{\mathcal{L}_h} := \{\mathbf{f} \in \mathcal{F}_h \mid \langle \mathbf{f}, \mathbf{v} \rangle = 0 \quad \forall\, \mathbf{v} \in \mathcal{L}_h\}\,.$$

As is well-known, there is an isometric isomorphism between the space $\mathcal{F}_h$, dual to $\mathcal{L}_h$, and the quotient space $\mathcal{F}/\overset{\circ}{\mathcal{L}_h}$, see e.g. Proposition I.9.18 (p.75) in [159]. It is often convenient to identify the spaces $\mathcal{F}_h$ and $\mathcal{F}/\overset{\circ}{\mathcal{L}_h}$. This identification allows to provide a straightforward and simple interpretation of force systems acting on the discrete model and it will be adopted without further future advice. The advantages related to the identification between force systems on the discrete model and the affine manifolds of force systems on the continuous model, are apparent if the analysis carried out in the sequel is compared with the one illustrated in [20].

9.1.1 Equilibrium

Discrete operators, kinematic and static, in duality

$$\mathbf{B}_h \in BL\,(\mathcal{L}_h\,; \mathcal{H}/\mathcal{S}_h^{\perp})\,, \qquad \mathbf{B}'_h \in BL\,(\mathcal{S}_h\,; \mathcal{F}/\overset{\circ}{\mathcal{L}_h})$$

are defined by

$$\begin{aligned} \mathbf{B}_h \mathbf{u}_h &= \mathbf{B}\,\mathbf{u}_h + \mathcal{S}_h^{\perp}\,, \quad \forall\, \mathbf{u}_h \in \mathcal{L}_h\,, \\ \mathbf{B}'_h \boldsymbol{\sigma}_h &= \mathbf{B}' \boldsymbol{\sigma}_h + \overset{\circ}{\mathcal{L}_h}\,, \quad \forall\, \boldsymbol{\sigma}_h \in \mathcal{S}_h\,. \end{aligned}$$

- The subspace of *rigid discrete velocity fields* on the constrained discrete structure $\mathcal{M}(\boldsymbol{\Omega}, \mathcal{L}, \mathbf{B}, \mathcal{L}_h, \mathcal{S}_h)$ is

$$\mathbf{ker}(\mathbf{B}_h) = \{\, \mathbf{u}_h \in \mathcal{L}_h \;:\; \mathbf{B}\mathbf{u}_h \in \mathcal{S}_h^{\perp} \,\} = \mathcal{L}_h \cap (\mathbf{B}^{-1} \mathcal{S}_h^{\perp})\,.$$

- The subspace of *self-equilibrated discrete stresses* on the constrained discrete structure $\mathcal{M}(\boldsymbol{\Omega}, \mathcal{L}, \mathbf{B}, \mathcal{L}_h, \mathcal{S}_h)$

$$\mathbf{ker}(\mathbf{B}'_h) = \mathcal{S}_h \cap \left[{\mathbf{B}'_h}^{-1} \left[\frac{\overset{\circ}{\mathcal{L}} + \overset{\circ}{\mathcal{L}_h}}{\overset{\circ}{\mathcal{L}_h}} \right] \right] = \mathcal{S}_h \cap \left[{\mathbf{B}'}^{-1} \overset{\circ}{\mathcal{L}_h} \right]$$

is constituted by the discrete stresses $\boldsymbol{\sigma}_h \in \mathcal{S}_h$ in equilibrium with a system of *discrete reactive forces*

$$\mathbf{B}' \boldsymbol{\sigma}_h \in \mathcal{L}_h^{\circ} \subseteq \mathcal{F} \,.$$

- The variational condition

$$\langle \mathbf{f}, \mathbf{v}_h \rangle = 0 \,, \quad \forall \, \mathbf{v}_h \in \mathbf{ker}(\mathbf{B}_h) = \mathcal{L}_h \cap (\mathbf{B}^{-1} \mathcal{S}_h^{\perp})$$

assures that the equilibrium problem

$$(\!(\, \boldsymbol{\sigma}_h \, , \, \mathbf{B}_h \mathbf{v}_h \,)\!) = \langle \mathbf{f}, \mathbf{v}_h \rangle \,, \quad \boldsymbol{\sigma}_h \in \mathcal{S}_h \,, \quad \forall \, \mathbf{v}_h \in \mathcal{L}_h$$

admits a solution. The solution is unique if and only if the linear subspace $\mathbf{ker}(\mathbf{B}'_h)$ of self-equilibrated discrete stresses vanishes.

- Finally let $\mathbf{f} \in \mathcal{F}_{\mathcal{L}}$ be an active force system in equilibrium on the continuous structure $\mathcal{M}(\boldsymbol{\Omega}, \mathcal{L}, \mathbf{B})$, i.e. such that

$$\langle \mathbf{f}, \mathbf{v} \rangle = 0 \,, \quad \forall \, \mathbf{v} \in \mathbf{ker}(\mathbf{B}_{\mathcal{L}}) = \mathbf{ker}(\mathbf{B}) \cap \mathcal{L} \,.$$

Then the equilibrium condition on discrete structure

$$\langle \mathbf{f}, \mathbf{v}_h \rangle = 0 \,, \quad \forall \, \mathbf{v}_h \in \mathbf{ker}(\mathbf{B}_h) = \mathcal{L}_h \cap (\mathbf{B}^{-1} \mathcal{S}_h^{\perp}) \,,$$

is fulfilled if and only if the condition

$$\mathbf{ker}(\mathbf{B}_h) \subseteq \mathbf{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h \,,$$

holds. Let us note that, being $\mathbf{ker}(\mathbf{B}) = \mathbf{B}^{-1}\{\mathbf{o}\} \subseteq \mathbf{B}^{-1} \mathcal{S}_h^{\perp}$, we have that

$$\mathbf{ker}(\mathbf{B}_h) \supseteq \mathcal{L}_h \cap \mathbf{ker}(\mathbf{B}) = \mathbf{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h \,.$$

Proposition 9.1.1 *The equality* $\mathbf{ker}(\mathbf{B}_h) = \mathbf{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h$ *is equivalent to the property*

$$i) \qquad \forall \, \boldsymbol{\sigma} \in \mathcal{H} \quad \exists \, \boldsymbol{\sigma}_h \in \mathcal{S}_h \, : \, (\!(\, \boldsymbol{\sigma} - \boldsymbol{\sigma}_h \, , \, \mathbf{B} \mathbf{v}_h \,)\!) = 0 \,, \quad \forall \, \mathbf{v}_h \in \mathcal{L}_h \,.$$

Proof. Observing that

$$\mathbf{ker}(\mathbf{B}_h) = \mathbf{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h \iff \big[\mathbf{ker}(\mathbf{B}_h)\big]^{\perp} = \big[\mathbf{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h\big]^{\perp},$$

and being

$$\big[\mathbf{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h\big]^{\perp} = \mathcal{L}_h^{\circ} + \mathbf{B}'\mathcal{H},$$

$$\big[\mathbf{ker}(\mathbf{B}_h)\big]^{\perp} = \mathcal{L}_h^{\circ} + \mathbf{B}'\mathcal{S}_h,$$

we get

$$\mathbf{f} \in \big[\mathbf{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h\big]^{\perp} \iff \exists\, \boldsymbol{\sigma} \in \mathcal{H} : \langle \mathbf{f}, \mathbf{v}_h \rangle = (\!(\, \boldsymbol{\sigma}, \mathbf{B}\mathbf{v}_h \,)\!), \quad \forall\, \mathbf{v}_h \in \mathcal{L}_h,$$

$$\mathbf{f} \in \big[\mathbf{ker}(\mathbf{B}_h)\big]^{\perp} \iff \exists\, \boldsymbol{\sigma}_h \in \mathcal{S}_h : \langle \mathbf{f}, \mathbf{v}_h \rangle = (\!(\, \boldsymbol{\sigma}_h, \mathbf{B}\mathbf{v}_h \,)\!), \qquad \forall\, \mathbf{v}_h \in \mathcal{L}_h.$$

Hence the $i)$ holds. The converse implication is easily verifiable.

Remark 9.1.1 *A formulation and an alternative proof of proposition 9.1.1 are the following. The property $i)$ is equivalent to the condition $(\mathbf{B}\mathcal{L}_h)^{\perp} + \mathcal{S}_h = \mathcal{H}$ that in turn is equivalent to the condition $\mathbf{B}\mathcal{L}_h \cap \mathcal{S}_h^{\perp} = \mathbf{o}$, according to the proposition I.11.8 (p.88) in [159], given that the subspace $\mathbf{B}\mathcal{L}_h$ is finite dimensional and thus the sum subspace $\mathbf{B}\mathcal{L}_h + \mathcal{S}_h^{\perp}$ is closed in $\mathcal{H}$. Then we have that*

$$\begin{cases} \mathbf{u}_h \in \boldsymbol{ker}(\mathbf{B}_h) \\ \mathbf{B}\mathcal{L}_h \cap \mathcal{S}_h^{\perp} = \mathbf{o} \end{cases} \iff \begin{cases} \mathbf{B}\mathbf{u}_h \in \mathcal{S}_h^{\perp} \\ \mathbf{u}_h \in \mathcal{L}_h \\ \mathbf{B}\mathcal{L}_h \cap \mathcal{S}_h^{\perp} = \mathbf{o} \end{cases} \iff \begin{cases} \mathbf{B}\mathbf{u}_h = \mathbf{o} \\ \mathbf{u}_h \in \mathcal{L}_h. \end{cases}$$

Hence the condition $\mathbf{B}\mathcal{L}_h \cap \mathcal{S}_h^{\perp} = \mathbf{o}$ is necessary and sufficient so that the equality $\boldsymbol{ker}(\mathbf{B}_h) = \boldsymbol{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h$ holds.

9.1.2 Compatibility

Let us observe preliminarily that the *self-equilibreted discrete stress subspace* on the constrained discrete structure may be written as

$$\mathbf{ker}(\mathbf{B}'_h) = \mathcal{S}_h \cap \big[\mathbf{B}'^{-1}\mathcal{L}_h^{\circ}\big] = \mathcal{S}_h \cap \big[\mathbf{B}\,\mathcal{L}_h\big]^{\perp}.$$

Indeed

$$\mathbf{B}'\,\boldsymbol{\sigma}_h \in \mathcal{L}_h^{\circ} \iff \langle \mathbf{B}'\boldsymbol{\sigma}_h, \mathbf{v}_h \rangle = (\!(\, \boldsymbol{\sigma}_h, \mathbf{B}\mathbf{v}_h \,)\!) = 0 \quad \forall\, \mathbf{v}_h \in \mathcal{L}_h.$$

- The compatibility variational condition of a tangent deformation field $\boldsymbol{\varepsilon} \in \mathcal{H}$ on the continuous structure $\mathcal{M}(\boldsymbol{\Omega}, \mathcal{L}, \mathbf{B})$ is given by

  $$\langle \boldsymbol{\sigma}, \boldsymbol{\varepsilon} \rangle = 0\,, \quad \forall\, \boldsymbol{\sigma} \in \mathbf{ker}(\mathbf{B}'_{\mathcal{L}}) = (\mathbf{B}\mathcal{L})^{\perp}$$

  and it is equivalent to $\boldsymbol{\varepsilon} \in \mathbf{B}\mathcal{L}$.

- The compatibility variational condition on the discrete structure is

  $$\langle \boldsymbol{\sigma}_h, \boldsymbol{\varepsilon} \rangle = 0\,, \quad \forall\, \boldsymbol{\sigma}_h \in \mathbf{ker}(\mathbf{B}'_h) = \mathcal{S}_h \cap \big[\mathbf{B}\,\mathcal{L}_h\big]^{\perp}\,,$$

  that is equivalent to

  $$\boldsymbol{\varepsilon} \in \big[\mathcal{S}_h \cap (\mathbf{B}\,\mathcal{L}_h)^{\perp}\big]^{\perp} = \mathcal{S}_h^{\perp} + \mathbf{B}\,\mathcal{L}_h\,.$$

  The last equality holds since $\mathbf{B}\,\mathcal{L}_h$ is closed in $\mathcal{H}$ and the sum subspace $\mathcal{S}_h^{\perp} + \mathbf{B}\,\mathcal{L}_h$ is closed in $\mathcal{H}$ given that $\mathcal{S}_h^{\perp}$ is finite codimensional and $\mathbf{B}\,\mathcal{L}_h$ is finite dimensional (see section I.11 (p.81) in [159]).

Tangent deformation fields $\boldsymbol{\varepsilon} \in \mathcal{S}_h^{\perp} \subseteq \mathcal{H}$ are dubbed *free tangent deformations.* The compatibility property on the discrete structure follows by the compatibility property on the continuous structure if we have that

$$\mathbf{ker}(\mathbf{B}'_h) = \mathbf{ker}(\mathbf{B}'_{\mathcal{L}}) \cap \mathcal{S}_h\,.$$

The following property is perfectly analogous to the one in proposition 9.1.1.

Proposition 9.1.2 *The equality* $\boldsymbol{ker}(\mathbf{B}'_h) = \boldsymbol{ker}(\mathbf{B}'_{\mathcal{L}}) \cap \mathcal{S}_h$ *is equivalent to the property*

$$ii) \qquad \forall\, \mathbf{u} \in \mathcal{H} \quad \exists\, \mathbb{P}_h \mathbf{u} \in \mathcal{L}_h \;:\; (\!(\, \boldsymbol{\sigma}_h \,,\, \mathbf{B}\,(\mathbf{u} - \mathbb{P}_h \mathbf{u})\,)\!) = 0\,, \quad \forall\, \boldsymbol{\sigma}_h \in \mathcal{S}_h\,.$$

Proof. Observing that

$$\mathbf{ker}(\mathbf{B}'_h) = \mathbf{ker}(\mathbf{B}'_{\mathcal{L}}) \cap \mathcal{S}_h \iff \big[\mathbf{ker}(\mathbf{B}'_h)\big]^{\perp} = \big[\mathbf{ker}(\mathbf{B}'_{\mathcal{L}}) \cap \mathcal{S}_h\big]^{\perp}\,,$$

and being

$$\big[\mathbf{ker}(\mathbf{B}'_{\mathcal{L}}) \cap \mathcal{S}_h\big]^{\perp} = \mathcal{S}_h^{\perp} + \mathbf{B}\,\mathcal{L}\,,$$

$$\big[\mathbf{ker}(\mathbf{B}'_h)\big]^{\perp} = \mathcal{S}_h^{\perp} + \mathbf{B}\,\mathcal{L}_h\,,$$

we get

$$\boldsymbol{\varepsilon} \in \big[\mathbf{ker}(\mathbf{B}'_{\mathcal{L}}) \cap \mathcal{S}_h\big]^{\perp} \iff \exists\, \mathbf{u} \in \mathcal{L} \;:\; (\!(\, \boldsymbol{\sigma}_h \,,\, \boldsymbol{\varepsilon}\,)\!) = (\!(\, \boldsymbol{\sigma}_h \,,\, \mathbf{B}\mathbf{u}\,)\!) \quad \forall\, \boldsymbol{\sigma}_h \in \mathcal{S}_h\,,$$

$$\boldsymbol{\varepsilon} \in \big[\mathbf{ker}(\mathbf{B}'_h)\big]^{\perp} \iff \exists\, \mathbf{u}_h \in \mathcal{L}_h \;:\; (\!(\, \boldsymbol{\sigma}_h \,,\, \boldsymbol{\varepsilon}\,)\!) = (\!(\, \boldsymbol{\sigma}_h \,,\, \mathbf{B}\mathbf{u}_h\,)\!) \quad \forall\, \boldsymbol{\sigma}_h \in \mathcal{S}_h\,.$$

Hence the $ii)$ holds. The converse implication is easily verifiable.

