2

Geometric Basis of Human–Like Biomechanics

In this Chapter we develop the geometric structure of modern biomechanics.

2.1 Biomechanical Manifold M

The core of geometrodynamics is the concept of the *manifold*, the stage where our *covariant force law*, $F_i = mg_{ij}a^j$, works. To get some dynamical feeling before we dive into more serious geometry, let us consider a simple 3DOF biomechanical system (e.g., a representative point of the center of mass of the human body) determined by three *generalized coordinates* $q^i = \{q^1, q^2, q^3\}$. There is a unique way to represent this system as a 3D manifold, such that to each point of the manifold there corresponds a definite configuration of the biomechanical system with coordinates q^i; therefore, we have a geometric representation of the configurations of our biomechanical system. For this reason, the manifold is called the *configuration manifold*. If the biomechanical system moves in any way, its coordinates are given as the functions of the time. Thus, the motion is given by equations of the form: $q^i = q^i(t)$. As t varies we observe that the system's *representative point* in the configuration manifold describes a *curve* and $q^i = q^i(t)$ are the equations of this curve.

On the other hand, a *topological manifold* is a separable Hausdorff space M which is locally homeomorphic to $\mathbb{R}^n$ (see, e.g., [Tho79, Hir76, Hel01, Lee00, Lee02]). So, a topological manifold has the following properties:

1. *M is a Hausdorff space*: For every pair of points $m_1, m_2 \in M$, there are disjoint open subsets $U, V \subset M$ such that $m_1 \in U$ and $m_2 \in V$.
2. *M is second countable*: There exists a countable basis for the topology of M.
3. *M is locally Euclidean of dimension n*: Every point of M has a neighborhood that is homeomorphic to an open subset of $\mathbb{R}^n$.

This further implies that for any point $m \in M$ there is a homeomorphism $\phi : U \to \phi(U) \subseteq \mathbb{R}^n$, where U is an open neighborhood of m in M and $\phi(U)$

is an open subset in $\mathbb{R}^n$. The pair (U, ϕ) is called a *coordinate chart* at a point $m \in M$.

2.1.1 Definition of the Manifold M

Given a chart (U, ϕ), we call the set U a *coordinate domain*, or a coordinate neighborhood of each of its points. If in addition $\phi(U)$ is an open ball in $\mathbb{R}^n$, then U is called a *coordinate ball*. The map ϕ is called a (*local*) *coordinate map*, and the component functions $\left(x^1, ..., x^n\right)$ of ϕ, defined by $\phi(m) = \left(x^1(m), ..., x^n(m)\right)$, are called *local coordinates* on U.

Two charts (U_1, ϕ_1) and (U_2, ϕ_2) such that $U_1 \cap U_2 \neq \varnothing$ are called *compatible* if $\phi_1(U_1 \cap U_2)$ and $\phi_2(U_2 \cap U_1)$ are open subsets of $\mathbb{R}^n$. A family $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ of compatible charts on M such that the U_α form a *cover* of M is called an *atlas*. The maps $\phi_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_{\alpha\beta}) \to \phi_\beta(U_{\alpha\beta})$ are called the *chart changings*, or *transition maps*, for the atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A}$, where $U_{\alpha\beta} = U_\alpha \cap U_\beta$, so that we have a commutative triangle:

$$\begin{tikzcd} & U_{\alpha\beta} \subseteq M \arrow[dl, "\phi_\alpha"'] \arrow[dr, "\phi_\beta"] & \\ \phi_\alpha(U_{\alpha\beta}) \arrow[rr, "\phi_{\alpha\beta}"'] & & \phi_\beta(U_{\alpha\beta}) \end{tikzcd}$$

An atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ for a manifold M is said to be a C^k-*atlas*, if all transition maps $\phi_{\alpha\beta} : \phi_\alpha(U_{\alpha\beta}) \to \phi_\beta(U_{\alpha\beta})$ are differentiable of class C^k. Two C^k atlases are called C^k-*equivalent*, if their union is again a C^k-atlas for M. An equivalence class of C^k-atlases is called a C^k-*structure* on M. In other words, a smooth structure on M is a *maximal* smooth atlas on M, i.e., such an atlas that is not contained in any strictly larger smooth atlas. By a C^k-*manifold* M, we mean a topological manifold together with a C^k-structure and a chart on M will be a chart belonging to some atlas of the C^k-structure. Smooth manifold means $C^\infty-$manifold, and the word '*smooth*' is used synonymously for C^∞. However, for most of our biomechanical needs, the weaker requirement, C^k would be sufficient. In case of any doubt, we can simply replace C^k with C^∞.

Sometimes the terms 'local coordinate system' or 'parametrization' are used instead of charts. That M is not defined with any particular atlas, but with an equivalence class of atlases, is a mathematical formulation of the *general covariance* principle. Every suitable coordinate system is equally good. A Euclidean chart may well suffice for an open subset of $\mathbb{R}^n$, but this coordinate system is not to be preferred to the others, which may require many charts (as with polar coordinates), but are more convenient in other respects.

For example, the atlas of a $n-$sphere S^n has two charts. If $N = (1, 0, ..., 0)$ and $S = (-1, ..., 0, 0)$ are the north and south poles of S^n respectively, then the two charts are given by the stereographic projections from N and S:

$$\phi_1 : S^n \backslash \{N\} \to \mathbb{R}^n, \phi_1(x^1, ..., x^{n+1}) = (x^2/(1-x^1), \dots, x^{n+1}/(1-x^1)),$$

and

$$\phi_2 : S^n \backslash \{S\} \to \mathbb{R}^n, \phi_2(x^1, ..., x^{n+1}) = (x^2/(1+x^1), \dots, x^{n+1}/(1+x^1)),$$

and the overlap map $\phi_2 \circ \phi_1^{-1} : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}^n \backslash \{0\}$ is given by the diffeomorphism $(\phi_2 \circ \phi_1^{-1})(z) = z/||z||^2$, for z in $\mathbb{R}^n \backslash \{0\}$, from $\mathbb{R}^n \backslash \{0\}$ to itself.

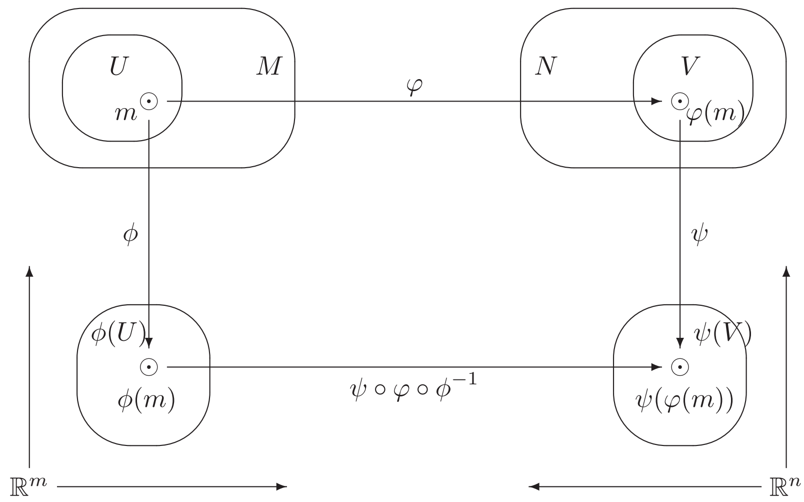
Various *additional structures* can be imposed on $\mathbb{R}^n$, and the corresponding manifold M will inherit them through its covering by charts. For example, if a covering by charts takes their values in a *Banach space* E, then E is called the *model space* and M is referred to as a C^k-*Banach manifold* modelled on E. Similarly, if a covering by charts takes their values in a *Hilbert space* $\mathcal{H}$, then $\mathcal{H}$ is called the *model space* and M is referred to as a C^k-*Hilbert manifold* modelled on $\mathcal{H}$. If not otherwise specified, we will consider M to be an Euclidean manifold, with its covering by charts taking their values in $\mathbb{R}^n$.

For a Hausdorff C^k-manifold the following properties are equivalent [KMS93]:

1. It is paracompact.
2. It is metrizable.
3. It admits a Riemannian metric.
4. Each connected component is separable.

2.1.2 Smooth Maps Between Manifolds

A map $\varphi : M \to N$ between two manifolds M and N, with $M \ni m \mapsto \varphi(m) \in N$, is called a *smooth map*, or C^k-map, if we have the following charting:



This means that for each $m \in M$ and each chart (V, ψ) on N with $\varphi(m) \in V$ there is a chart (U, ϕ) on M with $m \in U, \varphi(U) \subseteq V$, and $\Phi = \psi \circ \varphi \circ \phi^{-1}$ is C^k, that is, the following diagram commutes:

$$\begin{tikzcd} M \supseteq U \arrow[r, "\varphi"] \arrow[d, "\phi"'] & V \subseteq N \arrow[d, "\psi"] \\ \phi(U) \arrow[r, "\Phi"'] & \psi(V) \end{tikzcd}$$

Let M and N be smooth manifolds and let $\varphi : M \to N$ be a smooth map. The map φ is called a *covering*, or equivalently, M is said to *cover* N, if φ is surjective and each point $n \in N$ admits an open neighborhood V such that $\varphi^{-1}(V)$ is a union of disjoint open sets, each diffeomorphic via φ to V.

A C^k-map $\varphi : M \to N$ is called a C^k-*diffeomorphism* if φ is a bijection, $\varphi^{-1} : N \to M$ exists and is also C^k. Two manifolds are called diffeomorphic if there exists a diffeomorphism between them.

All smooth manifolds and smooth maps between them form the category $\mathcal{M}$.

The most important examples of biomechanical manifolds have also an additional group structure and thus belong to the *category of Lie groups* $\mathcal{G}$.

2.2 Biomechanical Bundles

In this section we introduce secondary concepts of *biomechanical bundles*, derived from the primary concept of the manifold.