9.2 Discrete primary mixed elastic problem

Let us consider a *discrete primary mixed elastic problem*

$$\begin{cases} \mathbf{k}\,(\mathbf{u}_h, \mathbf{v}_h) + \mathbf{b}\,(\boldsymbol{\sigma}_h, \mathbf{v}_h) = \langle \ell_{eq}, \mathbf{v}_h \rangle\,, & \begin{cases} \mathbf{u}_h \in \mathcal{L}_h\,, \\ \forall\, \mathbf{v}_h \in \mathcal{L}_h\,, \end{cases} \\ \mathbf{b}\,(\boldsymbol{\tau}_h, \mathbf{u}_h) - \mathbf{c}_o\,(\boldsymbol{\sigma}_h, \boldsymbol{\tau}_h) = (\!(\, \boldsymbol{\delta}_{eq}\,, \boldsymbol{\tau}_h \,)\!)\,, & \begin{cases} \boldsymbol{\sigma}_h \in \mathcal{S}_h\,, \\ \forall\, \boldsymbol{\tau}_h \in \mathcal{S}_h\,. \end{cases} \end{cases}$$

The problem $\mathbb{M}_h$ is always well-posedness since the playing spaces are of finite dimension. Let us consider the discrete operators:

- $\mathbf{K}_h \in BL\,(\mathcal{L}_h\,; \mathcal{F}/\mathcal{L}_h^{\circ})$ defined by

$$\mathbf{K}_h\,\mathbf{u}_h := \mathbf{K}\,\mathbf{u}_h + \mathcal{L}_h^{\circ}\,,$$

- $\mathbf{C}_{oh} \in BL\,(\mathcal{S}_h\,; \mathcal{H}/\mathcal{S}_h^{\perp})$ defined by

$$\mathbf{C}_{oh}\,\boldsymbol{\sigma}_h := \mathbf{C}_o\,\boldsymbol{\sigma}_h + \mathcal{S}_h^{\perp}\,.$$

The problem $\mathbb{M}_h$ admits an unique solution if

$$\begin{cases} \mathbf{ker}(\mathbf{K}_h) \cap \mathbf{ker}(\mathbf{B}_h) = \{\mathbf{o}\}\,, \\ \mathbf{ker}(\mathbf{B}'_h) \cap \mathbf{ker}(\mathbf{C}_{oh}) = \{\mathbf{o}\}\,. \end{cases}$$

- The explicit expression of the uniqueness condition of the state of discrete stress $\boldsymbol{\sigma}_h \in \mathcal{S}_h$

$$\mathbf{ker}(\mathbf{B}'_h) \cap \mathbf{ker}(\mathbf{C}_{oh}) = \mathbf{ker}(\mathbf{B}'_h) \cap \mathcal{S}_h \cap {\mathbf{C}_o}^{-1}\,\mathcal{S}_h^{\perp} = \{\mathbf{o}\}\,,$$

  shows that the condition is verified if $\mathbf{ker}(\mathbf{C}_o) = \{\mathbf{o}\}$.
  Indeed by the positivity of $\mathbf{C}_o \in BL\,(\mathcal{H}\,; \mathcal{H})$ we deduce that

$$\boldsymbol{\sigma}_h \in \mathcal{S}_h \cap {\mathbf{C}_o}^{-1}\mathcal{S}_h^{\perp} \implies \mathbf{C}_o\boldsymbol{\sigma}_h = \mathbf{o} \implies \boldsymbol{\sigma}_h = \mathbf{o}\,,$$

  i.e. $\mathbf{ker}(\mathbf{C}_{oh}) = \mathcal{S}_h \cap {\mathbf{C}_o}^{-1}\mathcal{S}_h^{\perp} = \{\mathbf{o}\}$.

- The explicit expression of the uniqueness condition for the displacement field $\mathbf{u}_h \in \mathcal{L}_h$ is

$$\mathbf{ker}(\mathbf{K}_h) \cap \mathbf{ker}(\mathbf{B}_h) = \mathcal{L}_h \cap \mathbf{K}^{-1}\mathcal{L}_h^{\circ} \cap \mathbf{B}^{-1}\mathcal{S}_h^{\perp} = \{\mathbf{o}\}\,.$$

By the positivity of $\mathbf{K}_h \in BL(\mathcal{L}_h\,;\mathcal{F}/\mathcal{L}_h^{\circ})$ we deduce that

$$\begin{cases}\mathbf{u}_h \in \mathcal{L}_h\,,\\ \mathbf{K}\,\mathbf{u}_h \in \mathcal{L}_h^{\circ}\,,\end{cases} \Longrightarrow \langle \mathbf{K}\,\mathbf{u}_h, \mathbf{u}_h \rangle = 0 \Longrightarrow \mathbf{K}\,\mathbf{u}_h = \{\mathbf{o}\}\,,$$

i.e.

$$\mathbf{ker}(\mathbf{K}_h) = \mathcal{L}_h \cap \mathbf{ker}(\mathbf{K})\,.$$

The uniqueness of the displacement field $\mathbf{u}_h \in \mathcal{L}_h$ holds if and only if

$$\begin{cases}\mathbf{B}\mathbf{u}_h \in \mathcal{S}_h^{\perp}\,,\\ \mathbf{u}_h \in \mathcal{L}_h \cap \mathbf{ker}(\mathbf{K})\,,\end{cases} \Longrightarrow \mathbf{u}_h = \mathbf{o}\,,$$

namely if

$$\mathcal{L}_h \cap (\mathbf{ker}(\mathbf{K})) \cap \mathbf{B}^{-1}\mathcal{S}_h^{\perp} = \{\mathbf{o}\}\,.$$

In mechanical terms we say that

- there are not elastically ineffective conforming discrete displacements that are not zero and generating free tangent deformations.

Remark 9.2.1 *The condition* $\boldsymbol{ker}(\mathbf{B}_h) = \mathbf{o}$ *imposes that* $\mathbf{B}\mathcal{L}_h \cap \mathcal{S}_h^{\perp} = \mathbf{o}$ *and thus, by observation 9.1.1, we get* $\boldsymbol{ker}(\mathbf{B}_h) = \boldsymbol{ker}(\mathbf{B}_{\mathcal{L}}) \cap \mathcal{L}_h\,.$ *In structural applications the condition* $\boldsymbol{ker}(\mathbf{B}_h) = \mathbf{o}$ *is the needful (and sufficient) to assure the uniqueness of the displacement field in absence of elastic constraints. It follows that the interpolating subspaces* $\mathcal{L}_h$ *and* $\mathcal{S}_h$ *have to be assumed so that the condition* $\mathbf{B}\mathcal{L}_h \cap \mathcal{S}_h^{\perp} = \mathbf{o}$ *is respected. In the finite element method the condition have to be substituted by a stronger one so that by imposing it to shape functions defined in the reference element.*

9.2.1 Error estimate

Error estimates in mixed elastostatics is a topic of great interest in computational mechanics. An assessment of the approximation energy error is provided in terms of a parameter h which is the elements' diameter in the finite element method. A sufficient condition for the convergence in energy of the approximate solution is expressed in terms of suitable properties of the interpolating subspaces. The result contributes an alternative form of the well known LBB condition. Let us assume that the uniqueness and well-posedness conditions of the continuous problem and the uniqueness condition of the displacement field of the discrete problem are fulfilled. We will provide an estimate of the *approximation error* in energy, defined by

$$\|\mathbf{u} - \mathbf{u}_h\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{H}} \,.$$

Following the treatment developed in [20], we employ the triangle inequality to conclude that

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}_h\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{H}} \ &\leq \|\mathbf{u} - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}} + \\ &+ \|\mathbf{u}_h - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \|\boldsymbol{\sigma}_h - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}} \,. \end{aligned}$$

$\forall\, \overline{\mathbf{u}_h} \in \mathcal{L}_h \,, \forall\, \overline{\boldsymbol{\sigma}_h} \in \mathcal{S}_h$. The first step consists in increasing the term $\|\mathbf{u}_h - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \|\boldsymbol{\sigma}_h - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}}$ by means of the distance $\|\mathbf{u} - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}}$. To this end we observe that by the problems $\mathbb{M}$ and $\mathbb{M}_h$ follows that

$$\mathbb{P}) \begin{cases} \mathbf{k}\,(\mathbf{u}_h - \overline{\mathbf{u}_h}, \mathbf{v}_h) + \mathbf{b}\,(\boldsymbol{\sigma}_h - \overline{\boldsymbol{\sigma}_h}, \mathbf{v}_h) = \mathbf{k}\,(\mathbf{u} - \overline{\mathbf{u}_h}, \mathbf{v}_h) + \mathbf{b}\,(\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}_h}, \mathbf{v}_h)\,, \\ \mathbf{b}\,(\boldsymbol{\tau}_h, \mathbf{u}_h - \overline{\mathbf{u}_h}) - \mathbf{c}_o\,(\boldsymbol{\sigma}_h - \overline{\boldsymbol{\sigma}_h}, \boldsymbol{\tau}_h) = \mathbf{b}\,(\boldsymbol{\tau}_h, \mathbf{u} - \overline{\mathbf{u}_h}) - \mathbf{c}_o\,(\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}_h}, \boldsymbol{\tau}_h)\,. \end{cases}$$

The known terms are continuous linear functionals on $\mathcal{L}_h$ and on $\mathcal{S}_h$.

- Applying to the problem $\mathbb{P})$ the treatment of subsection **??** of chapter **??** we deduce that the estimate holds

  $$\|\mathbf{u}_h - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \|\boldsymbol{\sigma}_h - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}} \ \leq \ m_h \left[\|\mathbf{u} - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}} \right].$$

  m_h is a positive and bounded nonlinear function of

  $$\|\mathbf{c}_o\|\,, \quad \|\mathbf{k}\|\,, \quad c_{\mathbf{B}h}\,, \quad c_{\mathbf{k}h}\,, \quad \alpha_h\,,$$

  on bounded subsets.

  By the triangle inequality we deduce that

  $$\|\mathbf{u} - \mathbf{u}_h\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{H}} \ \leq \ (1 + m_h) \left[\|\mathbf{u} - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}} \right],$$

  $\forall\, \overline{\mathbf{u}_h} \in \mathcal{L}_h \,, \forall\, \overline{\boldsymbol{\sigma}_h} \in \mathcal{S}_h$. Setting $c_h = 1 + m_h$ we conclude that

  $$\|\mathbf{u} - \mathbf{u}_h\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\mathcal{H}} \leq c_h \left[\inf_{\overline{\mathbf{u}_h} \in \mathcal{L}_h} \|\mathbf{u} - \overline{\mathbf{u}_h}\|_{\mathcal{L}} + \inf_{\overline{\boldsymbol{\sigma}_h} \in \mathcal{S}_h} \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}_h}\|_{\mathcal{H}} \right].$$

9.2.2 LBB condition and convergence

If the constant c is independent of h, the convergence in energy of the approximate solution to the exact one is ensured if there are sufficient properties of interpolation of the discrete subspaces. In the literature a condition which guarantees such properties is referred to as Ladyzhenskaya-Babuška-Brezzi condition (LBB *condition*, see [88], [10], [11], [18], [136], [20]). An alternative form of LBB condition is provided in the next theorem.

Theorem 9.2.1 *Let the mixed elastic problem be well-posed with an unique solution and the elasticity of the structure be not singular so that* $\boldsymbol{ker}(\mathbf{C}_o) = \{\mathbf{o}\}$ *with the kinematic operator a* Korn*'s operator. Further, let us assume that the families of the interpolating linear subspaces* $\mathcal{L}_h \subset \mathcal{L}$ *and* $\mathcal{S}_h \subset \mathcal{H}$ *meet the conditions*

$$a) \quad \mathbf{B}\,\mathcal{L}_h \cap \overset{\circ}{\mathcal{S}_h} = \{\mathbf{o}\}\,,$$

$$b) \quad \mathbf{B}\,\mathcal{L}_h + \overset{\circ}{\mathcal{S}_h} \quad \textit{uniformly closed in} \quad \mathcal{H}\,.$$

Then an asymptotic estimate of the approximation error can be inferred from an asymptotic estimate of the interpolation error.

Proof. Let us preliminarily observe that the condition $a)$ is equivalent to $\mathbf{ker}(\mathbf{B}_h) = \mathbf{ker}(\mathbf{B}) \cap \mathcal{L}_h$. The uniqueness of the displacement solution of the continuous problem, given that $\mathbf{ker}(\mathbf{K}) \cap \mathbf{ker}(\mathbf{B}) = \{\mathbf{o}\}$, implies $\mathbf{ker}(\mathbf{K}_h) \cap \mathbf{ker}(\mathbf{B}_h) = \mathbf{ker}(\mathbf{K}) \cap \mathbf{ker}(\mathbf{B}) \cap \mathcal{L}_h = \{\mathbf{o}\}$. Hence the uniqueness of solution of the discrete problem $\mathbb{MV}_h)$ in terms of interpolating displacement fields is met. The ellipticity condition on $\mathbf{ker}(\mathbf{B})$ of the bilinear form $\mathbf{k}$

$$\mathbf{k}\,(\mathbf{u},\mathbf{u}) \geq\; c_{\mathbf{k}}\,\|\mathbf{u}\|^2_{\mathcal{L}/(\mathbf{Ker\,K}\cap\mathbf{Ker\,B})}\,, \qquad \forall\,\mathbf{u} \in \mathbf{ker}(\mathbf{B})\,,$$

can be rewritten as $\mathbf{k}\,(\mathbf{u},\mathbf{u}) \geq\; c_{\mathbf{k}}\,\|\mathbf{u}\|^2_{\mathcal{L}}$ for any $\mathbf{u} \in \mathbf{ker}(\mathbf{B})$, so that:

$$\mathbf{k}\,(\mathbf{u}_h,\mathbf{u}_h) \geq\; c_{\mathbf{k}}\,\|\mathbf{u}_h\|^2_{\mathcal{L}} \quad \forall\,\mathbf{u}_h \in \mathbf{ker}(\mathbf{B}_h) = \mathbf{ker}(\mathbf{B}) \cap \mathcal{L}_h\,,$$

i.e. the uniform ellipticity on $\mathbf{ker}(\mathbf{B}_h)$ of the bilinear form $\mathbf{k}$. The condition $b)$ is equivalent to the uniform closure of the family of subspaces $\operatorname{Im}\mathbf{B}_h = \mathbf{B}\mathcal{L}_h + \overset{\circ}{\mathcal{S}_h}$ which is expressed by the inequality

$$\sup_{\boldsymbol{\tau}_h \in \mathcal{S}_h} \frac{(\!(\,\boldsymbol{\tau}_h\,,\,\mathbf{B}\mathbf{u}_h\,)\!)}{\|\boldsymbol{\tau}_h\|_{\mathcal{H}}} \geq c_{\mathbf{b}}\,\|\mathbf{u}_h\|_{\mathcal{H}/\mathrm{Ker}\,\mathbf{B}_h} \quad \forall\,\mathbf{u}_h \in \mathcal{L}_h\,,$$

with $c_{\mathbf{b}}$ independent of h. Then the inequality above together with the problem $\mathbb{P})$ allows us to state that

$$\|\mathbf{u}_h - \overline{\mathbf{u}}_h\|_{\mathcal{L}} + \|\boldsymbol{\sigma}_h - \overline{\boldsymbol{\sigma}}_h\|_{\mathcal{H}} \leq m\,(\|\mathbf{u} - \overline{\mathbf{u}}_h\|_{\mathcal{L}} + \|\boldsymbol{\sigma} - \overline{\boldsymbol{\sigma}}_h\|_{\mathcal{H}})\,,$$

where m is a nonlinear function of $\|\mathbf{c}_o\|$, $\|\mathbf{k}\|$, γ, $c_{\mathbf{k}}$, α, and is positive and bounded on bounded subsets [159]. By the triangle inequality we deduce that

$$\|\mathbf{u}-\mathbf{u}_h\|_{\mathcal{L}}+\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathcal{H}} \leq (1+m)\left[\|\mathbf{u}-\overline{\mathbf{u}}_h\|_{\mathcal{L}}+\|\boldsymbol{\sigma}-\overline{\boldsymbol{\sigma}}_h\|_{\mathcal{H}}\right],$$

for any $\overline{\mathbf{u}}_h \in \mathcal{L}_h$ and $\overline{\boldsymbol{\sigma}}_h \in \mathcal{S}_h$. Setting $c = 1+m$ we conclude that

$$\|\mathbf{u}-\mathbf{u}_h\|_{\mathcal{L}}+\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathcal{H}} \leq c\left[\inf_{\overline{\mathbf{u}}_h\in\mathcal{L}_h}\|\mathbf{u}-\overline{\mathbf{u}}_h\|_{\mathcal{L}}+\inf_{\overline{\boldsymbol{\sigma}}_h\in\mathcal{S}_h}\|\boldsymbol{\sigma}-\overline{\boldsymbol{\sigma}}_h\|_{\mathcal{H}}\right]. \qquad \blacksquare$$

Remark 9.2.2 *Observing that* $\mathbf{B}\,\mathcal{L}_h \cap \mathcal{S}_h^{\circ} = \{\mathbf{o}\}$, *the uniform closure condition in* $\mathcal{H}$ *of the family* $\mathbf{B}\,\mathcal{L}_h + \mathcal{S}_h^{\circ}$ *can be expressed by* $\|\mathbf{\Pi}\mathbf{B}\mathbf{u}_h\|_{\mathcal{H}} \geq c\,\|\mathbf{B}\mathbf{u}_h\|_{\mathcal{H}}$ *for any* $\mathbf{u}_h \in \mathcal{L}_h$, *in which* $\mathbf{\Pi} \in BL(\mathcal{H};\mathcal{H})$ *is the orthogonal projector on* $\mathcal{S}_h \subset \mathcal{H}$ *[159]. Hence this condition is an alternative expression of the LBB condition.* $\blacksquare$

Theorem 9.2.1 shows that the approximation error is bounded above by the interpolation error. The asymptotic estimate, i.e. as $h \to 0$, of the decrease rate of the interpolation error

$$\mathbf{Err}(h) = \inf_{\overline{\mathbf{u}}_h\in\mathcal{L}_h}\|\mathbf{u}-\overline{\mathbf{u}}_h\|_{\mathcal{L}}+\inf_{\overline{\boldsymbol{\sigma}}_h\in\mathcal{S}_h}\|\boldsymbol{\sigma}-\overline{\boldsymbol{\sigma}}_h\|_{\mathcal{H}}$$

is provided by the polynomial interpolation theory that leads to the exponential formula:

$$\mathbf{Err}(h) \leq \beta\, h^k\,(\|\mathbf{u}\|_{\mathcal{L}}+\|\boldsymbol{\sigma}\|_{\mathcal{H}})\,.$$

In a two-logaritmic scale the exponential law with exponent k transforms to the linear law with slope equal to k that is

$$\ln\left(\mathbf{Err}(h)\right) \leq \ln(\beta\,(\|\mathbf{u}\|_{\mathcal{L}}+\|\boldsymbol{\sigma}\|_{\mathcal{H}}))+k\,\ln(h)\,.$$

Chapter 10

Subdifferential Calculus

10.1 Introduction

Subdifferential calculus is nowadays a well developed chapter of non-smooth analysis which is recognized for its many applications to optimization theory. The very definition of subdifferential and the basic results concerning the addition and the chain rule of subdifferential calculus were first established in the early sixties by ROCKAFELLAR [193] with reference to convex functions on $\Re^n$. A comprehensive treatment of the subject has been provided by himself in the later book on convex analysis [194]. The theory was developed further by MOREAU [195] in the context of linear topological vector spaces and applied to problems of unilateral mechanics [196]. A summary of basic mathematical results can also be found in the book by LAURENT [197] and in an introductory chapter of the book by EKELAND and TEMAM [198]. In the early seventies different attempts were initiated to extend the range of validity of subdifferential calculus to non-convex functions, mainly by ROCKAFELLAR and his school. In this context saddle functions were considered by MCLINDEN [199]. Significant advances were made by CLARKE [200, 201] who set up a definition of subdifferential for arbitrary lower semicontinuous functions on $\Re^n$ and extended the validity of the rules of subdifferential calculus to this non-convex context. His results were later further developed and extended by ROCKAFELLAR [202, 203], who has also provided a nice exposition of the state of art, up to the beginning of eighties, in [204]. A review of the main results and applications in different areas of mathematical physics can be found in a recent book by PANAGIOTOPOULOS

[205]. A different treatment of the subject is presented in the book by Ioffe and Tihomirov [206], who introduce the notion of regular local convexity to deal with the non-convex case. A careful review of all these contributions to subdifferential calculus leads however to the following considerations. The results provided up to now to establish the validity of the addition and of the chain rule for subdifferentials do appear to rely upon sufficient but largely not necessary assumptions. In fact a number of important situations, in which the results do hold true, are beyond of the target of existing theorems. On the other hand the author has realized the lack of a chain rule concerning the very important case of convex functionals which are expressed as the composition of a monotone convex function and another convex functional. The first observation in this respect was made with reference to positively homogeneous convex functionals of order greater than one or, more generally, to convex functionals which are composed by a Young function and a sublinear functional (gauge-like functionals in Rockafellar's terminology). The theorems presented in this paper are intended to contribute to the filling of these gaps; progress is provided in two directions. The first concerns the chain rule pertaining to the composition of a convex functional and a differentiable operator. We have addressed the question of finding out a necessary and sufficient condition for its validity. The theorem provided here shows that this task can be accomplished to within a closure operation; the proof is straightforward and relies on a well-known lemma of convex analysis concerning sublinear functionals. The result obtained must be considered as optimal; a simple counterexample reveals indeed that there is no hope of dropping the closure operation. On the contrary, to establish a perfect equality (one not requiring closures) in the chain rule formula, classical treatments were compelled to set undue restrictions on the range of validity of the result. In this respect it has to be remarked that classical conditions were *global* in character, in the sense that validity of chain and addition rules where ensured at all points. The new results provided here are instead based upon *local* conditions which imply validity of the rules only at the very point where subdifferentials have to be evaluated. It follows that classical conditions can be verified *a priori* while the new conditions must be checked *a posteriori* at the point of interest. The second contribution consists in establishing a new chain rule formula concerning functionals which are formed by the composition of a monotone convex function and a convex functional. A natural application of these results can be exploited in convex optimization problems. It is shown in fact that the Kuhn and Tucker multipliers theory can be immediately derived from the above theorems and the existence proof can be performed under assumptions less stringent then the classical Slater conditions [207]. Computation of the subdifferentials involved

in the proof requires considering the following two special cases of the new chain rules of subdifferential calculus contributed here. In the first case we have to deal with the composition of the indicator of the zero and of an affine functional. In the second one we must consider a functional formed by the composition of the indicator of non-positive reals and of a convex functional. Both cases were not covered by previous results.

10.2 Local convexity and subdifferentials

Let (X, X') be a pair of locally convex topological vector spaces (l.c.t.v.s.) placed in separating duality by a bilinear form $\langle \cdot, \cdot \rangle$ and $f : X \mapsto \Re \cup \{+\infty\}$ an extended real valued functional with a nonempty effective domain:

$$\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\} .$$

The one-sided directional derivative of f at the point $x \in \operatorname{dom} f$, along the vector $h \in X$, is defined by the limit

$$df(x; h) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \Big[f(x + \varepsilon h) - f(x) \Big] .$$

The derivative of f at x is then the extended real valued functional $p : X \mapsto \{-\infty\} \cup \Re \cup \{+\infty\}$ defined by

$$p(h) := df(x; h) ,$$

which is easily seen to be positively homogeneous in h. The functional f is said to be *locally convex* at x when p is sublinear in h, that is:

$$\begin{cases} p(\alpha h) = \alpha p(h) , & \forall\, \alpha \geq 0 \qquad \text{(positive homogeneity)} , \\ p(h_1) + p(h_2) \geq p(h_1 + h_2) , & \forall\ h_1, h_2 \in X \qquad \text{(subadditivity)} . \end{cases}$$

The epigraph of p is then a convex cone in $X \times \Re$. A locally convex functional f is said to be *locally subdifferentiable* at x if its one-sided derivative p is a proper sublinear functional, i.e., if it is nowhere $-\infty$. In fact, denoting by $\overline{p}$ the closure of p defined by the limit formula:

$$\overline{p}(h) = \liminf_{z \to h} p(z) , \quad \forall\, h \in X ,$$

a well-known result of convex analysis ensures that the proper lower semicontinuous (l.s.c.) sublinear functional $\overline{p}$, turns out to be the support functional of a nonempty closed convex set K^*, that is:

$$\overline{p}(h) = \sup\{\langle x^*, h\rangle \ : \ x^* \in K^*\},$$

with

$$K^* = \{x^* \in X' \ : \ p(h) \geq \langle x^*, h\rangle , \quad \forall\, h \in X\}.$$

The *local subdifferential* of the functional f is then defined by:

$$\partial f(x) := K^* .$$

A relevant special case, which will be referred to in the sequel, occurs when, the one-sided derivative of f at x, turns out to be l.s.c. so that $p = \overline{p}$. The functional f is then said to be *regularly locally subdifferentiable* at x. When the functional f is differentiable at $x \in X$ the local subdifferential is a singleton and coincides with the usual differential. For a convex functional $f : X \mapsto \Re \cup \{+\infty\}$, the difference quotient in the definition of one-sided directional derivative does not increase as ε decreases to zero [194, 206]. Hence the limit exists at every point $x \in \operatorname{dom} f$ along any direction $h \in X$ and the following formula holds:

$$df(x; h) = \inf_{\varepsilon > o} \frac{1}{\varepsilon}\Big[f(x + \varepsilon h) - f(x)\Big].$$

A simple computation shows that the directional derivative of f is convex as a function of h and hence sublinear. Moreover the definition of local subdifferential turns out to be equivalent to

$$x^* \in \partial f(x) \iff f(y) - f(x) \geq \langle x^*, y - x\rangle \qquad \forall\, y \in X,$$

which is the usual definition of subdifferential in convex analysis [194].

10.3 Classical Subdifferential Calculus

Let $f_1, f_2 : X \mapsto \Re \cup \{+\infty\}$ and $f : Y \mapsto \Re \cup \{+\infty\}$ be convex functionals and $L : X \mapsto Y$ a continuous linear operator. From the definition of local subdifferential it follows easily that:

$$\partial(\lambda f)(x) = \lambda \partial f(x), \quad \lambda \geq 0,$$
$$\partial(f_1 + f_2)(x) \supseteq \partial f_1(x) + \partial f_2(x),$$
$$\partial(f \circ L)(x) \supseteq L'\, \partial f(Lx),$$

where L' denotes the dual of L.

As remarked in [198] equality in the last two relations is far from being always realized. The aim of subdifferential calculus has thus primarily consisted in providing conditions sufficient to ensure that the converse of the last two inclusions does hold true. In convex analysis this task has been classically accomplished by the following kind of results [194, 197, 198, 206].

Theorem 10.3.1 (Additivity) *If $f_1, f_2 : X \mapsto \Re \cup \{+\infty\}$ are convex and at least one of them is continuous at a point of $dom\, f_1 \cap dom\, f_2$, then*

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \quad \forall\, x \in X\,.$$

■

Theorem 10.3.2 (Chain-rule) *Given a continuous linear operator $L : X \mapsto Y$ and a convex functional $f : Y \mapsto \Re \cup \{+\infty\}$ which is continuous at a point of $dom\, f \cap Im\, L$, it results that*

$$\partial(f \circ L)(x) = L' \, \partial f(\,Lx\,)\,, \quad \forall\, x \in X\,.$$

■

The chain-rule equality above can be equivalently written with the more familiar notation

$$\partial(f \circ L)(x) = \partial f(\,Lx\,) \circ L\,, \quad \forall\, x \in X\,.$$

A generalization of the previous results can be performed to get a chain rule involving a locally convex functional and a nonlinear differentiable operator. Given a nonlinear differentiable operator $A : X \mapsto Y$ and a functional $f : Y \mapsto \Re \cup \{+\infty\}$ which is locally convex at $y_o = A(x_o)$, we have to prove the following equality:

$$\partial(f \circ A)(x_o) = \partial f[A(x_o)] \circ dA(x_o) = [dA(x_o)]' \, \partial f[A(x_o)]\,,$$

where $dA(x_o)$ is the derivative of the operator A at $x_o \in X$.