2.2.1 The Tangent Bundle of the Manifold *M*

Recall that if $[a, b]$ is a closed interval, a C^0-map $\gamma : [a, b] \to M$ is said to be *differentiable* at the endpoint a if there is a chart (U, ϕ) at $\gamma(a)$ such that the following limit exists and is finite [AMR88]:

$$\frac{d}{dt}(\phi \circ \gamma)(a) \equiv (\phi \circ \gamma)'(a) = \lim_{t \to a} \frac{(\phi \circ \gamma)(t) - (\phi \circ \gamma)(a)}{t - a}. \tag{2.1}$$

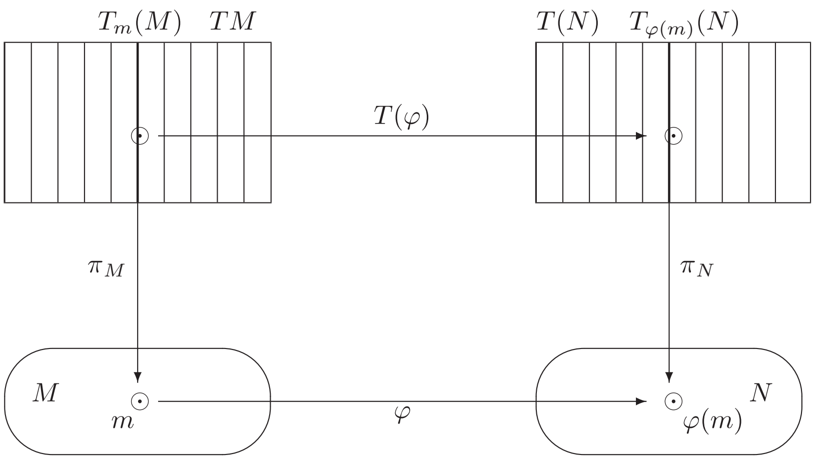
Generalizing (2.1), we get the notion of the *curve on a manifold*. For a smooth manifold M and a point $m \in M$ a curve at m is a C^0-map $\gamma : I \to M$ from an interval $I \subset \mathbb{R}$ into M with $0 \in I$ and $\gamma(0) = m$.

Two curves γ_1 and γ_2 passing though a point $m \in U$ are *tangent at* m with respect to the chart (U, ϕ) if $(\phi \circ \gamma_1)'(0) = (\phi \circ \gamma_2)'(0)$. Thus, two curves are tangent if they have identical tangent vectors (same direction and speed) in a local chart on a manifold.

For a smooth manifold M and a point $m \in M$, the *tangent space* $T_m M$ to M at m is the *set of equivalence classes* of curves at m:

$$T_m M = \{[\gamma]_m : \gamma \text{ is a curve at a point } m \in M\}.$$

A C^k-map $\varphi : M \ni m \mapsto \varphi(m) \in N$ between two manifolds M and N induces a linear map $T_m \varphi : T_m M \to T_{\varphi(m)} N$ for each point $m \in M$, called a *tangent map*, if we have:



i.e., the following diagram commutes:

$$\begin{tikzcd}
T_mM \arrow[r, "T_m\varphi"] \arrow[d, "\pi_M"'] & T_{\varphi(m)}N \arrow[d, "\pi_N"] \\
M \ni m \arrow[r, "\varphi"'] & \varphi(m) \in N
\end{tikzcd}$$

with the *natural projection*, or *tangent bundle projection*, $\pi_M : TM \to M$, given by $\pi_M\,(T_mM) = m$, that takes a tangent vector v to the point $m \in M$ at which the vector v is attached i.e., $v \in T_mM$.

For a smooth manifold M of dimension n, its *tangent bundle* TM is the disjoint union of all its tangent spaces T_mM at all points $m \in M$, i.e., $TM = \bigsqcup_{m \in M} T_mM$.

If M is an $n-$manifold, then TM is a $2n-$manifold. To define the smooth structure on TM, we need to specify how to construct local coordinates on TM. To do this, let $\left(x^1(m), ..., x^n(m)\right)$ be local coordinates of a point m on M and let $\left(v^1(m), ..., v^n(m)\right)$ be components of a tangent vector in this coordinate system. Then the $2n$ numbers $\left(x^1(m), ..., x^n(m),\, v^1(m), ..., v^n(m)\right)$ give a local coordinate system on TM. This is the basic idea one uses to prove that indeed TM is a $2n-$manifold [MR99].

$TM = \bigsqcup_{m \in M} T_mM$ defines a family of vector spaces parameterized by M. The inverse image $\pi_M^{-1}(m)$ of a point $m \in M$ under the natural projection π_M is the tangent space T_mM. This space is called the *fibre* of the tangent bundle over the point $m \in M$ [Sti51].

A C^k-map $\varphi : M \to N$ between two manifolds M and N induces a linear *tangent map* $T\varphi : TM \to TN$ between their tangent bundles, i.e., the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "T\varphi"] \arrow[d, "\pi_M"'] & TN \arrow[d, "\pi_N"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

All tangent bundles and their tangent maps form the category $\mathcal{TB}$. The category $\mathcal{TB}$ is the natural framework for *Lagrangian biomechanics*.

Now, we can formulate the *global version of the chain rule*. If $\varphi : M \to N$ and $\psi : N \to P$ are two smooth maps, then we have $T(\psi \circ \varphi) = T\psi \circ T\varphi$ (see [KMS93]). In other words, we have a functor $T : \mathcal{M} \Rightarrow \mathcal{TB}$ from the category $\mathcal{M}$ of smooth manifolds to the category $\mathcal{TB}$ of their tangent bundles:

$$\begin{tikzcd} & M \arrow[dl, "\varphi"'] \arrow[dr, "(\psi\circ\varphi)"] & \\ N \arrow[rr, "\psi"'] & & P \end{tikzcd} \quad \overset{T}{\Longrightarrow} \quad \begin{tikzcd} & TM \arrow[dl, "T\varphi"'] \arrow[dr, "T(\psi\circ\varphi)"] & \\ TN \arrow[rr, "T\psi"'] & & TP \end{tikzcd}$$

2.2.2 The Cotangent Bundle of the Manifold M

The *dual* notion to the tangent space T_mM to a smooth manifold M at a point m is its *cotangent space* T^*_mM at the same point m. Similarly to the tangent bundle, for a smooth manifold M of dimension n, its *cotangent bundle* T^*M is the disjoint union of all its cotangent spaces T^*_mM at all points $m \in M$, i.e., $T^*M = \bigsqcup_{m\in M} T^*_mM$. Therefore, the cotangent bundle of an $n-$manifold M is the vector bundle $T^*M = (TM)^*$, the (real) dual of the tangent bundle TM.

If M is an $n-$manifold, then T^*M is a $2n-$manifold. To define the smooth structure on T^*M, we need to specify how to construct local coordinates on T^*M. To do this, let $\left(x^1(m), ..., x^n(m)\right)$ be local coordinates of a point m on M and let $(p_1(m), ..., p_n(m))$ be components of a covector in this coordinate system. Then the $2n$ numbers $\left(x^1(m), ..., x^n(m),\, p_1(m), ..., p_n(m)\right)$ give a local coordinate system on T^*M. This is the basic idea one uses to prove that indeed T^*M is a $2n-$manifold.

$T^*M = \bigsqcup_{m\in M} T^*_mM$ defines a family of vector spaces parameterized by M, with the *conatural projection*, or *cotangent bundle projection*, $\pi^*_M : T^*M \to M$, given by $\pi^*_M\left(T^*_mM\right) = m$, that takes a covector p to the point $m \in M$ at which the covector p is attached i.e., $p \in T^*_mM$. The inverse image $\pi^{-1}_M(m)$ of a point $m \in M$ under the conatural projection π^*_M is the cotangent space T^*_mM. This space is called the *fibre* of the cotangent bundle over the point $m \in M$.

In a similar way, a C^k-map $\varphi : M \to N$ between two manifolds M and N induces a linear *cotangent map* $T^*\varphi : T^*M \to T^*N$ between their cotangent bundles, i.e., the following diagram commutes:

$$\begin{tikzcd} T^*M \arrow[r, "T^*\varphi"] \arrow[d, "\pi^*_M"'] & T^*N \arrow[d, "\pi^*_N"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

All cotangent bundles and their cotangent maps form the category $\mathcal{T}^*\mathcal{B}$. The category $\mathcal{T}^*\mathcal{B}$ is the natural stage for *Hamiltonian biomechanics*.

Now, we can formulate the *dual version of the global chain rule.* If $\varphi : M \to N$ and $\psi : N \to P$ are two smooth maps, then we have $T^*(\psi \circ \varphi) = T^*\psi \circ T^*\varphi$. In other words, we have a cofunctor $T^* : \mathcal{M} \Rightarrow \mathcal{T}^*\mathcal{B}$ from the category $\mathcal{M}$ of smooth manifolds to the category $\mathcal{T}^*\mathcal{B}$ of their cotangent bundles:

$$\begin{tikzcd} & M \arrow[dl, "\varphi"'] \arrow[dr, "(\psi \circ \varphi)"] & \\ N \arrow[rr, "\psi"'] & & P \end{tikzcd} \quad \overset{T^*}{\Longrightarrow} \quad \begin{tikzcd} & T^*M & \\ T^*N \arrow[ur, "T^*\varphi"] & & T^*P \arrow[ul, "T^*(\psi \circ \varphi)"'] \arrow[ll, "T^*\psi"] \end{tikzcd}$$

2.2.3 Fibre Bundles

Vector Bundles

Both the tangent bundle (TM, π_M, M) and the cotangent bundle (T^*M, π^*_M, M) are examples of a more general notion of *vector bundle* (E, π, M) of a manifold M, which consists of manifolds E (the total space) and M (the base), as well as a smooth map $\pi : E \to M$ (the projection) together with an equivalence class of vector bundle atlases (in this section we follow [KMS93]). A vector bundle atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ for (E, π, M) is a set of pairwise compatible vector bundle charts (U_α, ϕ_α) such that $(U_\alpha)_{\alpha \in A}$ is an open cover of M. Two vector bundle atlases are called equivalent, if their union is again a vector bundle atlas.

On each fibre $E_m = \pi^{-1}(m)$ corresponding to the point $m \in M$ there is a unique structure of a real vector space, induced from any vector bundle chart (U_α, ϕ_α) with $m \in U_\alpha$. A section u of (E, π, M) is a smooth map $u : M \to E$ with $\pi \circ u = Id_M$.

Let (E, π_M, M) and (F, π_N, N) be vector bundles. A *vector bundle homomorphism* $\Phi : E \to F$ is a fibre respecting, fibre linear smooth map induced by the smooth map $\varphi : M \to N$ between the base manifolds M and N, i.e., the following diagram commutes:

$$\begin{tikzcd} E \arrow[r, "\Phi"] \arrow[d, "\pi_M"'] & F \arrow[d, "\pi_N"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

We say that Φ *covers* φ. If Φ is invertible, it is called a *vector bundle isomorphism*.

All smooth vector bundles together with their homomorphisms form a category $\mathcal{VB}$.

If (E, π, M) is a vector bundle which admits a vector bundle atlas $(U_\alpha, \phi_\alpha)_{\alpha \in A}$ with the given open cover, then, we have $\phi_\alpha \circ \phi_\beta^{-1}(m, v) = \left(m, \phi_{\alpha\beta}(m)v\right)$ for C^k-transition functions $\phi_{\alpha\beta} : U_{\alpha\beta} = U_\alpha \cap U_\beta \to GL(V)$ (where we have fixed a standard fibre V). This family of transition maps satisfies the *cocycle condition*

$$\begin{cases} \phi_{\alpha\beta}(m) \cdot \phi_{\beta\gamma}(m) = \phi_{\alpha\gamma}(m) \quad \text{for each } m \in U_{\alpha\beta\gamma} = U_\alpha \cap U_\beta \cap U_\gamma, \\ \qquad\qquad\qquad \phi_{\alpha\alpha}(m) = e \quad \text{for all } m \in U_\alpha. \end{cases}$$

The family $(\phi_{\alpha\beta})$ is called the *cocycle* of transition maps for the vector bundle atlas (U_α, ϕ_α).

Now, let us suppose that the same vector bundle (E, π, M) is described by an equivalent vector bundle atlas $(U_\alpha, \psi_\alpha)_{\alpha \in A}$ with the same open cover (U_α). Then the vector bundle charts (U_α, ϕ_α) and (U_α, ψ_α) are compatible for each α, so $\psi_\alpha \circ \phi_\beta^{-1}(m, v) = (m, \tau_\alpha(m)v)$ for some $\tau_\alpha : U_\alpha \to GL(V)$. We get

$$\tau_\alpha(m)\, \phi_{\alpha\beta}(m) = \phi_{\alpha\beta}(m)\, \tau_\beta(m) \quad \text{for all } m \in U_{\alpha\beta},$$

and we say that the two cocycles $(\phi_{\alpha\beta})$ and $(\psi_{\alpha\beta})$ of transition maps over the cover (U_α) are *cohomologous*. If $GL(V)$ is an Abelian group, i.e., if the standard fibre V is of real or complex dimension 1, then the cohomology classes of cocycles $(\phi_{\alpha\beta})$ over the open cover (U_α) form a usual cohomology group $H^1(M, GL(V))$ with coefficients in the sheaf $GL(V)$ [KMS93].

Let (E, π, M) be a vector bundle and let $\varphi : N \to M$ be a smooth map between the base manifolds N and M. Then there exists the *pull–back vector bundle* $(\varphi^* E, \varphi^* \pi, \varphi^* N)$ with the same typical fibre and a vector bundle homomorphism, given by the commutative diagram [KMS93]:

$$\begin{tikzcd} \varphi^* E \arrow[r, "\pi^* \varphi"] \arrow[d, "\varphi^* \pi"'] & E \arrow[d, "\pi"] \\ N \arrow[r, "\varphi"'] & M \end{tikzcd}$$

The vector bundle $(\varphi^* E, \varphi^* \pi, \varphi^* N)$ is constructed as follows. Let $E = VB(\phi_{\alpha\beta})$ denote that E is described by a cocycle $(\phi_{\alpha\beta})$ of transition maps over an open cover (U_α) of M. Then $(\phi_{\alpha\beta} \circ \varphi)$ is a cocycle of transition maps over the open cover $\big(\varphi^{-1}(U_\alpha)\big)$ of N and the bundle is given by $\varphi^* E = VB(\phi_{\alpha\beta} \circ \varphi)$.

The Second Vector Bundle of the Manifold M

Let (E, π, M) be a vector bundle over the biomechanical manifold M with fibre addition $+_E : E \times_M E \to E$ and fibre scalar multiplication $m_t^E : E \to E$. Then (TE, π_E, E), the tangent bundle of the manifold E, is itself a vector bundle, with fibre addition denoted by $+_{TE}$ and scalar multiplication denoted by m_t^{TE}. The *second vector bundle* structure on $(TE, T\pi, TM)$, is the 'derivative' of the original one on (E, π, M). In particular, the space $\{\Xi \in TE : T\pi.\Xi = 0 \in TM\} = (Tp)^{-1}(0)$ is denoted by VE and is called the *vertical bundle* over E. Its main characteristics are *vertical lift* and *vertical projection* (see [KMS93] for details).

All of this is valid for the *second tangent bundle* $T^2M = TTM$ of a manifold, but here we have one more natural structure at our disposal. The *canonical flip* or *involution* $\kappa_M : T^2M \to T^2M$ is defined locally by

$$(T^2\phi \circ \kappa_M \circ T^2\phi^{-1})(x, \xi; \eta, \zeta) = (x, \eta; \xi, \zeta).$$

where (U, ϕ) is a local chart on M (this definition is invariant under changes of charts). The flip κ_M has the following properties (see [KMS93]):

1. $\kappa_M \circ T^2 f = T^2 f \circ \kappa_M$ for each $f \in C^k(M, N)$;
2. $T(\pi_M) \circ \kappa_M = \pi_{TM}$;
3. $\pi_{TM} \circ \kappa_M = T(\pi_M)$;
4. $\kappa_M^{-1} = \kappa_M$;
5. κ_M is a linear isomorphism from the bundle $(TTM, T(\pi_M), TM)$ to (TTM, π_{TM}, TM), so it interchanges the two vector bundle structures on TTM;
6. κ_M is the unique smooth map $TTM \to TTM$ which, for each $\gamma : \mathbb{R} \to M$, satisfies
$$\partial_t \partial_s \gamma(t, s) = \kappa_M \partial_t \partial_s \gamma(t, s).$$

In a similar way the second cotangent bundle of a manifold M can be defined. Even more, for every manifold there is a *geometric isomorphism* between the bundles $TT^*M = T(T^*M)$ and $T^*TM = T^*(TM)$ [MS78].

General Fibre Bundles

A vector bundle is a special case of a more general structure, a *fibre bundle* [Sti51], a topological construction which itself is a class of *fibrations*.

Let $I = [0, 1]$. A map $p : E \to B$ is said to have the *homotopy lifting property* (HLP) with respect to a topological space X if for every map $f : X \to E$ and homotopy $G : X \times I \to B$ of $p \circ f$ there is a homotopy $F : X \times I \to E$ with $f = F_0$ and $p \circ F = G$. F is said to be a *lifting* of G. p is called a *fibration* if it has the HLP for all spaces X and a weak fibration if it has the HLP for all disks D^n, $n \geq 0$. If $b_0 \in B$ is the base point, then the space $F = p^{-1}(b_0)$ is called the *fibre* of p. The projection onto the first factor, $p_B : B \times F \to B$, is clearly a fibration and is called the *trivial fibration* over B with fibre F [Swi75].

A *fibre bundle* is a quadruple (B, p, E, F) where the space B is called the *base space*, E is the *total space*, and the vector spaces $F = p^{-1}(b)$ are the fibers. Here the projection $p : E \to B$ is such a map that B has an open covering $\{U_\alpha\}_{\alpha \in A}$, and for each $\alpha \in A$ there is a homeomorphism $\phi_\alpha : U_\alpha \times F \to p^{-1}U_\alpha$ such that $p \circ \phi_\alpha = p_{U_\alpha} : U_\alpha \times F \to U_\alpha$ [Sti51]. In other words, locally $p : E \to B$ looks like a trivial fibration. If B is paracompact, one can show that $p : E \to B$ is a fibration. If (B, p, E, F) is a fibre bundle, then $p : E \to B$ is a weak fibration [Swi75]. A map $p : E \to B$ has a *local cross–section* at a point $x \in B$ if there is a neighborhood U of x in B and a map $\lambda : U \to E$ with $p \circ \lambda = 1_U$.

A fibre bundle (B, p, E, F) with F discrete is called a covering of B. p is called a a *covering projection* and E a covering space over B. For example, the $n-$torus T^n is the $n-$fold Cartesian product $S^1 \times S^1 \times ... \times S^1$. The map $p : \mathbb{R}^n \to T^n$ defined by $p(r_1, r_2, ..., r_n) = \left(e^{2\pi i r_1}, e^{2\pi i r_2}, ..., e^{2\pi i r_n}\right)$ is a covering projection. The fibre is the set of integer lattice points in $\mathbb{R}^n$. Since $\mathbb{R}^n$ is contractible, it follows that its kth homotopy group $\pi_k(T^n) = 0$ for $k \geq 2$ [Swi75].

All smooth fibre bundles together with their homomorphisms form a category $\mathcal{FB}$.

Our vector bundle defined above represents an important class of fibre bundles for which every fibre has the structure of a vector space in a way which is compatible on neighboring fibres. Let F denote $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ – the real, complex or quaternionic numbers. An nD $F-$vector bundle is a fibre bundle $\xi = (B, p, E, F^n)$ in which each fibre $p^{-1}(b)$, $b \in B$, has the structure of a vector space over F such that there is an open covering $\{U_\alpha : \alpha \in A\}$ of B and for each $\alpha \in A$ a homeomorphism $\phi_\alpha : U_\alpha \times F^n \to p^{-1}U_\alpha$ with $p \circ \phi_\alpha = p_{U_\alpha}$ and $(\phi_\alpha | \{b\} \times F^n) : \{b\} \times F^n \to p^{-1}(b)$ a vector space isomorphism for each $b \in U_\alpha$. We speak of real, complex or quaternionic vector bundles according to whether $F = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ [Swi75].

For example, for any space B the trivial nD $F-$vector bundle is $(B, p_B, B \times F^n, F^n)$.

If we let E be the quotient space of $I \times \mathbb{R}$ under the identifications $(0, t) \sim (1, -t)$, then the projection $I \times \mathbb{R} \to I$ induces a map $p : E \to S^1$ which is a 1D vector bundle, or *line bundle*. Since E is homeomorphic to a Mőbius band with its boundary circle deleted, we call this bundle the *Mőbius bundle* [Hat02].

For any $n \geq 1$ the *tangent bundle* TS^n of the unit $n-$sphere $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$ is the fibre bundle $(S^n, p, E, \mathbb{R}^n)$, where $E = \{(x, y) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} : \|x\| = 1,\, x \cdot y = 0\}$ and $p : E \to S^n$ is defined by $p(x, y) = x$.

For any $n \geq 1$ the *normal bundle* NS^n of the $n-$sphere S^n is the fibre bundle $(S^n, p', E', \mathbb{R}^1)$, where $E' = \{(x, y) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} : \|x\| = 1, y = \lambda x, \lambda \in \mathbb{R}^1\}$ and $p' : E' \to S^n$ is defined by $p'(x, y) = x$ [Swi75].

The only two vector bundles with base space B a circle and 1D fibre F are the Mőbius band and the annulus, but the classification of all the different vector bundles over a given base space with fibre of a given dimension is quite difficult in general. For example, when the base space is a high–dimensional sphere and the dimension of the fibre is at least three, then the classification is of the same order of difficulty as the fundamental but still largely unsolved problem of computing the homotopy groups of spheres [Hat02].