The task can be accomplished by first providing conditions sufficient to guarantee the validity of the chain-rule identity for one-sided directional derivatives:

$$d(f \circ A)(x_o\,; x) = df[\,A(x_o)\,; dA(x_o)x\,]\,, \quad \forall\, x \in X\,,$$

which is easily seen to hold trivially when A is an affine operator. Then, setting $L := dA(x_o)$, we consider the sublinear functionals

$$p(y) := df(A(x_o); y) \quad \text{and} \quad q(x) = d(f \circ A)(x_o\,; x)\,.$$

The identity above ensures that $q = p \circ L$; further, observing that, by definition,

$$\partial p(0) = \partial f(A(x)) \quad \text{and} \quad \partial q(0) = \partial(p \circ L)(0) = \partial(f \circ A)(x)\,,$$

the equality to be proved can then be rewritten as:

$$\partial(p \circ L)(0) = L' \partial p(0)\,.$$

This result can be inferred from the chain-rule theorem concerning convex functionals by assuming that the sublinear functional p is continuous at a point of $\operatorname{dom} p \cap \operatorname{Im} L$. In this respect we remark that it has been shown in [206] that, assuming the functional f to be *regularly locally convex* at $A(x) \in Y$, that is locally convex and uniformly differentiable in all directions at $A(x)$, its derivative p turns out to be continuous in the whole space Y. Therein it is also proved that a convex functional is regularly locally convex at a point if and only if it is continuous at that point. An analogous generalization can be performed for the addition formula of subdifferential calculus. A different and more general treatment of the nonconvex case has been developed, on the basis of Clarke's [200, 201] contributions, by Rockafellar [202, 203]. According to his approach the validity of the chain rule was proved by assuming that the operator A is strictly differentiable at $x \in X$, that f is finite, directionally Lipschitzian and subdifferentially regular at $A(x)$ and that the interior of the domain of the one-sided derivative of f at $x \in X$ has a non-empty intersection with the range of $dA(x)$. Reference is made to the quoted papers for a precise assessment of definitions and proofs.

10.4 New results

As illustrated above, all the contributions provided to subdifferential calculus until now have directed their efforts in the direction of finding conditions directly sufficient to ensure the validity of the equality sign in the relevant relations. This approach has led to the formulation of very stringent conditions which rule out a number of significant situations. In the next subsection we propose an alternative approach to the assessment of the chain rule pertaining to the composition of a convex functional and a differentiable operator. Further we derive the addition rule as a special case of this chain rule. In the second subsection we present the proof of a new product-rule formula of subdifferential calculus which deals with the composition of a monotone convex function and a convex functional. These results are applied in the last subsection to assess the existence of KUHN and TUCKER multipliers in convex optimization problems, under assumption less stringent than the classical SLATER conditions [207] (see also [194] and [204]).

10.4.1 Classical addition and chain rule formulas

The new approach to classical rules of subdifferential calculus consists in splitting the procedure into two steps. It has in fact been realized that getting the equality at once in the related relations requires too stringent assumptions and allows less deep insight into the problem.

The classical chain rule requires the equality of the subdifferential of a composite function, which is a closed convex set, to the image of the subdifferential of a convex function through a linear operator. Since in general the image of a closed convex set fails to be closed too, it is natural to look first for conditions apt to provide equality of the former set to the closure of the latter one, leaving to a subsequent step the answer about the closedness of the latter set.

The first step is performed by means of the following result.

Theorem 10.4.1 (New proof of the classical chain rule) *Let $A : X \mapsto Y$ be a nonlinear operator which is differentiable at a point $x_o \in X$ with derivative $dA(x_o) : X \mapsto Y$ linear and continuous. Let further $f : Y \mapsto \Re \cup \{+\infty\}$ be a functional which is locally subdifferentiable at $A(x_o) \in Y$ and assume that $f \circ A : X \mapsto \Re \cup \{+\infty\}$ is locally subdifferentiable at $x_o \in X$. Then we have that:*

$$\partial(f \circ A)(x_o) = \overline{\partial f[A(x_o)] \circ [dA(x_o)]} = \overline{[dA(x_o)]' \partial f[A(x_o)]}$$

if and only if

$$\overline{q}(x) = \overline{p}(Lx) \,, \quad \forall\, x \in X \,,$$

where $q(\cdot) := d(f \circ A)(x_o; \cdot)$, $p(\cdot) := df[A(x_o); \cdot]$ and $L := dA(x_o)$, a superimposed bar denoting the closure.

Proof. f being locally subdifferentiable at $A(x_o) \in Y$, its directional derivative $p : Y \mapsto \Re \cup \{+\infty\}$ is a proper sublinear functional, so that:

$$\overline{p}(y) = \sup\{\langle y^*, y \rangle \mid y^* \in K^*\}$$

where:

$$K^* = \partial p(0) := \{y^* \in Y' \mid p(y) \geq \langle y^*, y \rangle \quad \forall\, y \in Y\}$$

is a nonempty, closed convex set. Then we have:

$$\overline{p}(Lx) = \sup\{\langle y^*, Lx \rangle \mid y^* \in K^*\} = \sup\{\langle x^*, x \rangle \mid x^* \in L'K^*\}\,.$$

Similarly, $f \circ A$ being locally subdifferentiable at $x_o \in X$ its directional derivative $q : X \mapsto \Re \cup \{+\infty\}$ is a proper sublinear functional, so that

$$\overline{q}(x) = \sup\{\langle x^*, x \rangle \mid x^* \in C^*\}\,,$$

where

$$C^* = \{x^* \in X' \mid q(x) \geq \langle x^*, x \rangle\,, \quad \forall\, x \in X\}$$

is a nonempty, closed convex set in X'.

Comparison of the two expressions above leads directly to the following conclusion:

$$\overline{q}(x) = \overline{p}(Lx) \quad \text{if and only if} \quad C^* = \overline{L'K^*}\,.$$

The statement of the theorem is then inferred by observing that:

$$\begin{aligned} \partial f[A(x_o)] &= \{y^* \in Y' \mid df[(A)(x_o); y] \geq \langle y^*, y \rangle\,, \quad \forall\, y \in Y\} \\ &= \{y^* \in Y' \mid p(y) \geq \langle y^*, y \rangle\,, \quad \forall\, y \in Y\} = \partial p(0) = K^* \end{aligned}$$

and

$$\begin{aligned} \partial (f \circ A)(x_o) &= \{x^* \in X' \mid d(f \circ A)(x_o; x) \geq \langle x^*, x \rangle\,, \quad \forall\, x \in X\} \\ &= \{x^* \in X' \mid q(x) \geq \langle x^*, x \rangle\,, \quad \forall\, x \in X\} = \partial q(0) = C^* \end{aligned}$$

and the proof is complete. ∎

An useful variant is stated in the following:

Corollary 10.4.1 *Let $A : X \mapsto Y$ be a nonlinear operator which is differentiable at a point $x_o \in X$ with derivative $dA(x_o) : X \mapsto Y$ linear and continuous. Let further $f : Y \mapsto \Re \cup \{+\infty\}$ be a functional which is locally subdifferentiable at $A(x_o) \in Y$ and assume that the following identity holds:*

$$d(f \circ A)(x_o; x) = df[A(x_o); dA(x_o)x]\,, \quad \forall\, x \in X\,.$$

Then we have

$$\partial (f \circ A)(x_o) = \overline{\partial f[A(x_o)] \circ [dA(x_o)]} = \overline{[dA(x_o)]' \partial f[A(x_o)]}$$

if and only if it results:

$$(\overline{p \circ L})(x) = \overline{p}(Lx)\,, \quad \forall\, x \in X\,,$$

where $p(y) := df[A(x_o); y]$ *and* $L := dA(x_o)$, *a superimposed bar denoting the closure.*

Proof. The result is directly inferred from the theorem above by noting that the assumed identity amounts to require that $q = p \circ L$. ■

It has to be remarked that the chain-rule for one-sided directional derivatives assumed in the statement of the corollary holds trivially for every affine operator A. Moreover the necessary and sufficient condition is fulfilled when the sublinear functional p is closed. The next result shows that the addition rule for subdifferentials can be directly derived by applying the result provided by the chain-rule theorem.

Theorem 10.4.2 (New proof of the classical addition rule) *We consider the functionals* $f_i : X \mapsto \Re \cup \{+\infty\}$ *with* $i = 1, \cdots, n$ *and assume that they are locally subdifferentiable at* $x_o \in X$ *with* $p_i(x) := df_i(x_o; x)$. *The following addition rule then holds:*

$$\partial(\textstyle\sum_{i=1}^n f_i)(x_o) = \overline{\textstyle\sum_{i=1}^n \partial f_i(x_o)}$$

if and only if

$$(\overline{\textstyle\sum_{i=1}^n p_i})(x) = \textstyle\sum_{i=1}^n \overline{p_i}(x)\,, \quad \forall\, x \in X\,.$$

Proof. Let $A : X \mapsto X^n$ be the iteration operator defined as

$$Ax = |x_i|\,, \quad x_i = x\,, \quad i = 1, \cdots, n\,.$$

The dual operator $A' : X' \mapsto (X^n)'$ meets the identity

$$\langle A'|x_i^*|, x \rangle = \langle |x_i^*|, Ax \rangle = \textstyle\sum_{i=1}^n \langle x_i^*, x \rangle = \langle \sum_{i=1}^n x_i^*, x \rangle\,, \quad \forall\, x \in X\,,$$

and hence is the addition operator

$$A'|x_i^*| = \textstyle\sum_{i=1}^n x_i^*\,.$$

Defining the functional $f : X^n \mapsto \Re \cup \{+\infty\}$ as: $f(|x_i|) := \sum_{i=1}^n f_i(x_i)$, we have $(f \circ A)(x) = \sum_{i=1}^n f_i(x)$ and hence

$$\partial(f \circ A)(x_o) = \partial(\textstyle\sum_{i=1}^n f_i)(x_o)\,.$$

On the other hand,

$$\begin{aligned} df(Ax_o; |x_i|) &= \lim_{\alpha \downarrow 0} \frac{1}{\alpha} [f(Ax_o + \alpha |x_i|) - f(Ax_o)] \\ &= \lim_{\alpha \downarrow 0} \frac{1}{\alpha} [\textstyle\sum_{i=1}^n [f_i(x_o + \alpha x_i) - f_i(x_o)] = \sum_{i=1}^n df_i(x_o; x_i) \,. \end{aligned}$$

By the definition of local subdifferential we then get

$$|x_i^*| \in \partial f(Ax_o) \iff x_i^* \in \partial f_i(x_o)$$

so that

$$A' \partial f(Ax_o) = \textstyle\sum_{i=1}^n \partial f_i(x_o) \,,$$

Finally, noticing that

$$p(|x_i|) := df[Ax_o; |x_i|] = \textstyle\sum_{i=1}^n df_i[x_o; x_i] := \sum_{i=1}^n p_i(x_i) \,,$$

$$(\overline{p \circ A})(x) = (\overline{\textstyle\sum_{i=1}^n p_i})(x) \,,$$

$$\overline{p}(Ax) = \textstyle\sum_{i=1}^n \overline{p_i}(x) \,,$$

the proof follows from the result contributed in the chain-rule theorem above.

■

Corollary 10.4.2 *We consider the functionals* $f_i : X \mapsto \Re \cup \{+\infty\}$ *with* $i = 1, \cdots, n$, *and assume that they are regularly locally subdifferentiable at* $x_o \in X$. *The following addition rule then holds:*

$$\partial (\textstyle\sum_{i=1}^n f_i)(x_o) = \overline{\textstyle\sum_{i=1}^n \partial f_i(x_o)} \,.$$

Proof. The result follows at once by theorem 10.4.2, observing that:

$$p_i \quad (i = 1, \cdots, n) \quad \text{closed} \quad \Longrightarrow \quad \textstyle\sum_{i=1}^n p_i \quad \text{closed} \,.$$

■

We now derive a special case of the chain-rule formula which is referred to later when dealing with the existence of Kuhn and Tucker vectors in convex optimization.

A special case. Let $A : X \mapsto Y$ be a continuous affine operator, that is,

$$A(x) = L(x) + c$$

with $L : X \mapsto Y$ linear and continuous and $c \in Y$. Let further $f : Y \mapsto \Re \cup \{+\infty\}$ be the convex indicator of the point $\{A(x_o)\}$:

$$f(y) = \mathbf{ind}_{\{A(x_o)\}}(y)\,, \quad \forall\, y \in Y\,.$$

The chain-rule for one-sided directional derivatives holds true since A is affine. Moreover the functionals:

$$p(y) := df[A(x_o); y] = \mathbf{ind}_{\{0\}}(y)\,,$$

$$(p \circ L)(x) := df[A(x_o); Lx] = \mathbf{ind}_{\{0\}}(Lx) = \mathbf{ind}_{\{Ker L\}}(x)\,,$$

turn out to be sublinear, proper and closed. On the basis of the corollary to the chain-rule theorem provided above we may then that

$$\partial(f \circ A)(x_o) = \overline{L'Y'} = \overline{\operatorname{Im} L'}\,.$$

The particular case when $Y = \Re$ will be of special interest in the sequel. In this case we may write:

$$A(x) = \langle a^*, x \rangle + c \quad \text{with} \quad a^* \in X', \quad c \in \Re\,.$$

Note that now $L = a^* : X \mapsto \Re$ and $L' : \Re \mapsto X'$ with $Lx = \langle a^*, x \rangle$ and $L'\alpha = \alpha a^*$. It follows that $\operatorname{Im} L' = \operatorname{Lin}\{a^*\}$ is a closed subspace and hence

$$\partial(f \circ A)(x_o) = \operatorname{Im} L' = \operatorname{Lin}\{a^*\} = L'\partial f[A(x_o)] = \partial f[A(x_o)]L = \Re a^*$$

which is the formula of future interest. Two significant examples are reported hereafter to enlighten the meaning of the conditions required for the validity of the chain-rule formula.

Examples. The first example shows that, when the necessary and sufficient condition for the validity of the chain-rule formula is not satisfied, the two convex sets involved in the formula can in fact be quite different from one another.

Let f be the convex indicator of a circular set in $\Re^2$ centered at the origin and let $(x_o, 0)$ be a point on its boundary. The one-sided directional derivative of f at $(x_o, 0)$ is the proper sublinear functional $p : \Re^2 \mapsto \Re$ given by

$$p(x,y) = \begin{cases} 0 & \text{for } x < 0 \text{ and at the origin}\,, \\ +\infty & \text{elsewhere.} \end{cases}$$

Denoting the orthogonal projector on the axis $\Re_y$ by $L = L'$ we have

$$(p \circ L) = \mathbf{ind}_{\{\Re_x\}} \quad \text{and then} \quad \partial(p \circ L)(0,0) = \Re_y\,.$$

On the other hand,

$$\partial p(0,0) = \Re_x^+ \quad \text{so that} \quad L' \partial p(0,0) = L' \Re_x^+ = (0,0)\,.$$

The second example provides a situation in which all the assumptions set forth in the corollary are met but still the two convex sets fail to be equal since the second one is nonclosed. Let K^* be the iperbolic convex set in $\Re^2$ defined by

$$K^* := \{(x^*, y^*) \in \Re^2 \mid x^* y^* \geq 1\}$$

and let p be its support functional:

$$p(x,y) := \sup\{\langle x^*, x\rangle + \langle y^*, y\rangle \mid (x^*, y^*) \in K^*\}\,.$$

Denoting the orthogonal projector on the axis $\Re_y$ again by $L = L'$ we then have

$$(p \circ L)(x,y) = \begin{cases} 0 & \text{on} \quad \Re_x \times \Re_y^-\,, \\ +\infty & \text{elsewhere}\,. \end{cases}$$

Hence $K^* = \partial(p \circ L)(0,0) = \Re_y^+$ but $L' \partial p(0,0) = L' K^* = \Re_y^+ - (0,0)$ which is open.