Now, there is a natural direct sum operation for vector bundles over a fixed base space X, which in each fibre reduces just to direct sum of vector spaces. Using this, one can obtain a weaker notion of isomorphism of vector bundles by defining two vector bundles over the same base space X to be stably isomorphic if they become isomorphic after direct sum with product vector bundles $X \times \mathbb{R}^n$ for some n, perhaps different n's for the two given vector bundles. Then it turns out that the set of stable isomorphism classes of vector bundles over X forms an Abelian group under the direct sum operation, at least if X is compact Hausdorff. The traditional notation for this group is $\widetilde{KO}(X)$. In the case of spheres the groups $\widetilde{KO}(S^n)$ have the quite unexpected property of being periodic in n. This is called *Bott periodicity*, and the values of $\widetilde{KO}(S^n)$ are given by the following table [Hat02]:

$n \bmod 8$	1	2	3	4	5	6	7	8
$\widetilde{KO}(S^n)$	$\mathbb{Z}_2$	$\mathbb{Z}_2$	0	$\mathbb{Z}$	0	0	0	$\mathbb{Z}$

For example, $\widetilde{KO}(S^1)$ is $\mathbb{Z}_2$, a cyclic group of order two, and a generator for this group is the Mőbius bundle. This has order two since the direct sum of two copies of the Mőbius bundle is the product $S^1 \times \mathbb{R}^1$, as one can see by embedding two Mőbius bands in a solid torus so that they intersect orthogonally along the common core circle of both bands, which is also the core circle of the solid torus.

The complex version of $\widetilde{KO}(X)$, called $\widetilde{K}(X)$, is constructed in the same way as $\widetilde{KO}(X)$ but using vector bundles whose fibers are vector spaces over $\mathbb{C}$ rather than $\mathbb{R}$. The complex form of Bott Periodicity asserts simply that $\widetilde{K}(S^n)$ is $\mathbb{Z}$ for n even and 0 for n odd, so the period is two rather than eight.

The groups $\widetilde{K}(X)$ and $\widetilde{KO}(X)$ for varying X share certain formal properties with the cohomology groups studied in classical algebraic topology. Using a more general form of Bott periodicity, it is in fact possible to extend the groups $\widetilde{K}(X)$ and $\widetilde{KO}(X)$ to a full cohomology theory, families of Abelian groups $\widetilde{K^n}(X)$ and $\widetilde{KO}^n(X)$ for $n \in \mathbb{Z}$ that are periodic in n of period two and eight, respectively. There is more algebraic structure here than just the additive group structure, however. Tensor products of vector spaces give rise to tensor products of vector bundles, which in turn give product operations in both real and complex K–theory similar to cup product in ordinary cohomology. Furthermore, exterior powers of vector spaces give natural operations within K–theory [Hat02].

Tensor Fields as Sections of the Vector Bundle

A *tensor–field* $\tau \in \Gamma(\mathcal{F}(M))$ of type (p,q) (see Appendix) on a smooth $n-$manifold M is a *smooth section of the vector bundle*

$$\bigotimes^{q} T^*M \otimes \bigotimes^{p} TM = TM \overbrace{\otimes ... \otimes}^{p \text{ times}} TM \otimes T^*M \overbrace{\otimes ... \otimes}^{q \text{ times}} T^*M.$$

The coefficients of the tensor–field τ are C^k functions on U, with p indices up and q indices down. The classical position of indices can be explained in modern terms as follows. If (U,ϕ) is a chart at a point $m \in M$ with local coordinates $\left(x^1,...,x^n\right)$, we have the *holonomous frame field*

$$\partial_{x^{i_1}} \otimes \partial_{x^{i_2}} \otimes ... \otimes \partial_{x^{i_p}} \otimes dx^{j_1} \otimes dx^{j_2} ... \otimes dx^{j_q},$$

for $i \in \{1,...,n\}^p$, $j = \{1,...,n\}^q$, over U of this tensor bundle, and for any $(p,q)-$tensor–field τ we have

$$\tau|U = \tau^{i_1...i_p}_{j_1...j_q} \partial_{x^{i_1}} \otimes \partial_{x^{i_2}} \otimes ... \otimes \partial_{x^{i_p}} \otimes dx^{j_1} \otimes dx^{j_2} ... \otimes dx^{j_q}.$$

For such tensor–fields the Lie derivative along any vector–field is defined, and it is a *derivation* (i.e., both linearity and Leibniz rules hold) with respect to the tensor product. This natural bundle admits many natural transformations. For example, a ‘contraction’ like the trace $T^*M \otimes TM = L(TM,TM) \to M \times \mathbb{R}$, but applied just to one specified factor of type T^*M and another one of type TM, is a natural transformation. And any ‘permutation of the same kind of factors’ is a natural transformation.

The tangent bundle $\pi_M : TM \to M$ of a manifold M is a vector bundle over M such that, given an atlas $\{(U_\alpha, \varphi_\alpha)\}$ of M, TM is provided with the *holonomic atlas*

$$\Psi = \{(U_\alpha, \varphi_\alpha = T\varphi_\alpha)\}.$$

The associated linear bundle coordinates are the induced coordinates $(\dot{x}^\lambda)$ at a point $m \in M$ with respect to the *holonomic frames* $\{\partial_\lambda\}$ in tangent spaces T_mM. Their transition functions read (see Appendix)

$$\dot{x}'^\lambda = \frac{\partial x'^\lambda}{\partial x^\mu} \dot{x}^\mu.$$

The tangent bundle TM is a fibre bundle with the structure group $GL(\dim M, \mathbb{R})$.

The cotangent bundle of M is the dual T^*M of TM. It is equipped with the induced coordinates $(\dot{x}_\lambda)$ at a point $m \in M$ with respect to *holonomic coframes* $\{dx^\lambda\}$ dual of $\{\partial_\lambda\}$. Their transition functions read

$$\dot{x}'_\lambda = \frac{\partial x'^\mu}{\partial x^\lambda}\dot{x}_\mu.$$

The tensor products

$$(\overset{m}{\otimes} TX) \otimes (\overset{k}{\otimes} T^* X)$$

of the tangent and cotangent bundles are called *tensor bundles*.

The Natural Vector Bundle

In this section we mainly follow [Mic01, KMS93].

A *vector bundle functor* or *natural vector bundle* is a functor $\mathcal{F}$ which associates a vector bundle $(\mathcal{F}(M), \pi_M, M)$ to each $n-$manifold M and a vector bundle homomorphism

$$\begin{tikzcd} \mathcal{F}(M) \arrow[r, "\mathcal{F}(\varphi)"] \arrow[d, "\pi_M"'] & \mathcal{F}(N) \arrow[d, "\pi_N"] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd}$$

to each $\varphi : M \to N$ in $\mathcal{M}$, which covers φ and is fiberwise a linear isomorphism. Two common examples of the vector bundle functor $\mathcal{F}$ are tangent bundle functor T (subsection 2.2.1) and cotangent bundle functor T^* (subsection 2.2.2).

The space of all smooth sections of the vector bundle (E, π_M, M) is denoted by $\Gamma(E, \pi_M, M)$. Clearly, it is a vector space with fiberwise addition and scalar multiplication.

Let $\mathcal{F}$ be a vector bundle functor on $\mathcal{M}$. Let M be a smooth manifold and let $X \in \mathcal{X}(M)$ be a vector–field on M. Then the flow F_t of X for fixed t, is a diffeomorphism defined on an open subset of M. The map

$$\begin{tikzcd} \mathcal{F}(M) \arrow[r, "\mathcal{F}(F_t)"] \arrow[d, "\pi_M"'] & \mathcal{F}(M) \arrow[d, "\pi_M"] \\ M \arrow[r, "F_t"'] & M \end{tikzcd}$$

is then a vector bundle isomorphism, defined over an open subset of M.

We consider a tensor–field τ (2.2.3), which is a section $\tau \in \Gamma(\mathcal{F}(M))$ of the vector bundle $(\mathcal{F}(M), \pi_M, M)$ and we define for $t \in \mathbb{R}$

$$F_t^* \tau = \mathcal{F}(F_{-t}) \circ \tau \circ F_t,$$

a local section of the bundle $\mathcal{F}(M)$. For each point $m \in M$ the value $F_t^* \tau(x) \in \mathcal{F}(M)_m$ is defined, if t is small enough (depending on x). So, in the vector space $\mathcal{F}(M)_m$ the expression $\frac{d}{dt}|_{t=0} F_t^* \tau(x)$ makes sense and therefore the section

$$\mathcal{L}_X \tau = \frac{d}{dt}|_{t=0} F_t^* \tau$$

is globally defined and is an element of $\Gamma(\mathcal{F}(M))$. It is called the *Lie derivative of the tensor–field τ along a vector–field $X \in \mathcal{X}(M)$* (see subsection 2.4.1, for details on Lie derivative).

In this situation we have:

1. $F_t^* F_r^* \tau = F_{t+r}^* \tau$, whenever defined.
2. $\frac{d}{dt} F_t^* \tau = F_t^* \mathcal{L}_X \tau = \mathcal{L}_X (F_t^* \tau)$, so $[\mathcal{L}_X, F_t^*] = \mathcal{L}_X \circ F_t^* - F_t^* \circ \mathcal{L}_X = 0$, whenever defined.
3. $F_t^* \tau = \tau$ for all relevant t iff $\mathcal{L}_X \tau = 0$.

Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two vector bundle functors on $\mathcal{M}$. Then the (fiberwise) tensor product $(\mathcal{F}_1 \otimes \mathcal{F}_2)(M) = \mathcal{F}_1(M) \otimes \mathcal{F}_2(M)$ is again a vector bundle functor and for $\tau_i \in \Gamma(\mathcal{F}_i(M))$ with $i = 1, 2$, there is a section $\tau_1 \otimes \tau_2 \in \Gamma(\mathcal{F}_1 \otimes \mathcal{F}_2)(M)$, given by the pointwise tensor product.

Also in this situation, for $X \in \mathcal{X}(M)$ we have

$$\mathcal{L}_X (\tau_1 \otimes \tau_2) = \mathcal{L}_X \tau_1 \otimes \tau_2 + \tau_1 \otimes \mathcal{L}_X \tau_2.$$

In particular, for $f \in C^k(M, \mathbb{R})$ we have $\mathcal{L}_X (f \tau) = df(X) \tau + f \mathcal{L}_X \tau$.