10.4.2 A new product rule formula

We present here the proof of a new product-rule formula of subdifferential calculus which deals with the composition of a monotone convex function and a convex functional. The original interest of the author for this kind of product rule arose in connection with subdifferential relations involving gauge-like functionals [194] which are composed by a monotone convex Young function and a sublinear MINKOWSKY functional. The new product-rule formula turns out to be of the utmost interest in dealing with minimization problems involving convex constraints expressed in terms of level sets of convex functionals. A new approach to KUHN and TUCKER theory of convex optimization can be founded upon these results is carried out in the next subsection. Two introductory lemmas, which the main theorem resorts to, are preliminarily reported hereafter.

Lemma 10.4.1 *Let $I = [\lambda_1, \lambda_2]$ be an interval belonging to the nonnegative real line and let C be a weakly compact convex set in X. Then the set IC is convex and closed if either a) $0 \notin C$, or b) I is compact (i.e., bounded).*

Proof. We first prove that C being convex, the set IC is convex too.

If $\bar{x}_1, \bar{x}_2 \in IC$ and $\alpha_1, \alpha_2 \geq 0$ with $\alpha_1 + \alpha_2 = 1$, we have

$$\alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2 = \alpha_1 m_1 x_1 + \alpha_2 m_2 x_2 \quad \text{with} \quad m_1, m_2 \in I, \quad x_1, x_2 \in C.$$

Now, by the convexity of C ([194], th. 3.2),

$$\alpha_1 m_1 x_1 + \alpha_2 m_2 x_2 \in \alpha_1 m_1 C + \alpha_2 m_2 C = (\alpha_1 m_1 + \alpha_2 m_2) C \subseteq IC,$$

the last inclusion holding true since $\alpha_1 m_1 + \alpha_2 m_2 \in I$, by the convexity of I.

To prove the weak closedness of IC, we consider a weak limit point z of IC and a sequence $\{a_k x_k\}$, with $a_k \in I$ and $x_k \in C$, converging weakly to z:

$$\langle x^*, a_k x_k \rangle \mapsto \langle x^*, z \rangle, \quad \forall\, x^* \in X'.$$

C being weakly compact in X, we may assume that the sequence $\{x_k\}$ is weakly convergent to a point $x \in C$. Under assumption a) we then infer that $x \neq 0$ so that there will exist an $\overline{x}^*$ such that:

$$\langle \overline{x}^*, x_k \rangle \mapsto \langle \overline{x}^*, x \rangle > 0.$$

For a sufficiently large k, $\langle \overline{x}^*, x_k \rangle \geq \xi > 0$, and hence the sequence $\{a_k\}$ cannot be unbounded. In fact otherwise $\langle x^*, a_k x_k \rangle \geq a_k \xi \mapsto +\infty$, contrary to the assumption that $a_k x_k \overset{\text{w}}{\mapsto} z$. Under assumption b) the boundedness of the sequence $\{a_k\}$ is a trivial consequence of the boundedness of I. In both cases we may then assume that $a_k \mapsto a \in I$ and $x_k \overset{\text{w}}{\mapsto} x \in C$. As a consequence we get that $\{a_k x_k\} \overset{\text{w}}{\mapsto} ax$ and hence $z = ax \in IC$. ■

Lemma 10.4.2 *Let $f : X \mapsto \Re$ be a continuous nonconstant convex functional. Denoting its zero level set by N, if there is a vector $x_- \in N$ such that $f(x_-) < 0$ then it results:*

$$int\, N = N_- := \{x \in X \mid f(x) < 0\},$$

$$bnd\,N = N_o := \{x \in X \mid f(x) = 0\}\,,$$

and both sets turn out to be nonempty.

Proof. Since $f(x_-) < 0$, by the continuity of f a neighbourhood $\mathcal{N}(x_-)$ exists such that $f(x) < 0\,, \quad \forall x \in \mathcal{N}(x_-)$. Hence $\mathcal{N}(x_-) \subset N$ so that $x_- \in \mathrm{int}\,N$. Further, f being nonconstant and negative at x_-, by convexity there will be an $x_o \in X$ such that $f(x_o) = 0$. Let $S(x_o; x_-) \subset N$ be the segment joining x_o and x_- and let $L(x_o; x_-)$ be the line generated by $S(x_o; x_-)$ (see fig. 10.1(a)). Setting $f_L(t) = f[\hat{x}(t)]$ with $\hat{x}(t) = (1-t)x_o + tx_-$, $t \in \Re\}$ we have

$$f_L(0) = 0 \quad \text{and} \quad f_L(1) < 0\,.$$

Hence, by convexity (see fig. 10.1(a)):

$$f_L(t) < 0 \ \text{ for } 0 < t \leq 1 \quad \text{and} \quad f_L(t) > 0 \ \text{ for } t < 0\,.$$

We may then conclude that

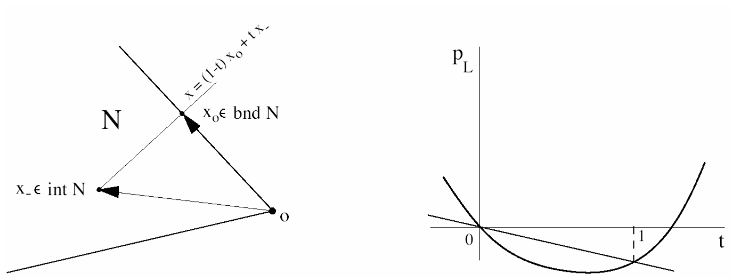

Figure 10.1: (a) The zero level set of $f(x)$ - (b) The graph of $f_L(t)$.

$$N_- \subseteq \mathrm{int}\,N \quad \text{and} \quad N_o \subseteq \mathrm{bnd}\,N$$

and the relations:

$$\mathrm{int}\,N = N \backslash \mathrm{bnd}\,N \subseteq N \backslash N_o = N_-\,,$$

$$\mathrm{bnd}\,N = N \backslash \mathrm{int}\,N \subseteq N \backslash N_- = N_o$$

yield the converse inclusions. ■

The main theorem providing the new product rule can now be stated.

Theorem 10.4.3 (The new product rule) *Let* $m : \Re \cup \{+\infty\} \mapsto \Re \cup \{+\infty\}$ *be a monotone convex function with* $m(+\infty) = +\infty$ *and let* $k : X \mapsto \Re \cup \{+\infty\}$ *a proper convex functional continuous at* $x \in X$ *. Then, if* x *is not a minimum point of* k *and* m *is subdifferentiable at* $k(x)$ *, setting* $f = m \circ k$ *, results in*

$$\partial f(x) = \partial m[k(x)] \partial k(x) \,.$$

Proof. The proof is carried out in two steps. First we provide a representation formula for the closure of the one-sided directional derivative of f ; then recourse to the two preliminary lemmas will yield the result. To provide the representation formula, given a director $h \in X - \{0\}$, we define the convex real function $\chi : \Re^+ \mapsto \Re$ as the restriction of k to the half-line starting at x and directed along h : $\chi(\alpha) := k(x + \alpha h)$ so that $\chi'(0) := dk(x; h)$. In investigating the behavior of $df(x; \cdot)$ it is basic to consider the zero level set of $dk(x; \cdot)$. First we observe that the continuity of k at x implies [206] the continuity of the sublinear function $dk(x; h)$ as a function of h . Its zero level set $N = \{h \in X \mid dk(x; h) \leq 0\}$ is then a closed convex cone. Since by assumption x is not a minimum point for k , the preliminary lemma, lemma 10.4.2, states that the interior and the boundary of N are not empty, being $dk(x; \cdot) < 0$ in $\operatorname{int} N$ and $dk(x; \cdot) = 0$ on $\operatorname{bnd} N$. The derivative $df(x; h)$ of the product functional $f = m \circ k$ can be immediately computed along the directions $h \in \operatorname{int} N$ and $h \notin N$. In fact if $dk(x; h) = \chi'(0)$ does not vanish, $\alpha \downarrow 0$ implies that definitively either $\chi(\alpha) \downarrow \chi(0)$ if $\chi'(0) > 0$ or $\chi(\alpha) \uparrow \chi(0)$ if $\chi'(0) < 0$ (see fig. 10.2). Hence, denoting the right and left derivates of m at the point $k(x) = \chi(0)$ by ${m'}_+$ and ${m'}_-$, it will be seen that

$$\begin{aligned} df(x; h) &= \lim_{\alpha \downarrow 0} \frac{1}{\alpha} \Big[f(x + \alpha h) - f(x) \Big] = \lim_{\alpha \downarrow 0} \frac{1}{\alpha} \Big[m[k(x + \alpha h)] - m[k(x)] \Big] \\ &= \lim_{\alpha \downarrow 0} \frac{1}{\alpha} \Big[m[\chi(\alpha)] - m[\chi(0)] \Big] \\ &= \lim_{\alpha \downarrow 0} \frac{m[\chi(\alpha)] - m[\chi(0)]}{\chi(\alpha) - \chi(0)} \, \frac{\chi(\alpha) - \chi(0)}{\alpha} \\ &= \lim_{\chi(\alpha) \downarrow \chi(0)} \frac{m[\chi(\alpha)] - m[\chi(0)]}{\chi(\alpha) - \chi(0)} \lim_{\alpha \downarrow 0} \frac{\chi(\alpha) - \chi(0)}{\alpha} \\ &= {m'}_+ \, \chi'(0) \,, \end{aligned}$$

if $\chi'(0) > 0$. Apparently $df(x; h) = {m'}_- \, \chi'(0)$ if $\chi'(0) < 0$.

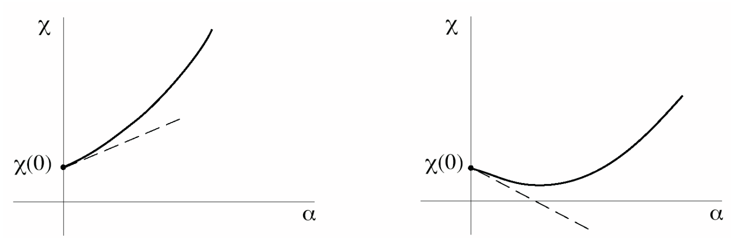

Figure 10.2: (a) $\chi'(0) > 0$ and (b) $\chi'(0) < 0$.

A more detailed discussion has to be made when $h \in \mathrm{bnd}\, N$, so that $\chi'(0) = 0$. In this case, as shown in fig. 10.3, the convexity of k implies that either $\chi(\alpha)$ goes to $\chi(0)$ with a strict monotonic descent or it attains the value $\chi(0)$ for some $\alpha > 0$ and then remains definitively constant.

In both cases $df(x; h) = 0$ if the right derivative m'_+ is finite. In fact, in the case of figure 10.3(a), the formula $df(x; h) = m'_+ \, dk(x; h)$ holds with $dk(x; h) = 0$; in the case of figure 10.3(b) the conclusion is trivial.

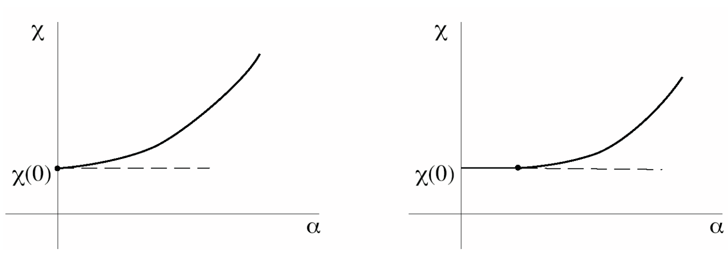

Figure 10.3: Graphs of $\chi(\alpha)$ for $\chi'(0) = 0$). (a) Monotonic descent and (b) definitive constancy.

We may then conclude that:

$$df(x; \cdot) = \begin{cases} 0 & \text{on} \quad \mathrm{bnd}\, N, \\ m'_- \, dk(x; \cdot) \le 0 & \text{in} \quad \mathrm{int}\, N, \\ m'_+ \, dk(x; \cdot) \ge 0 & \text{outside} \quad N, \end{cases}$$

so that the following formula holds:

$$df(x; h) = \sup_{\lambda \in I} \lambda dk(x; h) \quad \text{with } I = [m'_-, m'_+].$$

An indecisive situation occurs instead when $m'_+ = +\infty$ since, in the case of figure 10.3a, $df(x;h) = +\infty$.

Noticing that $\partial m[k(x)] = I = [m'_-, m'_+]$, the assumed subdifferentiability of m at $k(x)$ ensures that $m'_- < +\infty$. Hence we get:

$$df(x;\cdot) = \begin{cases} 0 \ \text{ or} + \infty & \text{on} \quad \text{bnd}\, N, \\ m'_-\, dk(x;\cdot) \leq 0 & \text{in} \quad \text{int}\, N\,, \\ m'_+\, dk(x;\cdot) = +\infty & \text{outside} \quad N\,. \end{cases}$$

To resolve the indecisive situation on $\text{bnd}\, N$ we observe that, $dk(x;\cdot)$ being continuous in X and vanishing on $\text{bnd}\, N$, the restriction of $df(x;\cdot)$ to $\text{int}\, N$ can be extended by continuity to zero on $\text{bnd}\, N$.

As a consequence the closure of $df(x;\cdot)$ will vanish on $\text{bnd}\, N$, being equal to $df(x;\cdot)$ elsewhere:

$$\overline{df(x;\cdot)} = \begin{cases} 0 & \text{on} \quad \text{bnd}\, N\,, \\ m'_-\, dk(x;\cdot) \leq 0 & \text{in} \quad \text{int}\, N\,, \\ m'_+\, dk(x;\cdot) = +\infty & \text{outside} \quad N\,. \end{cases}$$

From the analysis above we infer then the general validity of the formula

$$\overline{df(x;h)} = \sup_{\lambda \in I} \lambda dk(x;h) \quad \text{with } I = [m'_-, m'_+]\,,$$

holding whether m'_+ is finite or not.

To get the product rule we finally remark that, by the continuity of $dk(x;\cdot)$,

$$dk(x;h) = \sup\{\langle x^*, h\rangle \mid x^* \in \partial k(x)\}$$

so that the formula above may be rewritten as

$$\overline{df(x;h)} = \sup_{\lambda \in I}\{\lambda \sup\{\langle x^*, h\rangle \mid x^* \in \partial k(x)\}\} = \sup\{\langle x^*, h\rangle \mid x^* \in I\partial k(x)\}\,.$$

The set $I\partial k(x) = \partial m[k(x)]\partial k(x)$ being convex by lemma 10.4.1, we then get

$$\partial f(x) = \overline{\partial m[k(x)]\partial k(x)}\,.$$

Finally we observe that, by the continuity of k at x, the convex set $\partial k(x)$ is nonempty, closed and weakly compact in X' ([195], prop. 10.c.); further it does not contain the origin since x is not a minimum point for k. By lemma 10.4.1 we may then infer the closure of the set $\partial m[k(x)]\partial k(x)$ and the proof is complete. ∎

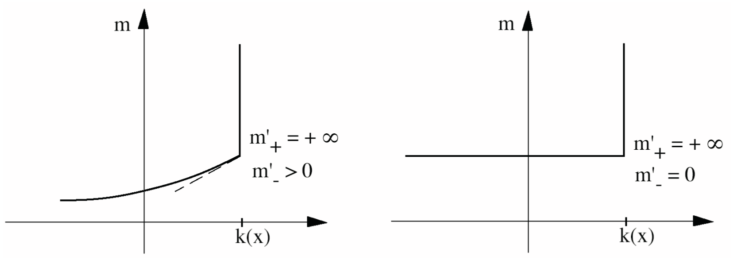

Figure 10.4: Graphs of m when $m'_+ = +\infty$.

Typical shapes of the monotone convex function m in the case when $m'_+ = +\infty$ are shown in figure 10.4 depending on whether $m'_- > 0$ or $m'_- = 0$. The latter case reveals that a significant special choice for m is the convex indicator of the nonpositive real axis. This is in fact the choice to be made in discussing convex optimization problems.

10.4.3 Applications to convex optimization

Given a proper convex functional $f : X \mapsto \Re \cup \{+\infty\}$ let us consider the following convex optimization problem:

$$\inf\{f(x) \mid x \in C\}$$

where C is the feasible set, defined by

$$C = C_g \cap C_h$$

with

$$\begin{aligned} C_g &= \{\cap C_i \mid \ i = 1, \cdots, n_1\}\,, \\ C_h &= \{\cap C_j \mid \ j = 1, \cdots, n_2\}\,, \\ C_i &= \{x \in X \mid g_i(x) \le 0\}\,, \\ C_j &= \{x \in X \mid h_j(x) = 0\}\,. \end{aligned}$$

In order that the optimization problem above be meaningful, we have to assume that the intersection between the feasible set and the domain of the objective functional is not empty, i.e., $\mathrm{dom}\, f \cap C \neq \emptyset$. Here $g_i : X \mapsto \Re$ are n_1 continuous convex functionals and $h_j : X \mapsto \Re$ are n_2 continuous affine functionals, that is, $h_j(x) = \langle a^*{}_j, x \rangle + c$ with $a^*{}_j \in X'$ and $c \in \Re$. Without loss of generality the functionals g_i and h_j can be assumed to be nonconstant; further it is natural to assume that each of the convex functionals g_i do assume negative values.

The following preliminary result is easily proved.

Lemma 10.4.3 *Let $g : X \mapsto \Re$ be a nonconstant continuous convex functional. Denoting its zero level set by N, if a vector $x_- \in N$ exists such that $g(x_-) < 0$ we have that*

$$\partial(\textbf{\textit{ind}}_{\{\Re^-\}} \circ g)(x) = \partial \textbf{\textit{ind}}_{\{\Re^-\}}[g(x)] \partial g_i(x) \,, \quad \forall x \in N \,.$$

Proof. By lemma 10.4.2 it follows that

$$\operatorname{int} N = N_- := \{x \in X \mid g(x) < 0\} \,,$$

$$\operatorname{bnd} N = N_o := \{x \in X \mid g(x) = 0\} \,,$$

and both sets turn out to be nonempty.

Now, if $x \in N_o$ the properties ensuring the validity of the new product-rule formula proved in the subsection 10.4.2 are fulfilled. On the other hand, if $x \in N_-$, by the continuity of g there is a neighborhood of x in which g is negative. The formula above then follows by observing that in this case $\partial \mathbf{ind}_{\{\Re^-\}}[g(x)] = 0$. ■

We are now ready to discuss the convex optimization problem considered above which can be conveniently reformulated as

$$\inf \psi(x) \quad \text{with} \quad \psi(x) = f(x) + \textstyle\sum_{i=1}^n \mathbf{ind}_{\{\Re^-\}}[g_i(x)] + \sum_{j=1}^m \mathbf{ind}_{\{0\}}[h_j(x)] \,.$$

Convex analysis tells us that

$$x_o = \arg\min \psi(x) \iff 0 \in \partial \psi(x_o) \quad \text{or explicitly:}$$

$$0 \in \partial \Big[f(x_o) + \textstyle\sum_{i=1}^n \mathbf{ind}_{\{\Re^-\}}[g_i(x_o)] + \sum_{j=1}^m \mathbf{ind}_{\{0\}}[h_j(x_o)] \Big] \,.$$

Under the validity of the addition rule of subdifferential calculus the extremum condition becomes

$$0 \in \partial f(x_o) + \textstyle\sum_{i=1}^n \partial(\mathbf{ind}_{\{\Re^-\}} \circ g_i)(x_o) + \sum_{j=1}^m \partial(\mathbf{ind}_{\{0\}} \circ h_j)(x_o) \,.$$

Here we apply the result contributed above in Lemma 10.4.3 to compute the subdifferentials related to inequality constraints:

$$\partial(\mathbf{ind}_{\{\Re^-\}} \circ g_i)(x_o) = \partial\mathbf{ind}_{\{\Re^-\}}[g_i(x_o)]\partial g_i(x_o)\,.$$

The new proof of the chain-rule provided in subsection 10.4.1 allows us to carry out computation of the subdifferentials related to equality constraints:

$$\partial(\mathbf{ind}_{\{0\}} \circ h_j)(x_o) = \partial\mathbf{ind}_{\{0\}}[h_j(x_o)]\partial h_j(x_o)\,.$$

Finally we observe that

$$\partial\,\mathbf{ind}_{\{\Re^-\}}[g_i(x_o)] = N_{\{\Re^-\}}[g_i(x_o)]\,,$$

$$\partial\,\mathbf{ind}_{\{0\}}[h_j(x_o)] = \Re\,, \quad \text{and} \quad \partial h_j(x_o) = {a^*}_j\,,$$

where $N_{\{\Re^-\}}[g_i(x_o)]$ is the normal cone to $\Re^-$ at the point $g_i(x_o)$. It turns out to be equal to $\{0\}$ when $g_i(x_o) < 0$ and to $\Re^+$ when $g_i(x_o) = 0$. The extremum condition above can thus be restated explicitly in terms of Kuhn and Tucker complementarity relations:

$$\begin{cases} \lambda_i \in \Re^+\,, \quad g_i(x_o) \in \Re^-\,,\ \lambda_i g_i(x_o) = 0 \quad i = 1, \cdots, n \\ \mu_j \in \Re\,, \quad\quad j = 1, \cdots, m \end{cases}$$

and of the related stationarity condition

$$0 \in \partial f(x_o) + \textstyle\sum_{i=1}^n \lambda_i \partial g_i(x_o) + \sum_{j=1}^m \mu_j {a^*}_j\,.$$

The corresponding Lagrangian is given by

$$L(x, \lambda_i, \mu_j) = f(x) + \sum_{i=1}^n \lambda_i g_i(x) + \sum_{j=1}^m \mu_j h_j(x) - \sum_{i=1}^n \mathbf{ind}_{\{\Re^+\}}(\lambda_i) - \sum_{j=1}^m \mathbf{ind}_{\{\Re\}}(\mu_j)\,,$$

where the last inessential term has been added for formal symmetry.

The Kuhn and Tucker conditions above are easily seen to be equivalent to the existence of a saddle point for the Lagrangian.

Classically the existence of Kuhn and Tucker multipliers is ensured by the fulfillment of Slater's conditions [207, 194]

$$\exists\,\overline{x} \in X \ \text{ such that } \ f(\overline{x}) < +\infty \ \text{ and } \ \begin{cases} g_i(\overline{x}) < 0\,, & i = 1, \cdots, n\,, \\ h_j(\overline{x}) = 0\,, & j = 1, \cdots, m\,, \end{cases}$$

i.e., by assuming that the intersection between the domain of the objective functional and the interior of the set C_g is nonempty.

According to the treatment developed in this paper the existence of KUHN and TUCKER multipliers can in fact be assessed under by far less stringent conditions; these amount in the obvious minimal requirement that the optimization problem is well posed (i.e., the intersection between the domain of the objective functional and the feasible set is nonempty) and in the further assumption that, at the optimal point, the property ensuring the validity of the addition rule is satisfied.

The graphical sketches in figure 10.5 exemplify the different assumptions about the feasible set $C = C_1 \cap C_2$ in the special case $n_1 = 2$ and $n_2 = 0$. SLATER condition is easily seen to be a straightforward consequence of the

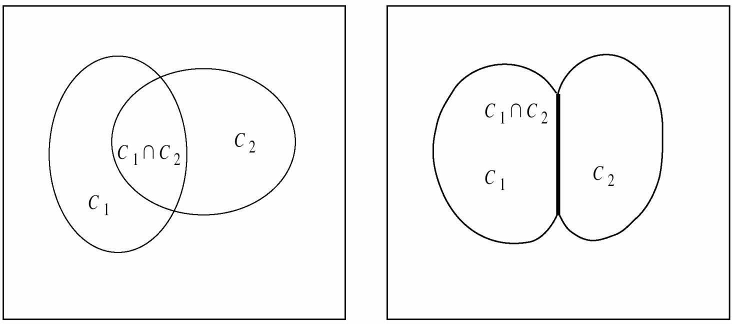

Figure 10.5: Assumption about the feasible set. (a) Slater's condition - (b) new requirement.

classical theorems on addition rule for subdifferentials [194, 204]. The new condition is based on the results provided in the present paper. Validity of the addition rule cannot however be imposed *a priori* but has to be verified *a posteriori* at the extremal point.

In this respect it has to be pointed out that when the optimal point x_o lies on the boundary of the set C_g the simple sufficient condition provided by corollary 10.4.2 results in a special requirement on the local shape of C_g around x_o.

In fact, when x_o belongs to the boundary of one of the sets C_i, the one-sided directional derivative $d(\mathbf{ind}_{\{\Re^-\}} \circ g_i)(x_o; \cdot)$ will be l.s.c. if and only if the monotone convex function $\mathbf{ind}_{\{\Re^-\}}$ is definitively constant towards zero along any direction h such that $dg(x_o; h) = 0$, i.e., $h \in \text{bnd}\, C_i$ (see lemma 10.4.2, theorem 10.4.3 and figure 10.3). This means that there must be an $\alpha_o > 0$ such that $g(x_o + \alpha h) = 0$ for $\alpha_o \geq \alpha \geq 0$. The boundary of C_g must thus have a conical shape around x_o as sketched in figure 10.6.

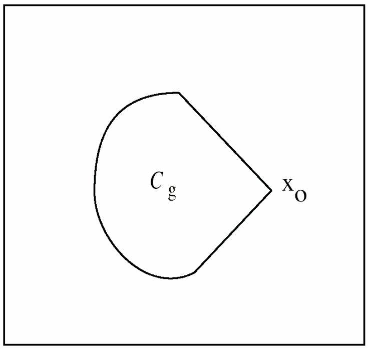

Figure 10.6: Local conical shape around the optimal point.

We finally provide an example in which SLATER's condition fails, but the existence of KUHN-TUCKER multipliers can still be assessed on the basis of the new results contributed above. To this end we consider a two-dimensional optimization problem for the convex function $f(x, y) = \frac{1}{2}(x^2 + y^2)$ under the following inequality constraints: $h_1 = x - 1 \leq 0$; $h_2 = -x + 1 \leq 0$; $h_3 = y - 2 \leq 0$; $h_4 = -y + 2 \leq 0$. It is apparent that the feasible set does have an empty interior so that SLATER's condition is not fulfilled. On the contrary the differentiability of the contraint functions ensures the validity of the addition rule so that the new requirement is satisfied. The feasible set C is depicted in figure 10.7 and a set of KUHN-TUCKER multipliers at the optimal point $x = 1$; $y = 1$ is given by: $\lambda_1 = 0\,, \lambda_2 = 1\,, \lambda_3 = 0\,, \lambda_4 = 1$.

10.5 Conclusions

The new approach to classical chain and addition rules of subdifferential calculus and the new product-rule formula presented in this paper have been shown to provide a useful and simple tool in the analysis of convex optimization problems. KUHN-TUCKER optimality conditions have been proved under minimal assumptions on the data. Further applications of the results contributed here

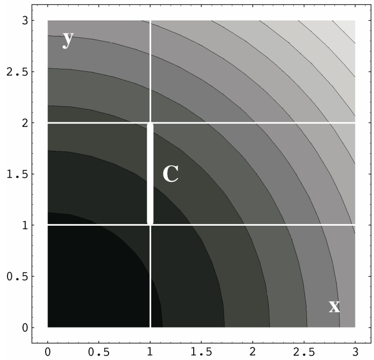

Figure 10.7: Feasible set and contour plot of the objective function.

can be envisaged in different areas of mathematical physics. The original motivation for the study stemmed from problems in the theory of plasticity. In fact, starting from the classical normality rule of the plastic flow to the convex domain of admissible static states, the new product rule provides a simple and effective tool to derive the equivalent expression of the flow in terms of plastic multipliers and gradients of the yield modes. A comprehensive treatment of the subject can be found in three papers by the author and coworkers [208, 209, 210].

Bibliography

[1] Abraham R., Marsden J.E.: Foundations of Mechanics, Benjamin, New York (1967).

[2] Abraham R., Marsden J.E.: Foundations of Mechanics, second edition, the Benjamin/Cummings Publishing Company, Reading Massachusetts, (1978).

[3] Abraham R., Marsden J.E., Ratiu T.: Manifolds, Tensor Analysis, and Applications, second edition, Springer Verlag, New York (1988), third edition, Springer Verlag, New York (2002).

[4] Ambrose W., Palais R.S. & Singer I.M., Sprays, Anais da Acad. Bras. de Ciencias, 32-2, 163-178 (1960).

[5] Antman S. S., Marlow R. S.: Material Constraints, Lagrange Multipliers, and Compatibility. Applications to Rod and Shell Theories, Arch. Rational Mech. Anal., vol.116, 257-299 (1991).

[6] Arnold V.I.: Mathematical methods of classical mechanics, Springer Verlag, New York (1989), translated from Mir publishers, Moscow (1979).

[7] Arnold V.I., Kozlov V.V., Neishtadt A.I.: Dynamical Systems, III, Encyclopaedia of Mathematical Sciences, V.I. Arnold (Ed.), Vol. 3, Springer-Verlag, New York-Berlin, (1988).

[8] Avez A.: Ergodic theory of dynamical systems, Minneapolis (1967).

[9] Arnold V.I., Avez A.: Ergodic problems of classical mechanics, Benjamin, New york (1968).

[10] Babuška I., Aziz A. K.: The Mathematical Foundations of the Finite Element Method with Application to Partial Differential Equations - Part I, Academic Press, New York (1972).

[11] Babuška I.: The finite element method with lagrangian multipliers, Numer. Math., vol.20, pp.179-192, (1973).

[12] Banach S.: Théorie des Opérations Linéaires, Monografje Matematyczne, Warsaw (1932). Reprinted by Chelsea, New York (1955).

[13] Barenblatt G.I.: Mathematical Theory of Equilibrium Cracks in Brittle Fracture, Advances in Applied Mechanics, vol.7, Academic Press (1962).

[14] Bernardi C., Canuto C., Maday Y.: Generalized inf-sup conditions for Chebyshev spectral approximation of the Stokes problem, SIAM J. Numer. Anal., vol.25, pp.1237-1271, (1988).