For any vector bundle functor $\mathcal{F}$ on $\mathcal{M}$ and $X, Y \in \mathcal{X}(M)$ we have:

$$[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X = \mathcal{L}_{[X,Y]} : \Gamma(\mathcal{F}(M)) \to \Gamma(\mathcal{F}(M)).$$

The Pull–Back and Push–Forward

In this subsection we define two important operations, following [AMR88], which will be used in the further text.

Let $\varphi : M \to N$ be a C^k map of manifolds and $f \in C^k(N, \mathbb{R})$. Define the *pull–back* of f by φ by

$$\varphi^* f = f \circ \varphi \in C^k(M, \mathbb{R}).$$

If f is a C^k diffeomorphism and $X \in \mathcal{X}^k(M)$, the *push–forward* of X by φ is defined by

$$\varphi_* X = T\varphi \circ X \circ \varphi^{-1} \in \mathcal{X}^k(N).$$

If x^i are local coordinates on M and y^j local coordinates on N, the preceding formula gives the components of $\varphi_* X$ by

$$(\varphi_* X)^j(y) = \frac{\partial \varphi^j}{\partial x^i}(x) \, X^i(x), \qquad \text{where} \qquad y = \varphi(x).$$

We can interchange pull–back and push–forward by changing φ to φ^{-1}, that is, defining φ_* (resp. φ^*) by $\varphi_* = (\varphi^{-1})^*$ (resp. $\varphi^* = (\varphi^{-1})_*$). Thus the *push–forward* of a function f on M is $\varphi_* f = f \circ \varphi^{-1}$ and the *pull–back* of a vector–field Y on N is $\varphi^* Y = (T\varphi)^{-1} \circ Y \circ \varphi$.

Notice that φ must be a diffeomorphism in order that the pull–back and push–forward operations make sense, the only exception being pull–back of functions. Thus vector–fields can only be pulled back and pushed forward by diffeomorphisms. However, even when φ is not a diffeomorphism we can talk about $\varphi-$*related vector–fields* as follows.

Let $\varphi : M \to N$ be a C^k map of manifolds. The vector–fields $X \in \mathcal{X}^{k-1}(M)$ and Y$\in$ $\mathcal{X}^{k-1}(N)$ are called $\varphi-$*related*, denoted $X \sim_\varphi Y$, if $T\varphi \circ X = Y \circ \varphi$.

Note that if φ is diffeomorphism and X and Y are $\varphi-$related, then $Y = \varphi_* X$. In general however, X can be $\varphi-$related to more than one vector–field on N. $\varphi-$relatedness means that the following diagram commutes:

$$\begin{tikzcd} TM \arrow[r, "T\varphi"] & TN \\ M \arrow[u, "X"] \arrow[r, "\varphi"'] & N \arrow[u, "Y"'] \end{tikzcd}$$

The behavior of flows under these operations is as follows: Let $\varphi : M \to N$ be a C^k-map of manifolds, $X \in \mathcal{X}^k(M)$ and $Y \in \mathcal{X}^k(N)$. Let F_t and G_t denote the flows of X and Y respectively. Then $X \sim_\varphi Y$ iff $\varphi \circ F_t = G_t \circ \varphi$. In particular, if φ is a diffeomorphism, then the equality $Y = \varphi_* X$ holds iff the flow of Y is $\varphi \circ F_t \circ \varphi^{-1}$ (This is called the push–forward of F_t by φ since it is the natural way to construct a diffeomorphism on N out of one on M). In particular, $(F_t)_* X = X$. Therefore, the flow of the push–forward of a vector–field is the push–forward of its flow.

2.2.4 Jet Bundles

Roughly speaking, two maps $f, g : M \to N$ are said to determine the same $r-$jet at $x \in M$, if they have the rth order contact at x [KMS93]. To make this idea precise, we first define the rth order contact of two curves on a manifold. We recall that a smooth function $\mathbb{R} \to \mathbb{R}$ is said to vanish to rth order at a point, if all its derivatives up to order r vanish at this point. Two curves $\gamma, \delta : \mathbb{R} \to M$ have the rth contact at zero, if for every smooth function φ on M the difference $\varphi \circ \gamma - \varphi \circ \delta$ vanishes to rth order at $0 \in \mathbb{R}$. In this case we write $\gamma \sim_r \delta$. Clearly, $\sim_r$is an equivalence relation. For $r = 0$ this relation means $\gamma(0) = \delta(0)$. If $\gamma \sim_r \delta$, then $f \circ \gamma \sim_r f \circ \delta$ for every map $f : M \to N$.

Two maps $f, g : M \to N$ are said to determine the same $r$jet at $x \in M$, if for every curve $\gamma : \mathbb{R} \to M$ with $\gamma(0) = x$ the curves f and g have the rth order contact at zero. In such a case we write $j^r_x f = j^r_x g$ or $j^r f(x) = j^r g(x)$.

An equivalence class of this relation is called an $r$jet of M into N. The set of all $r$jets of M into N is denoted by $J^r(M,N)$. For $X = j^r_x f \in J^r(M,N)$, the point $x = \alpha X$ is the *source* of X and the point $f(x) = \beta X$ is the *target* of X. We denote by π^r_s, $0 \leq s \leq r$, the *projection* $j^r_x f \mapsto j^s_x f$ of $r$jets into $s$jets. By $J^r_x(M,N)$ or $J^r(M,N)_y$ we mean the set of all $r$jets of M into N with source $x \in M$ or target $y \in N$, respectively, and we write $J^r_x(M,N)_y = J^r_x(M,N) \cap J^r(M,N)_y$. The map $j^r f : M \to J^r(M,N)$ is called the rth *jet prolongation* of $f : M \to N$ [KMS93].

Since the composition of maps is associative, the same holds for $r$jets. Hence all $r$jets form a category $\mathcal{J}$, the units of which are the $r$jets of the identity maps of manifolds. Then also J^r is a *jet bifunctor* defined on the product category $\mathcal{M}_n \times \mathcal{M}$, with the values in the category of fibre bundles $\mathcal{FB}$ i.e., $J^r : \mathcal{M}_n \times \mathcal{M} \to \mathcal{FB}$.

Next, we are going to describe the coordinate expression of $r-$jets. By

$$D_\alpha f = \frac{\partial^{|\alpha|} f}{(\partial x^1)^{\alpha_1} ... (\partial x^m)^{\alpha_m}},$$

where $\alpha = (\alpha_1, ..., \alpha_m)$ a multiindex of range m, we denote the partial derivative with respect to the multiindex α of a function $f : U \subset \mathbb{R}^m \to \mathbb{R}$. Given a local coordinate system x^i on M in a neighborhood of x and a local coordinate system y^p on N in a neighborhood of $f(x)$, two maps $f, g : M \to N$ satisfy $j^r_x f = j^r_x g$ iff all the partial derivatives up to order r of the components f^p and g^p of their coordinate expressions coincide at x [KMS93]. If we have the curves $x^i = a^i t$ with arbitrary a^i, then the coordinate condition for $f \circ \gamma \sim_r g \circ \gamma$ reads $(D_\alpha f^p(x)) a^\alpha = (D_\alpha g^p(x)) a^\alpha$.

Now, the auxiliary relation $\gamma \sim_r \delta$ can be expressed in terms of $r-$jets, namely two curves $\gamma, \delta : \mathbb{R} \to M$ satisfy $\gamma \sim_r \delta$ iff $j^r_0 \gamma = j^r_0 \delta$.

The elements of $L^r_{m,n} = J^r_0(\mathbb{R}^m, \mathbb{R}^n)$ can be identified with the rth order Taylor expansions of the generating maps, i.e., with the $n$tuples of polynomials of degree r in m variables without absolute term. Such an expression $a^p_\alpha x^\alpha$ is called the polynomial representative of an $r$jet. Hence $L^r_{m,n}$ is a numerical space of the variables a^p_α. $\dim L^r_{m,n} = n \left[\begin{pmatrix} m+r \\ m \end{pmatrix} - 1 \right]$.

The projection $\pi^r_s : L^r_{m,n} \to L^s_{m,n}$ consists in suppressing all terms of degree $> s$. The jet composition $L^r_{m,n} \times L^r_{n,q} \to L^r_{m,q}$,also called truncated polynomial composition, is evaluated by taking the composition of the polynomial representatives and suppressing all terms of degree higher than r. The sets $L^r_{m,n}$ represent the sets of morphisms of a category $\mathcal{L}^r$ over non–negative integers, the composition in which is the jet composition. The set of all invertible elements of $L^r_{m,m}$ with the jet composition is a Lie group G^r_m called the rth differential group or the rth jet group in dimension m. For $r = 1$ the group G^1_m is identified with the general linear group $GL(m, \mathbb{R})$ [KMS93]

The elements of the manifold $T^r_k M = J^r_0(\mathbb{R}^k, M)$ are said to be the $k-$*dimensional velocities* of order r on M, in short $(k,r)-$*velocities*. The inclusion $T_k^r M \subset J^r(\mathbb{R}^m, M)$ defines the structure of a smooth fibre bundle on $T_k^r M \to M$. Every smooth map $f : M \to N$ is extended into an $\mathcal{FB}-$morphism $T_k^r f : T_k^r M \to T_k^r N$ defined by $T_k^r f(j_0^r g) = j_0^r(f \circ g)$. Hence T_k^r is a functor $\mathcal{M} \to \mathcal{FB}$. Since every map $\mathbb{R}^k \to M_1 \times M_2$ coincides with a pair of maps $\mathbb{R}^k \to M_1$ and $\mathbb{R}^k \to M_2$, functor T_k^r preserves products. For $k = r = 1$ we get another definition of the *tangent functor* $T = T_1^1$ [KMS93].

Analogously, the space $T_k^{r*} M = J^r(M, \mathbb{R}^k)_0$ is called the space of all $(k, r)-$co-velocities on M. For $k = 1$ we write in short $T_k^{r*} = T^{r*}$. Since $\mathbb{R}^k$ is a vector space, $T_k^{r*} M \to M$ is a vector bundle with $j_x^r \varphi(u) + j_x^r \psi(u) = j_x^r(\varphi(u) + \psi(u))$, $u \in M$, and $k j_x^r \varphi(u) = j_x^r k\varphi(u)$, $k \in \mathbb{R}$. Every local diffeomorphism $f : M \to N$ is extended to a vector bundle morphism $T_k^{r*} f : T_k^{r*} M \to T_k^{r*} N, j_x^r \varphi \mapsto j_{f(x)}^r(\varphi \circ f^{-1})$, where f^{-1} is constructed locally. In this sense T_k^{r*} is a functor on $\mathcal{M}_n$. For $k = r = 1$ we get the construction of the cotangent bundles as a functor $T_1^{1*} = T^*$ on $\mathcal{M}_n$.

The projection $\pi_{r-1}^r : T^{r*} M \to T^{r-1*} M$ is a linear morphism of vector bundles. Its kernel is described by the following exact sequence of vector bundles over M

$$0 \to S^r T^* M \to T^{r*} M \overset{\pi_{r-1}^r}{\to} T^{r-1*} M \to 0,$$

where S^r indicates the rth symmetric tensor power [KMS93].

Let $\hat{y}$ denote the constant map of M into $y \in N$. The subspace $\left(\pi_{r-1}^r\right)^{-1}(j_x^r \hat{y}) \subset J_x^r(M, N)_y$ is canonically identified with $T_y N \otimes S^r T_x^* M$. For $r = 1$ we have a distinguished element $j_x^1 \hat{y}$ in every fibre of $J^1(M, N) \to M \times N$. This identifies $J^1(M, N)$ with $TN \otimes T^* M$ [KMS93].

2.3 Sections of Biomechanical Bundles

In this section we introduce *sections of biomechanical bundles*, including vector (and tensor) fields and their flows, as well as exterior differential forms.

2.3.1 Biomechanical Evolution and Flow

As a motivational example, consider a biomechanical system that is capable of assuming various states described by points in a set U. For example, U might be $\mathbb{R}^3 \times \mathbb{R}^3$ and a state might be the positions and momenta (x^i, p_i) of a particle moving under the influence of the central force field, with $i = 1, 2, 3$. As time passes, the state evolves. If the state is $\gamma_0 \in U$ at time s and this changes to γ at a later time t, we set

$$F_{t,s}(\gamma_0) = \gamma,$$

and call $F_{t,s}$ the *evolution operator*; it maps a state at time s to what the state would be at time t; that is, after time $t - s$. has elapsed. Determinism is expressed by the *Chapman–Kolmogorov law* [AMR88]:

$$F_{\tau,t} \circ F_{t,s} = F_{\tau,s}, \qquad F_{t,t} = \text{identity}. \tag{2.2}$$

The evolution laws are called *time independent*, or *autonomous*, when $F_{t,s}$ depends only on $t - s$. In this case the preceding law (2.2) becomes the *group property*:

$$F_t \circ F_s = F_{t+s}, \qquad F_0 = \text{identity}. \tag{2.3}$$

We call such an F_t a *flow* and $F_{t,s}$ a *time–dependent flow*, or an evolution operator. If the system is irreversible, that is, defined only for $t \geq s$, we speak of a *semi–flow* [AMR88].

Usually, instead of $F_{t,s}$ the *laws of motion* are given in the form of ODEs that we must solve to find the flow. These equations of motion have the form:

$$\dot{\gamma} = X(\gamma), \qquad \gamma(0) = \gamma_0,$$

where X is a (possibly time–dependent) vector–field on U.

As a continuation of the previous example, consider the motion of a particle of mass m under the influence of the central force field (like gravity, or Coulombic potential) F^i $(i = 1, 2, 3)$, described by the Newtonian equation of motion:

$$m\ddot{x}^i = F^i(x). \tag{2.4}$$

By introducing momenta $p_i = m\dot{x}^i$, equation (6.13) splits into two Hamiltonian equations:

$$\dot{x}^i = p_i/m, \qquad \dot{p}_i = F_i(x). \tag{2.5}$$

Note that in Euclidean space we can freely interchange subscripts and superscripts. However, in general case of a Riemannian manifold, $p_i = mg_{ij}\dot{x}^j$ and (2.5) properly reads

$$\dot{x}^i = g^{ij}p_j/m, \qquad \dot{p}_i = F_i(x). \tag{2.6}$$

The phase–space here is the Riemannian manifold $(\mathbb{R}^3\backslash\{0\}) \times \mathbb{R}^3$, that is, the cotangent bundle of $\mathbb{R}^3\backslash\{0\}$, which is itself the manifold for the central force field. The r.h.s of equations (2.6) define a Hamiltonian vector–field on this 6D manifold by

$$X(x,p) = \left((x^i, p_i), (p_i/m, F_i(x))\right). \tag{2.7}$$

Integration of equations (2.6) produces trajectories (in this particular case, planar conic sections). These trajectories comprise the flow F_t of the vector–field $X(x,p)$ defined in (2.7).

2.3.2 Vector–Fields and Their Flows

Vector–Fields on M

A *vector–field* X on U, where U is an open chart in $n-$manifold M, is a *smooth function* from U to M assigning to each point $m \in U$ a vector at that point, i.e., $X(m) = (m, X(m))$. If $X(m)$ is tangent to M for each $m \in M$, X is said to be a *tangent vector–field* on M. If $X(m)$ is orthogonal to M (i.e., $X(p) \in M_m^{\perp}$) for each $X(m) \in M$, X is said to be a *normal vector–field* on M.

In other words, let M be a C^k-manifold. A C^k-vector–field on M is a C^k-section of the tangent bundle TM of M. Thus a vector–field X on a manifold M is a C^k-map $X : M \to TM$ such that $X(m) \in T_m M$ for all points $m \in M$,and $\pi_M \circ X = Id_M$. Therefore, a vector–field assigns to each point m of M a vector based (i.e., bound) at that point. The set of all C^k vector–fields on M is denoted by $\mathcal{X}^k(M)$.

A vector–field $X \in \mathcal{X}^k(M)$ represents a field of direction indicators [Thi79]: to every point m of M it assigns a vector in the tangent space $T_m M$ at that point. If X is a vector–field on M and (U, ϕ) is a chart on M and $m \in U$, then we have $X(m) = X(m)\,\phi^i \frac{\partial}{\partial \phi^i}$. Following [KMS93], we write $X|_U = X\,\phi^i \frac{\partial}{\partial \phi^i}$.

Let M be a connected $n-$manifold, and let $f : U \to \mathbb{R}$ (U an open set in M) and $c \in \mathbb{R}$ be such that $M = f^{-1}(c)$ (i.e., M is the *level set* of the function f at *height* c) and $\nabla f(m) \neq 0$ for all $m \in M$. Then there exist on M exactly two smooth unit normal vector–fields $N_{1,2}(m) = \pm \frac{\nabla f(m)}{|\nabla f(m)|}$ (here $|X| = (X \cdot X)^{1/2}$ denotes the norm or length of a vector X, and $(\cdot)$ denotes the scalar product on M) for all $m \in M$, called *orientations* on M.

Let $\varphi : M \to N$ be a smooth map. Recall that two vector–fields $X \in \mathcal{X}^k(M)$ and $Y \in \mathcal{X}(N)$ are called $\varphi-$related, if $T\varphi \circ X = Y \circ \varphi$ holds, i.e., if the following diagram commutes:

$$\begin{tikzcd}
TM \arrow[r, "T\varphi"] & TN \\
M \arrow[u, "X"] \arrow[r, "\varphi"'] & N \arrow[u, "Y"']
\end{tikzcd}$$

In particular, a diffeomorphism $\varphi : M \to N$ induces a linear map between vector–fields on two manifolds, $\varphi^* : \mathcal{X}^k(M) \to \mathcal{X}(N)$, such that $\varphi^* X = T\varphi \circ X \circ \varphi^{-1} : N \to TN$, i.e., the following diagram commutes:

$$\begin{tikzcd}
TM \arrow[r, "T\varphi"] & TN \\
M \arrow[u, "X"] \arrow[r, "\varphi"'] & N \arrow[u, "\varphi^* X"']
\end{tikzcd}$$

The correspondences $M \to TM$ and $\varphi \to T\varphi$ obviously define a functor $T : \mathcal{M} \Rightarrow \mathcal{M}$ from the category of smooth manifolds to itself. T is another special case of the vector bundle functor (2.2.3), and closely related to the tangent bundle functor (2.2.1).

A C^k *time–dependent vector–field* is a C^k-map $X : \mathbb{R} \times M \to TM$ such that $X(t,m) \in T_m M$ for all $(t,m) \in \mathbb{R} \times M$, i.e., $X_t(m) = X(t,m)$.

Integral Curves as Biomechanical Trajectories

Recall (2.2.1) that a curve γ at a point m of an $n-$manifold M is a C^0-map from an open interval I of $\mathbb{R}$ into M such that $0 \in I$ and $\gamma(0) = m$. For such a curve we may assign a tangent vector at each point $\gamma(t)$, $t \in I$, by $\dot\gamma(t) = T_t\gamma(1)$.

Let X be a smooth tangent vector–field on the smooth $n-$manifold M, and let $m \in M$. Then there exists an open interval $I \subset \mathbb{R}$ containing 0 and a parameterized curve $\gamma : I \to M$ such that:

1. $\gamma(0) = m$;
2. $\dot\gamma(t) = X(\gamma(t))$ for all $t \in I$; and
3. If $\beta : \tilde{I} \to M$ is any other parameterized curve in M satisfying (1) and (2), then $\tilde{I} \subset I$ and $\beta(t) = \gamma(t)$ for all $t \in \tilde{I}$.

A parameterized curve $\gamma : I \to M$ satisfying condition (2) is called an *integral curve* of the tangent vector–field X. The unique γ satisfying conditions (1)–(3) is the *maximal integral curve* of X through $m \in M$.

In other words, let $\gamma : I \to M$, $t \mapsto \gamma(t)$ be a smooth curve in a manifold M defined on an interval $I \subseteq \mathbb{R}$. $\dot\gamma(t) = \frac{d}{dt}\gamma(t)$ defines a smooth vector–field along γ since we have $\pi_M \circ \dot\gamma = \gamma$. Curve γ is called an *integral curve* or *flow line* of a vector–field $X \in \mathcal{X}^k(M)$ if the tangent vector determined by γ equals X at every point $m \in M$, i.e.,

$$\dot\gamma = X \circ \gamma,$$

or, if the following diagram commutes:

$$\begin{tikzcd}
TI \arrow[r, "Tu"] & TM \\
I \arrow[u, "1"] \arrow[ur, "\dot\gamma"'] \arrow[r, "\gamma"'] & M \arrow[u, "X"']
\end{tikzcd}$$

On a chart (U,ϕ) with coordinates $\phi(m) = \big(x^1(m), ..., x^n(m)\big)$, for which $\varphi \circ \gamma : t \mapsto \gamma_i(t)$ and $T\varphi \circ X \circ \varphi^{-1} : x^i \mapsto \big(x^i, X_i(m)\big)$, this is written

$$\dot\gamma_i(t) = X_i(\gamma(t)), \text{ for all } t \in I \subseteq \mathbb{R}, \tag{2.8}$$

which is an ordinary differential equation of first order in n dimensions.