[15] Bossavit A.: Applied Differential Geometry (A compendium), available on the web, pp.1-30, (2002).

[16] Brauer R.: On the relation between the orthogonal group and the unimodular group, Arch. Rational Mech. Anal., vol.18, pp.97-99, (1965).

[17] Brezis H.: Analyse Fonctionnelle, Théorie et applications, Masson Editeur, Paris (1983).

[18] Brezzi F.: On the existence, uniqueness and approximation of saddle point problems arising from lagrangian multipliers, R.A.I.R.O. Anal. Numer., vol.8, pp.129-151, (1974).

[19] Brezzi F., Gilardi G.: FEM Mathematics, in the Finite-Element Handbook, 1.1-1.76, Kardestuncer editor, McGraw-Hill, (1987).

[20] Brezzi F., Fortin M.: Mixed and Hybrid Finite Element Methods, Springer-Verlag, New York (1991).

[21] Brezzi F., Marini L.D.: On the numerical solution of plate bending problems by hybrid methods, R.A.I.R.O. Ser. Rouge Anal. Numér. R-3, pp.5-50, (1975).

[22] Brickell F., Clark R.S.: Differentiable Manifolds, van Nostrand Reinhold, London (1970).

[23] Boothby W.M.: An introduction to differentiable manifolds and riemannian geometry, Academic Press, Orlando, Florida (1986).

[24] Cartan E.: Leçons sur les invariants intégraux, Hermann, Paris (1922).

[25] Cartan H.: Notions d'algèbre différentielle, Colloque de topologie de Bruxelles, 1950, Masson (1951).

[26] Chern S.S.: Differential Manifolds, Mathematical Lecture Notes, University of Chicago (1959).

[27] Choquet-Bruhat Y.: Géométrie Différentielle et Systèmes extérieurs, Travaux et recherches matématiques, Collège de France, Dunod, Paris (1970).

[28] Choquet-Bruhat Y., DeWitt-Morette C.: Analysis, Manifolds and Physics, North-Holland Publishing Company, New York, (1982).

[29] Ciarlet P.G.: The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam-New York (1978).

[30] Ciarlet P.G.: Theory of Shells, Mathematical Elasticity, vol.III, North-Holand, Amsterdam (2000).

[31] Darboux G.: Sur le problème de Pfaff, Bull. Sci. Math. 6: 14-36, 49-68, (1882).

[32] Davini C.: Some remarks on the continuum theory of defects in solids, Int. Journal of Solids and Structures, vol.38, pp.1169-1182, (2001).

[33] de Leon M., Epstein M.: The geometry of uniformity in second grade elasticity, Acta Mechanica, vol.114, pp.217-224, (1996).

[34] de Rham G.: Differentiable Manifolds, Forms, Currents, Harmonic Forms, Springer-Verlag, Berlin (1984) translated from the French original edition: Varietés differentiables, Publisher Hermann, Paris (1955).

[35] Del Piero G.: A mathematical characterization of plastic phenomena in the context of Noll's new theory of simple materials, Engineering Fracture Mechanics, pp. 633-639, Vol. 21, No. 4 (1985).

[36] Deschamps G.A.: Exterior differential forms, Mathematics Applied to Physics, Ed. Roubine É., pp. 111-161, Springer-Verlag, Berlin (1970).

[37] Di Carlo A., Quiligotti S.: Growth and balance, Mechanics Research Communications, vol.29, pp.449-456, (2002).

[38] Dieudonne J.: Foundations of Modern Analysis, Academic Press, New York and London (1969).

[39] Duvaut G., Lions J.L.: Inequalities in Mechanics and Physics, Springer-Verlag, New York (1976), translated from Les Inéquations en Méchanique et en Physique, Dunod, Paris (1972).

[40] Ekeland I., Temam R.: Convex Analysis and Variational Problems, North-Holland Publishing Company, Amsterdam (1976).

[41] Epstein M., Maugin G.A.: The energy-momentum tensor and material uniformity in finite elasticity, Acta Mechanica, vol.83, pp.127-133, (1990).

[42] Epstein M., Maugin G.A.: On the geometrical material structure of anelasticity, Acta Mechanica, vol.115, pp.119-131, (1996).

[43] Epstein M., Maugin G.A.: Thermomechanics of volumetric growth in uniform bodies, Int. Journal of Plasticity, vol.16, 1-3, pp.951-978, (2000).

[44] Epstein M.: The Eshelby tensor and the theory of continuous distributions of inhomogeneities, Mechanics Research Communications, vol.29, pp.501-506, (2002).

[45] Ehresmann C.: Les connexions infinitésimal dans une espace fibré différentiable, Colloq. de Topologie Bruxelles, 1950 , G. Thone & Masson, pp. 29-55 (1951) .

[46] Ericksen J.: Tensor Fields, Handbuch der Physik, vol.III, 1, Springer-Verlag, Berlin (1960).

[47] Eshelby J.D.: The force on an elastic singularity, Phil. Trans. royal soc. London, vol.A-244, pp.87-112, (1951).

[48] Eshelby J.D.: The continuum theory of lattice defects, Solid State Phys., vol.3, pp.79-144, (1960).

[49] Eshelby J.D.: The elastic energy-momentum tensor, Journal of Elasticity, vol.5, pp.321-335, (1975).

[50] Esposito G., Marmo G., Sudarshan G.: From Classical to Quantum Mechanics, Cambridge University Press, (2004).

[51] Euler L.: Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti, additamentum II (1744), in: C. Carathéodory, ed. Opera Omnia, pp. LII-LV, 298-308, Fussli, Zürich, Switzerland (1952).

[52] Gray C.G., Karl G. and Novikov V.A.: The Four Variational Principles of Mechanics, Annals of Physics 251, 125 (1996).

[53] Gray C.G., Karl G. and Novikov V.A.: Progress in classical and quantum variational principles, Rep. Prog. Phys., 67, pp. 159-208 (2004).

[54] Fenchel W.: On conjugate convex functions, Canad. J. Math., vol.1, pp.73-77, (1949).

[55] Fenchel W.: Convex cones, Sets and Functions, Princeton Univ. Press, Princeton (1951).

[56] Fichera G., Existence Theorems in Elasticity, Handbuch der Physik, vol.VI/a, Springer-Verlag, Berlin (1972).

[57] Frölicher A., Nijenhuis A.: Theory of vector valued differential forms: Derivations in the graded ring of differential forms, Part I, Proc. Kon. Ned. Akad. A, 59 (1956) 338-359.

[58] Frölicher A., Nijenhuis A.: Theory of vector valued differential forms: Almost complex structures, Part II, Proc. Kon. Ned. Akad. A, 59 (1958) 414-429. ???????

[59] Gantmacher F.: Lectures in Analytical Mechanics, Mir publishers, Moscow (1970).

[60] Godbillon C.: Géometrie différentielle et Méchanique Analytique, Hermann, Paris (1969).

[61] Godbillon C.: The Principle of Maupertuis, in Chillingworth D. (ed.) Proc.symp. On diff. eq. and dyn. sys, Lec. Notes Math., Springer, New York, 91-94 (1971).

[62] Goldschmidt H. and Sternberg Sh.: The Hamilton-Cartan formalism in the calculus of variations, Ann. Inst. Fourier (Grenoble) 23, 203-267 (1973).

[63] Goldstein H.: Classical Mechanics, Addison Wesley, Reading, Mass. (1950).

[64] Gray C.G., Karl G., Novikov V.A.: Progress in Classical and Quantum Variational Principles, physics/0312071 Classical Physics, (2003).

[65] Green G.: On the Propagation of Light in Crystallized Media, Trans. Cambridge Phil. Soc., vol. vol.7, 1, pp.121-140, (1839). Papers, pp.293-311, (1841).

[66] Grifone J.: Structure presque tangente et connexions I, Annales de l'Institut Fourier, tome 22, 1, pp.287-334, (1972).

[67] Grifone J.: Structure presque tangente et connexions II, Annales de l'Institut Fourier, tome 22, 3, pp.291-338, (1972).

[68] Grifone J., Mehdi M.: On the geometry of Lagrangian mechanics with non-holonomic constraints, Journal of Geometry and Physics, 30, 187-203, (1999).

[69] Gurtin M.E.: An Introduction to Continuum Mechanics, Academic Press, San Diego (1981).

[70] Gurtin M.E.: Configurational forces as basics concepts of continuum physics, Applied Mathematical Sciences, vol.137, Springer, (2000).

[71] Gutowski R.: Analytical Mechanics, Foundations of Mechanics, part I, chapter p.108-119, Ed. Henryk Zorski, Elsevier, (1992).

[72] Hiriart-Urruty J.B., Lemaréchal C.: Convex Analysis and Minimization Algorithms, Springer, Berlin (1993).

[73] Ioffe A.D., Tihomirov V.M.: The Theory of Extremal Problems, Nauka, Moscow, English translation, North-Holland, Amsterdam (1979).

[74] Jacobi C.G.J.: Note sur l'intègration des équations differentielles de la dynamique, C.R Acad Sci. Paris, **5** (61) (1837),

[75] Jacobi C.G.J.: Vorlesungen über Dynamik, zweite ausgabe, herausgegeben von A. Clebsch, Verlag G. Reimer, Berlin 1884, Lectures on Dynamics, reprinted by Chelsea, New York 1969.

[76] John F.: Partial differential equations, Mathematics Applied to Physics, pp. 229-347, Ed. Roubine É., Springer-Verlag, Berlin (1970).

[77] Kline J.: Espaces variationnels et Méchanique, Annales de l'Institut Fourier, tome 12, (1962), p. 1-124.

[78] Kline J.: Les systémes dynamiques abstraits, Annales de l'Institut Fourier, tome 13, (1963), p. 191-202.

[79] Kline J., Voutier, A.: Formes extérieurs génératrices de sprays, Annales de l'Institut Fourier, tome 18, (1968), p. 241-260.

[80] Kolar I., Michor P.W., Slovak J.: Natural operations in differential geometry, Springer-Verlag, (1993).

[81] Kobayashi, S.: Theory of connections, Ann. Mat. Pura Appl., 42, 119-194 (1957).

[82] Kobayashi, S., Nomizu K.: Foundations of Differential Geometry, Vol. 1, Wiley Classics Library. New York, (1963), Wiley-Interscience (1996).

[83] Kondo K.: Geometry of elastic deformation and incompatibility, Memoirs of the unifying study of the basic problems in engineering science by means of geometry, Gakujutsu Benken Fukyu-Kai, Tokyo (1955).

[84] Koszul J.L.: Homologie et cohomologie des algebres de Lie, Bulletin de la Société Mathématique 78: 65-127 (1950).

[85] Kratochvil J.: On a finite strain theory of elastic inelastic materials, Acta Mechanica, vol.16, pp.127-142, (1973).

[86] Kröner E.: Allgemeine Kontinuumtheorie der Versetzungen und Eigenspannungen, Arch. Rational Mech. Anal., vol.4, pp.273-334, (1960).

[87] Kupka I., Oliva W.M.: The Non-Holonomic Mechanics, Journal of Differential Equations 169, 169-189 (2001).

[88] Ladyszhenskaya O.A., The Mathematical Theory of Viscous Incompressible Flow, Gordon and Breach, New York (1969).

[89] Lagrange J.L.: Mechanique Analytique, in: J.A. Serret, ed. Ouvres, Gauthiers-Villars Paris 1867.

[90] Landau L.D., Lifšits E.M.: Mechanika, Mir ed., Moscow (1976).

[91] Lang S.: Differential and Riemannian Manifolds. Springer-Verlag, (1995).

[92] Laurent J.P.: Approximation et Optimisation, Hermann, Paris (1972).

[93] Lee E.H.: Elastic-plastic deformations at finite strains, ASME Trans. J. Appl. Mech., vol.36, 1, (1969).

[94] Lei Y., O'Dowd N.P., Webster G.A.: Fracture mechanics analysis of a crack in a residual stress field, Int. Journal Fracture, vol.106, 195-206, (2000).

[95] Lei Y.: J-integral evaluation for cases involving non-proportional stressing, Engineering Fracture Mechanics, vol.72, 577-596, (2005).

[96] Libermann, P.: Sur les prolongements des fibrés principaux et grupoides, Sem. Anal. Glob. Montréal, 7-107 (1969).

[97] Libermann, P.: Sur les prolongements des fibrés principaux et grupoides differentiables Banachiques, Presses Université de Montréal (1971).

[98] Libermann, P.: Parallélismes, J. diff. Geom., 8, 511-539, (1973).

[99] Libermann, P.: Remarques sur les systèmes différentiels, Cahiers de Top. et Geom. Diff. 23, 1, 55-72 (1982).

[100] Lubliner J.: Plasticity theory, Macmillan Publishing C., New York (1990).

[101] Luenberger D.: Optimization by vector space methods, John Wyley & sons, Inc., New York (1968).

[102] Mandel J.: Plasticité Classique et Viscoplasticité, CISM Courses and Lectures, vol.97, Springer Verlag, Wien (1972).

[103] Mangiarotti, L., Modugno, M.: Graded Lie algebras and connections on a fibred space, Journ. Math. Pures et Appl., 83 , 111-120 (1984).

[104] Mangiarotti, L., Sardanashvily, G.: Connections in Classical and Quantum Field Theory, Word Scientific, Singapore (2000).

[105] Marlow R. S.: Global Material Constraints and their Associated Reactive Stress, Journal of Elasticity, vol.33, 203-212, (1993).

[106] Marsden J.E., Hughes T.J.R.: Mathematical Foundations of Elasticity, Prentice-Hall, Redwood City, Cal. (1983).

[107] Massa E. and Pagani E., A new look at Classical Mechanics of constrained systems, Ann. Inst. Henri Poincaré, 66, (1), 1-36 (1997).

[108] Maugin G.A., Trimarco C.: Pseudomomentum and material forces in nonlinear elasticity: variational formulations and application to brittle fracture, Acta Mechanica, 94, 1-28 (1992).

[109] Maugin G.: Material Inhomogeneities in Elasticity, Applied Mathematics and Mathematical Computation, vol.3, Boca Raton, Chapman & Hall (1993).

[110] Maugin G.A.: Material forces: concepts and applications, Appl. Mech. Rev., vol.48, 213 (1995).

[111] Maugin G.A., Epstein M.: Geometrical material structure of elastoplasticity, Int. Journal of Plasticity, 14, 1-3, 109-115 (1998).

[112] Maugin G.A.: Recent advances in M^3 mechanics (mechanics on the material manifold), Theoretical and Applied Mechanics, 28-29, 221-233 (2002).

[113] de Maupertuis P.L.M.: Essai de Cosmologie, Oeuvres, Tome I, Olms, Hildesheom, Germany (1759).

[114] de Maupertuis P.L.M.: Accord de différentes lois de la nature, Oeuvres, Tome IV, Olms, Hildesheom, Germany (1768).

[115] Michor P.W.: Graded derivations of the algebra of differential forms associated with a connection, Differential Geometry, Proceedings, Peniscola 1988, Springer Lecture Notes 1410, 249-261 (1989).