The *velocity* $\dot{\gamma}$ of the parameterized curve $\gamma(t)$ is a vector–field along γ defined by

$$\dot{\gamma}(t) = (\gamma(t), \dot{x}^1(t), \dots \dot{x}^n(t)).$$

Its length $|\dot{\gamma}| : I \to \mathbb{R}$, defined by $|\dot{\gamma}|(t) = |\dot{\gamma}(t)|$ for all $t \in I$, is a function along α. $|\dot{\gamma}|$ is called *speed* of γ [Arn89].

Each vector–field X along γ is of the form $X(t) = (\gamma(t), X_1(t), \dots, X_n(t))$, where each component X_i is a function along γ. X is *smooth* if each $X_i : I \to M$ is smooth. The *derivative* of a smooth vector–field X along a curve $\gamma(t)$ is the vector–field $\dot{X}$ along γ defined by

$$\dot{X}(t) = (\gamma(t), \dot{X}_1(t), \dots \dot{X}_n(t)).$$

$\dot{X}(t)$ measures the *rate of change of the vector part* $(X_1(t), \dots X_n(t))$ of $X(t)$ *along* γ. Thus, the *acceleration* $\ddot{\gamma}(t)$ of a parameterized curve $\gamma(t)$ is the vector–field along γ obtained by differentiating the velocity field $\dot{\gamma}(t)$.

Differentiation of vector–fields along parameterized curves has the following properties. For X and Y smooth vector–fields on M along the parameterized curve $\gamma : I \to M$ and f a smooth function along γ, we have:

1. $\frac{d}{dt}(X + Y) = \dot{X} + \dot{Y}$;
2. $\frac{d}{dt}(fX) = \dot{f}X + f\dot{X}$; and
3. $\frac{d}{dt}(X \cdot Y) = \dot{X}Y + X\dot{Y}$.

A *geodesic* in M is a parameterized curve $\gamma : I \to M$ whose acceleration $\ddot{\gamma}$ is everywhere orthogonal to M; that is, $\ddot{\gamma}(t) \in M^{\perp}_{\alpha(t)}$ for all $t \in I \subset \mathbf{R}$. Thus a geodesic is a curve in M which always goes 'straight ahead' in the surface. Its acceleration serves only to keep it in the surface. It has no component of acceleration tangent to the surface. Therefore, it also has a constant speed $\dot{\gamma}(t)$.

Let $v \in M_m$ be a vector on M. Then there exists an open interval $I \subset \mathbb{R}$ containing 0 and a geodesic $\gamma : I \to M$ such that:

1. $\gamma(0) = m$ and $\dot{\gamma}(0) = v$; and
2. If $\beta : \tilde{I} \to M$ is any other geodesic in M with $\beta(0) = m$ and $\dot{\beta}(0) = v$, then $\tilde{I} \subset I$ and $\beta(t) = \gamma(t)$ for all $t \in \tilde{I}$.

The geodesic γ is now called the *maximal geodesic* in M passing through m with initial velocity v.

By definition, a parameterized curve $\gamma : I \to M$ is a geodesic of M iff its acceleration is everywhere perpendicular to M, i.e., iff $\ddot{\gamma}(t)$ is a multiple of the orientation $N(\gamma(t))$ for all $t \in I$, i.e., $\ddot{\gamma}(t) = g(t)\, N(\gamma(t))$, where $g : I \to \mathbb{R}$. Taking the scalar product of both sides of this equation with $N(\gamma(t))$ we find $g = -\dot{\gamma}\dot{N}(\gamma(t))$. Thus $\gamma : I \to M$ is geodesic iff it satisfies the differential equation

$$\ddot{\gamma}(t) + \dot{N}(\gamma(t))\, N(\gamma(t)) = 0.$$

This vector equation represents the system of second order component ODEs

$$\ddot{x}^i + N_i(x+1,\dots,x^n)\frac{\partial N_j}{\partial x^k}(x+1,\dots,x^n)\,\dot{x}^j\dot{x}^k = 0.$$

The substitution $u^i = \dot{x}^i$ reduces this second order differential system (in n variables x^i) to the first order differential system

$$\dot{x}^i = u^i, \qquad \dot{u}^i = -N_i(x+1,\dots,x^n)\frac{\partial N_j}{\partial x^k}(x+1,\dots,x^n)\,\dot{x}^j\dot{x}^k$$

(in $2n$ variables x^i and u^i). This first order system is just the differential equation for the integral curves of the vector–field X in $U \times \mathbb{R}$ (U open chart in M), in which case X is called a *geodesic spray*.

Now, when an integral curve $\gamma(t)$ is the path a biomechanical system Ξ follows, i.e., the solution of the equations of motion, it is called a *trajectory*. In this case the parameter t represents time, so that (2.8) describes motion of the system Ξ on its configuration manifold M.

If $X_i\,(m)$ is C^0 the existence of a local solution is guaranteed, and a *Lipschitz condition* would imply that it is unique. Therefore, exactly one integral curve passes through every point, and different integral curves can never cross. As $X \in \mathcal{X}^k(M)$ is C^k, the following statement about the solution with arbitrary initial conditions holds [Thi79, Arn89]:

Theorem. Given a vector–field $X \in \mathcal{X}(M)$, for all points $p \in M$, there exist $\eta > 0$, a neighborhood V of p, and a function $\gamma : (-\eta, \eta) \times V \to M$, $\left(t, x^i\,(0)\right) \mapsto \gamma\left(t, x^i\,(0)\right)$ such that

$$\dot{\gamma} = X \circ \gamma, \qquad \gamma\left(0, x^i\,(0)\right) = x^i\,(0) \qquad \text{for all} \;\; x^i\,(0) \in V \subseteq M.$$

For all $|t| < \eta$, the map $x^i\,(0) \mapsto \gamma\left(t, x^i\,(0)\right)$ is a diffeomorphism f_t^X between V and some open set of M. For proof, see [Die69], I, 10.7.4 and 10.8.

This theorem states that trajectories that are near neighbors cannot suddenly be separated. There is a well–known estimate (see [Die69], I, 10.5) according to which points cannot diverge faster than exponentially in time if the derivative of X is uniformly bounded.

An integral curve $\gamma\,(t)$ is said to be *maximal* if it is not a restriction of an integral curve defined on a larger interval $I \subseteq \mathbb{R}$. It follows from the existence and uniqueness theorems for ODEs with smooth r.h.s and from elementary properties of Hausdorff spaces that for any point $m \in M$ there exists a maximal integral curve γ_m of X, passing for $t = 0$ through point m, i.e., $\gamma(0) = m$.

Theorem (Local Existence, Uniqueness, and Smoothness) [AMR88]. Let E be a Banach space, $U \subset E$ be open, and suppose $X : U \subset E \to E$ is of class C^k, $k \geq 1$. Then

1. For each $x_0 \in U$, there is a curve $\gamma : I \to U$ at x_0 such that $\dot{\gamma}(t) = X\left(\gamma(t)\right)$ for all $t \in I$.

2. Any two such curves are equal on the intersection of their domains.

3. There is a neighborhood U_0 of the point $x_0 \in U$, a real number $a > 0$, and a C^k map $F : U_0 \times I \to E$, where I is the open interval $]-a, a[$, such

that the curve $\gamma_u : I \to E$, defined by $\gamma_u(t) = F(u,t)$ is a curve at $u \in E$ satisfying the ODEs $\dot{\gamma}_u(t) = X(\gamma_u(t))$ for all $t \in I$.

Proposition (Global Uniqueness). Suppose γ_1 and γ_2 are two integral curves of a vector–field X at a point $m \in M$. Then $\gamma_1 = \gamma_2$ on the intersection of their domains [AMR88].

If for every point $m \in M$ the curve γ_m is defined on the entire real axis $\mathbb{R}$, then the vector–field X is said to be *complete*.

The *support* of a vector–field X defined on a manifold M is defined to be the closure of the set $\{m \in M | X(m) = 0\}$. A C^k vector–field with compact support on a manifold M is complete. In particular, a C^k vector–field on a compact manifold is complete. Completeness corresponds to well–defined dynamics persisting eternally.

Now, following [AMR88], for the *derivative* of a C^k function $f : E \to \mathbb{R}$ in the direction X we use the notation $X[f] = df \cdot X$, where *df* stands for the *derivative map*. In standard coordinates on $\mathbb{R}^n$ this is a standard *gradient*

$$df(x) = \nabla f = (\partial_{x^1} f, ..., \partial_{x^n} f), \qquad \text{and} \qquad X[f] = X^i \partial_{x^i} f.$$

Let F_t be the flow of X. Then $f(F_t(x)) = f(F_s(x))$ if $t \geq s$.

For example, Newtonian equations for a moving particle of mass m in a potential field V in $\mathbb{R}^n$ are given by $\ddot{q}^i(t) = -(1/m)\nabla V\left(q^i(t)\right)$, for a smooth function $V : \mathbb{R}^n \to \mathbb{R}$. If there are constants $a, b \in \mathbb{R}$, $b \geq 0$ such that $(1/m)V(q^i) \geq a - b\left\|q^i\right\|^2$, then *every solution exists for all time*. To show this, rewrite the second order equations as a first order system $\dot{q}^i = (1/m)\, p_i$, $\dot{p}_i = -V(q^i)$ and note that the energy $E(q^i, p_i) = (1/2m)\left\| p_i\right\|^2 + V(q)$ is a *first integral* of the motion. Thus, for any solution $\left(q^i(t), p_i(t)\right)$ we have $E\left(q^i(t), p_i(t)\right) = E\left(q^i(0), p_i(0)\right) = V\left(q(0)\right)$.

Let X_t be a C^k time–dependent vector–field on an $n-$manifold M, $k \geq 1$, and let m_0 be an *equilibrium* of X_t, that is, $X_t(m_0) = 0$ for all t. Then for any T there exists a neighborhood V of m_0 such that any $m \in V$ has integral curve existing for time $t \in [-T, T]$.