[116] Michor P.W.: Topics in Differential Geometry, available on the web at the URL http://www.mat.univie.ac.at/michor/dgbook.html, (Draft from september 30, 2007).

[117] Miehe C.: A constitutive frame of elastoplasticity at large strains based on the notion of a plastic metric, International Journal of Solids and Structures, 35, 3859-3897 (1998).

[118] Miehe C.: A formulation of finite elastoplasticity based on dual co- and contra-variant eigenvector triads normalized with respect to a plastic metric, Comp. Meth. Appl. Mech. Engrg., 159, 223-260 (1998).

[119] Milnor J.: Morse Theory, Princeton University Press, Princeton, New Jersey, USA, 1-153 (1963).

[120] Modugno M.: Formulation of analytical mechanics in general relativity, Annales de l'I.H.P., section A, tome 21, 2, 147-174 (1974).

[121] Modugno M.: Torsion and Ricci tensor for non linear connections, Diff. Geom. Appl., 1, 2, 177-192 (1991).

[122] Moreau J.J.: Fonctions convexes en dualité, Fac. des Sciences de Montpellier, Sem. de Mathématiques (1962).

[123] Moreau J.J.: Fonctionelles Convexes, Lecture Notes, Séminaire Equationes aux dérivées partielles, Collègie de France (1966).

[124] Moreau J.J.: On unilateral constraints, friction and plasticity, Centro internazionale matematico estivo (C.I.M.E.) Bressanone, 173-322 (1973).

[125] Mueller R., Maugin G.A.: On material forces and finite elements discretizations, Computational Mechanics, 29, 52-60 (2002).

[126] Nečas J.: Equations aux Dérivée Partielles, Presses de l'Université de Montréal (1965).

[127] Nelson E.: Topics in dynamics, I: Flows, Mathematical Notes, Princeton University Press, New Jersey (1969).

[128] Nicolaides R. A.: Existence, uniqueness and approximation for generalized saddle point problems, SIAM J. Numer. Anal., 19, 349-357 (1982).

[129] Noether E.A.: Invariante Variationsprobleme, Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse, 235-257 (1918); English translation: Tavel M.A., Transport Theory and Statistical Mechanics, 1, 3, 183-207 (1971).

[130] Noll W.: A mathematical theory of the mechanical behavior of continuous media, Arch. Rational Mech. Anal., 2, 197-226 (1958).

[131] Noll W.: Proof of the maximality of the the orthogonal group in the unimodular group, Arch. Rational Mech. Anal., 18, 100-102 (1965).

[132] Noll W.: Materially uniform simple bodies with inhomogeneities, Arch. Rational Mech. Anal., 27, 1-32 (1968).

[133] Noll W., Chiou S.: Geometry of Differentiable Manifolds, 1-95, posted on the website math.cmu.edu/wn0g/noll (2004).

[134] Noll W.: A Frame-Free Formulation of Elasticity, 1-16, posted on the website math.cmu.edu/wn0g/noll (2005).

[135] Noll W.: Finite-Dimensional Spaces: Algebra, Geometry, and Analysis Volume I, 1-398, posted on the website math.cmu.edu/wn0g/noll (2006).

[136] Oden J.T.: Problems with Constraints, The Finite Element Handbook, Ed. H. Kardestuncer, 9.4, 2.321-2.326, McGRAW-HILL, New York (1987).

[137] Oliva W.M.: Geometric Mechanics, Lecture Notes in Mathematics n.1798, Springer (2002). Available on line at the url: http://www.springerlink.com/home/main.mpx.

[138] Palais R.: Definition of the exterior derivative in terms of the Lie derivative, Proc. Am. Math. Soc., 5, 902-908 (1954).

[139] Podio Guidugli P., Vianello M.: Linearization of response mappings in constrained elasticity, Journal of elasticity, 34, 185-189 (1994).

[140] Panagiotopoulos P.D.: Inequality Problems in Mechanics and Applications, Birkhauser (1985).

[141] Peetre, J., Another approach to elliptic boundary problems, Comm. Pure Appl. Math. **14**, (1961) 711-731.

[142] Perzyna P.: The constitutive equations for rate sensitive plastic materials, Quart. Appl. Math., 20, 321-332 (1963).

[143] Petersen P.: Riemannian Geometry, Springer-Verlag, New York (1998).

[144] Piola G.: La meccanica dei corpi naturalmente estesi trattata col calcolo delle variazioni, Opusc. Mat. Fis. di Diversi Autori, 2, 201-236, Giusti, Milano (1833).

[145] Rice J.R.: A Path Independent Integral and the Approximate Analysis of Strain Concentration by Notches and Cracks, Journal of Applied Mechanics, 35, 379-386 (1968).

[146] Rice J.R.: Mathematical Analysis in Mechanics of Fracture, in Fracture an Advanced Treatise, vol.2-3, pp.191-311, Liebowitz H. editor, Academic Press, New York (1968).

[147] Rice J.R.: Inelastic constitutive relations for solids: an internal variable theory and its application to metal plasticity, J. Mech. Phys. Solids, 19, 433-455 (1971).

[148] Roberts J. E., Thomas J. M.: Mixed and Hybrid Methods, Handbook of Numerical Analysis, Finite Element Methods, (Part I), vol.II, ed. P.G. Ciarlet and J.L. Lions, North-Holland, Amsterdam (1989).

[149] Rockafellar R.T.: Convex Analysis, Princeton, New Jersey (1970).

[150] Romano G., Romano M.: Elastostatics of Structures with Unilateral Conditions on Stress and Displacements Fields, CISM Courses and Lectures, 288, 315-338, Ravello, Springer Verlag, New York - Wien (1985).

[151] Romano G., Rosati L.: Variational Principles in Convex, Structural Analysis, Volume in onore del Prof. Giulio Ceradini, Roma, luglio (1988).

[152] Romano G., Rosati L.: On Limit states, Plasticity and Locking, IX Congresso Naz. AIMETA, Bari, October 4-7 (1988).

[153] Romano G., Rosati L.: A Survey on Recent Advances in Convex Structural Analysis. Meccanica dei Materiali e delle Strutture, Atti del Convegno Nazionale in ricordo di Riccardo Baldacci e Michele Capurso, Roma, 25-27 ottobre (1989).

[154] Romano G., Rosati L., Marotti de Sciarra F., Bisegna P.: A Potential Theory for Monotone Multi-Valued Operators, Quart. Appl. Math. Anal., 4, 1.1, 613-631 (1993).

[155] Romano G.: New Results in Subdifferential Calculus with Applications to Convex Optimization, Appl. Math. Optim., 32, 213-234 (1995).

[156] Romano G., Marotti de Sciarra F., Diaco M.: Mixed Formulation, Localization and LBB condition, proceedings SACAM '98 University of Cape Town, South Africa (1998).

[157] Romano G., Rosati L., Diaco M.: Dual approaches to general mixed problems with applications to elasticity, preprint (1998).

[158] Romano G., Rosati L., Diaco M.: Well-Posedness of Mixed Formulations in Elasticity, ZAMM, 79, 7, 435-454 (1999).

[159] Romano G.: Theory of structural models, Part I, Elements of Linear Analysis, Università di Napoli Federico II, (2000).

[160] Romano G.: Theory of structural models, Part II, Structural models, Università di Napoli Federico II, (2000).

[161] Romano G.: On the necessity of Korn's inequality, STAMM 2000, Symposium on Trends in Applications of Mathematics to Mechanics, July 9th-14th, National University of Ireland, Galway (2000).

[162] Romano G.: Scienza delle Costruzioni, Tomo Zero, Hevelius, Benevento (2001).

[163] Romano G.: Scienza delle Costruzioni, Tomo I, Hevelius, Benevento (2001).

[164] Romano G.: Scienza delle Costruzioni, Tomo II, Hevelius, Benevento (2002).

[165] Romano G., Diaco M.: A Functional Framework for Applied Continuum Mechanics, New Trends in Mathematical Physics, World Scientific, 193-204, Singapore (2004).

[166] Romano G., Barretta R., Sellitto C.: On the evaluation of the elasto-plastic tangent stiffness. Computational Plasticity - Fundamentals and Applications- Owen D.R.J., Onate E., and Suarez B. (eds), CIMNE, Part. 2: ISBN 84-95999-79-X, 1118-1121, Barcelona (2005).

[167] Romano G., Barretta R.: On the metric theory of material behavior, Research Report, (2008).

[168] Romano G., Diaco M., Barretta R.: Basic issues in phase transition theory and fracture mechanics. Variational Formulations in Mechanics: Theory and Applications. Taroco E., de Souza Neto E.A. and Novotny A.A. (eds), CIMNE, ISBN 978-84-96736-32-0, 263-279, Barcelona (2007).

[169] Romano G., Diaco M., Barretta R.: On the theory of material inhomogeneities, Mechanics Research Communications, 33, 6, 758-763 (2006).

[170] Romano G., Sellitto C., Barretta R.: Nonlinear shell theory: a duality approach, Journal of Mechanics of Materials and Structures, 2, 7, 1207-1230 (2007).

[171] Romano G., Barretta R.: On connection curvature and torsion on fiber bundles, submitted to Differential Geometry and its Aplications (2008).

[172] Romano G., Diaco M., Barretta R.: The general law of dynamics in nonlinear manifolds and Noether's theorem. Memorie dell'Accademia di Scienze Fisiche e Matematiche. Mathematical Physics Models and Engineering Sciences, Liguori, ISBN-13978-88-207-4057-3, 439-453, Napoli (2008).

[173] Romano G., Barretta R.: Is vakonomic dynamics a variational mistake?, preprint (2008).

[174] Romano G., Barretta R., Diaco M., On the general form of the law of dynamics, submitted to Int. Journal of Non-linear Mechanics (2008).

[175] Romano G., Barretta R., Diaco M., On Continuum Dynamics, submitted to Journal of Mathematical Physics (2008).

[176] Rund H.: The Hamilton-Jacobi theory in the calculus of variations, D. van Nostrand company, London (1966).

[177] Saunders D.J.: The Geometry of Jet Bundles, London Mathematical Society Lecture Note Series. 142, Cambridge University Press (1989).

[178] Skrzypek J.J., Hetnarski R.B.: Plasticity and Creep, CRC Press, Boca Raton (1993).

[179] Simo J.C.: A Framework for Finite Strain Elastoplasticity based on Maximum Plastic Dissipation and the Multiplicative Decomposition: Continuum Formulation, Computer Methods in Applied Mechanics and Engineering, Part I, 66, 199-219 (1988).

[180] Spivak M.: A comprehensive Introduction to Differential Geometry, vol.I-V, Publish or Perish, Inc., Berkeley (1979).

[181] Synge, J.L.: Classical Dynamics, Handbuch der Physik, Band III/1, Springer, Berlin (1960).

[182] Temam R.: Navier-Stokes Equations, North-Holland, Amsterdam (1977)

[183] Tartar, L., Sur un lemme d'équivalence utilisé en Analyse Numérique, Calcolo **XXIV**, Fasc. II, (1987) 129-140.

[184] Temam R. Problèmes Mathématiques en Plasticité, Gauthier Villars, Paris (1983).

[185] Trautman A.: Invariance of Lagrangian systems, General relativity, Papers in honour of J.L. Synge, pp.85-99, Clarendon Press, Oxford (1972).

[186] Truesdell C., Noll W.: The non-linear field theories of mechanics, Handbuch der Physik, 1-591, Springer (1965).

[187] Truesdell C., Toupin R.: The Classical Field Theories, Encyclopædia of Physics, vol. III/1, 226-793, Springer-Verlag, Berlin (1960).

[188] Vainberg M.M.: Variational Methods for the Study of Nonlinear Operators, Holden-Day, INC. San Francisco (1964).

[189] Wang C.C.: On the geometric structure of simple bodies, a mathematical foundation for the theory of continuous distributions of dislocations, Arch. Rational Mech. Anal., 27, pp.33-93 (1968).

[190] Weiss P.: On the Hamilton-Jacobi Theory and Quantization of a Dynamical Continuum, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, Vol. 169, No. 936, Dec. 6, 102-119 (1938)

[191] Whittaker E.T.: A treatise on the analytical dynamics of particles and rigid bodies, 4th ed. Cambridge Univ. Press, Cambridge (1959).

[192] Yosida K.: Functional Analysis, Springer-Verlag, New York (1980).

[193] R.T. Rockafellar, Convex functions and dual extremum problems, Thesis, Harvard, 1963.

[194] R.T. Rockafellar, Convex analysis, Princeton University Press, Princeton, 1970.

[195] J.J. Moreau, Fonctionelles convexes, lecture notes, séminaire: equationes aux dérivées partielles, Collègie de France, 1966.

[196] J.J. Moreau, On unilateral constraints, friction and plasticity. New Variational Techniques in Mathematical Phisycs, C.I.M.E. , Bressanone, 171-322, 1973. Ed. Cremonese, Roma 1974.

[197] P.J. Laurent, Approximation et Optimisation, Hermann, Paris 1972.

[198] I. Ekeland, R. Temam, Analyse convexe et problèmes variationnels, Dunod, Paris-Bruxelles-Montreal, 1974.

[199] L. McLinden, Dual operations on saddle functions, Trans. Amer. Math. Soc. 179 (1973) 363-381.

[200] F.H. Clarke, Necessary conditions for nonsmooth problems in optimal control and the calculus of variations, Thesis, University of Washington, Seattle, 1973.

[201] F.H. Clarke, Optimization and Nonsmooth Analysis, J. Wiley & Sons, New York, 1983.

[202] R.T. Rockafellar, Directionally Lipschitzian functions and subdifferential calculus, Proc. London Math. Soc. 3 (39) (1979) 331-355.

[203] R.T. Rockafellar, Generalized directional derivatives and subgradients of nonconvex functions, Canad. J. Math. XXXII (2) (1980) 257-280.

[204] R.T. Rockafellar, The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions., Heldermann Verlag, Berlin, 1981.

[205] P.D. Panagiotopoulos, Inequality Problems in Mechanics and Applications., Birkhäuser, Boston, 1985.

[206] A.D. Ioffe, V.M. Tihomirov, The theory of extremal problems, Nauka, Moscow (English translation, North-Holland, Amsterdam), 1979.

[207] M. Slater, Lagrange multipliers revisited: a contribution to non-linear programming., Cowles Commission Discussion Paper, Math. 403 (1950).

[208] G. Romano, L. Rosati, F. Marotti de Sciarra, An internal variable theory of inelastic behaviour derived from the uniaxial rigid-perfectly plastic law, Internat. J. Engrg. Sci. 31 (8) (1992) 1105-1120.

[209] G. Romano, L. Rosati, F. Marotti de Sciarra, Variational principles for a class of finite-step elasto-plastic problems with non-linear mixed hardening, Comput. Methods Appl. Mech. Engrg. 109 (1993) 293-314.

[210] G. Romano, L. Rosati, F. Marotti de Sciarra, A variational theory for finite-step elasto-plastic problems, Internat. J. Solids 30 (17) (1993) 2317-2334.

Index