Biomechanical Flows on M

Recall (2.3.1) that the *flow* F_t of a C^k vector–field $X \in \mathcal{X}^k(M)$ is the *one–parameter group of diffeomorphisms* $F_t : M \to M$ such that $t \mapsto F_t\,(m)$ is the integral curve of X with initial condition m for all $m \in M$ and $t \in I \subseteq \mathbb{R}$. The flow $F_t(m)$ is C^k by induction on k. It is defined as [AMR88]:

$$\frac{d}{dt} F_t(x) = X(F_t(x)).$$

Existence and uniqueness theorems for ODEs guarantee that F_t is smooth in m and t. From uniqueness, we get the *flow property*:

$$F_{t+s} = F_t \circ F_s$$

along with the initial conditions $F_0 =$ identity. The flow property generalizes the situation where $M = V$ is a linear space, $X(x) = A\,x$ for a (bounded) linear operator A, and where $F_t(x) = e^{tA}x$ – to the nonlinear case. Therefore, the flow $F_t(m)$ can be defined as a *formal exponential*

$$F_t(m) = \exp(t\,X) = (I + t\,X + \frac{t^2}{2}X^2 + ...) = \sum_{k=0}^{\infty} \frac{X^k t^k}{k!}.$$

A *time–dependent vector–field* is a map $X : M \times \mathbb{R} \rightarrow TM$ such that $X(m,t) \in T_m M$ for each point $m \in M$ and $t \in \mathbb{R}$. An *integral curve* of X is a curve $\gamma(t)$ in M such that

$$\dot{\gamma}(t) = X\left(\gamma\left(t\right), t\right), \qquad \text{for all} \;\; t \in I \subseteq \mathbb{R}.$$

In this case, the flow is the one–parameter group of diffeomorphisms $F_{t,s} : M \rightarrow M$ such that $t \mapsto F_{t,s}(m)$ is the integral curve $\gamma(t)$ with initial condition $\gamma(s) = m$ at $t = s$. Again, the existence and uniqueness theorem from ODE–theory applies here, and in particular, uniqueness gives the time–dependent flow property, i.e., the *Chapman–Kolmogorov law*

$$F_{t,r} = F_{t,s} \circ F_{s,r}.$$

If X happens to be time independent, the two notions of flows are related by $F_{t,s} = F_{t-s}$ (see [MR99]).

Categories of ODEs

Ordinary differential equations are naturally organized into their categories (see [Koc81]). First order ODEs are organized into a category ODE_1. A first order ODE on a manifold–like object M is a vector–field $X : M \rightarrow TM$, and a *morphism of vector–fields* $(M_1, X_1) \rightarrow (M_2, X_2)$ is a map $f : M_1 \rightarrow M_2$ such that the following diagram commutes

$$\begin{tikzcd}
TM_1 \arrow[r, "Tf"] & TM_2 \\
M_1 \arrow[u, "X_1"] \arrow[r, "f"'] & M_2 \arrow[u, "X_2"']
\end{tikzcd}$$

A *global solution* of the differential equation (M, X), or a *flow line* of a vector–field X, is a morphism from $\left(\mathbb{R}, \frac{\partial}{\partial x}\right)$ to (M, X).

Similarly, second order ODEs are organized into a category ODE_2. A second order ODE on M is usually constructed as a vector–field on TM, $\xi : TM \rightarrow TTM$, and a morphism of vector–fields $(M_1, \xi_1) \rightarrow (M_2, \xi_2)$ is a map $f : M_1 \rightarrow M_2$ such that the following diagram commutes

$$\begin{tikzcd} TTM_1 \arrow[r, "TTf"] & TTM_2 \\ TM_1 \arrow[u, "\xi_1"] \arrow[r, "Tf"'] & TM_2 \arrow[u, "\xi_2"'] \end{tikzcd}$$

Unlike solutions for first order ODEs, solutions for second order ODEs are not in general homomorphisms from $\mathbb{R}$, unless the second order ODE is a *spray* [KR03].

2.3.3 Differential Forms on M

We are already familiar with the basic facts of exterior differential forms (see Introduction). To give a more precise exposition, here we start with 1−forms, which are dual to vector–fields, and after that introduce general $k-$forms.

1−Forms on M

Dual to the notion of a C^k vector–field X on an $n-$manifold M is a C^k *covector–field*, or a C^k 1−form α, which is defined as a C^k-section of the cotangent bundle T^*M, i.e., $\alpha : M \to T^*M$ is smooth and $\pi_M^* \circ X = Id_M$. We denote the set of all C^k 1−forms by $\Omega^1(M)$. A basic example of a 1−form is the differential *df* of a real–valued function $f \in C^k(M, \mathbb{R})$. With point wise addition and scalar multiplication $\Omega^1(M)$ becomes a vector space.

In other words, a C^k 1−form α on a C^k manifold M is a real–valued function on the set of all tangent vectors to M, i.e., $\alpha : TM \to \mathbb{R}$ with the following properties:

1. α is linear on the tangent space T_mM for each $m \in M$;
2. For any C^k vector–field $X \in \mathcal{X}^k(M)$, the function $f : M \to \mathbb{R}$ is C^k.

Given a 1−form α, for each point $m \in M$ the map $\alpha(m) : T_mM \to \mathbb{R}$ is an element of the dual space T_m^*M. Therefore, the space of 1−forms $\Omega^1(M)$ is dual to the space of vector–fields $\mathcal{X}^k(M)$.

In particular, the *coordinate* 1−*forms* $dx^1, ..., dx^n$ are locally defined at any point $m \in M$ by the property that for any vector–field $X = \left(X^1, ..., X^n\right) \in \mathcal{X}^k(M)$,

$$dx^i(X) = X^i.$$

The dx^i's form a *basis* for the 1−forms at any point $m \in M$, with local coordinates $\left(x^1, ..., x^n\right)$, so any other 1−form α may be expressed in the form

$$\alpha = f_i(m)\, dx^i.$$

If a vector–field X on M has the form $X(m) = \left(X^1(m), ..., X^n(m)\right)$, then at any point $m \in M$,

$$\alpha_m(X) = f_i(m)\, X^i(m),$$

where $f \in C^k(M, \mathbb{R})$.

The 1−forms on M are part of an algebra, called the *exterior algebra*, or *Grassmann algebra* on M. The multiplication $\wedge$ in this algebra is called *wedge product* (see (2.9) below), and it is skew–symmetric,

$$dx^i \wedge dx^j = -dx^j \wedge dx^i.$$

One consequence of this is that $dx^i \wedge dx^i = 0$.

k−Forms on *M*

A differential form, or an exterior form α of degree k, or a k−form for short, is a section of the vector bundle $\Lambda^k T^* M$, i.e., $\alpha : M \to \Lambda^k T^* M$. In other words, $\alpha(m) : T_m M \times ... \times T_m M \to \mathbb{R}$ (with k factors $T_m M$) is a function that assigns to each point $m \in M$ a skew–symmetric k−multilinear map on the tangent space $T_m M$ to M at m. Without the skew–symmetry assumption, α would be called a $(0, k)$−tensor–field. The space of all k−forms is denoted by $\Omega^k(M)$. It may also be viewed as the space of all skew symmetric $(0, k)$−tensor–fields, the space of all maps

$$\Phi : \mathcal{X}^k(M) \times ... \times \mathcal{X}^k(M) \to C^k(M, \mathbb{R}),$$

which are k−linear and skew–symmetric (see (2.9) below). We put $\Omega^k(M) = C^k(M, \mathbb{R})$.

In particular, a 2−form ω on an n−manifold M is a section of the vector bundle $\Lambda^2 T^* M$. If (U, ϕ) is a chart at a point $m \in M$ with local coordinates $\left(x^1, ..., x^n\right)$ let $\{e_1, ..., e_n\} = \{\partial_{x^1}, ..., \partial_{x^n}\}$ – be the corresponding basis for $T_m M$, and let $\left\{e^1, ..., e^n\right\} = \left\{dx^1, ..., dx^n\right\}$ – be the dual basis for $T_m^* M$. Then at each point $m \in M$, we can write a 2−form ω as

$$\omega_m(v, u) = \omega_{ij}(m)\, v^i u^j, \quad \text{where} \;\; \omega_{ij}(m) = \omega_m(\partial_{x^i}, \partial_{x^j}).$$

If each summand of a differential form $\alpha \in \Omega^k(M)$ contains k basis 1−forms dx^i's, the form is called a k−form. Functions $f \in C^k(M, \mathbb{R})$ are considered to be 0−forms, and any form on an n−manifold M of degree $k > n$ must be zero due to the skew–symmetry.

Any k−form $\alpha \in \Omega^k(M)$ may be expressed in the form

$$\alpha = f_I\, dx^{i_1} \wedge ... \wedge dx^{i_k} = f_I\, dx^I,$$

where I is a *multiindex* $I = (i_1, ..., i_k)$ of length k, and $\wedge$ is the wedge product which is associative, bilinear and anticommutative.

Just as 1−forms act on vector–fields to give real–valued functions, so k−forms act on k−tuples of vector–fields to give real–valued functions.

The *wedge product* of two differential forms, a $k-$form $\alpha \in \Omega^k(M)$ and an $l-$form $\beta \in \Omega^l(M)$ is a $(k+l)-$form $\alpha \wedge \beta$ defined as:

$$\alpha \wedge \beta = \frac{(k+l)!}{k!l!} \mathbf{A}(\alpha \otimes \beta), \tag{2.9}$$

where $\mathbf{A} : \Omega^k(M) \to \Omega^k(M)$, $\mathbf{A}\tau(e_1, ..., e_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sign} \sigma)\, \tau(e_{\sigma(1)}, ..., e_{\sigma(k)})$, where S_k is the permutation group on k elements consisting of all bijections $\sigma : \{1, ..., k\} \to \{1, ..., k\}$.

For any $k-$ form $\alpha \in \Omega^k(M)$ and $l-$form $\beta \in \Omega^l(M)$, the wedge product is defined fiberwise, i.e., $(\alpha \wedge \beta)_m = \alpha_x \wedge \beta_m$ for each point $m \in M$. It is also associative, i.e., $(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma)$, and graded commutative, i.e., $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$. These properties are proved in multilinear algebra. So $M \Longrightarrow \Omega^k(M)$ is a contravariant functor from the category $\mathcal{M}$ into the category of real graded commutative algebras [KMS93].

Let M be an $n-$manifold, $X \in \mathcal{X}^k(M)$, and $\alpha \in \Omega^{k+1}(M)$. The *interior product*, or *contraction*, $i_X \alpha = X \lrcorner \alpha \in \Omega^k(M)$ of X and α (with *insertion operator* i_X) is defined as

$$i_X \alpha(X^1, ..., X^k) = \alpha(X, X^1, ..., X^k).$$

Insertion operator i_X of a vector–field $X \in \mathcal{X}^k(M)$ is natural with respect to the pull–back F^* of a diffeomorphism $F : M \to N$ between two manifolds, i.e., the following diagram commutes:

$$\begin{tikzcd}
\Omega^k(N) \arrow[r, "F^*"] \arrow[d, "i_X"'] & \Omega^k(M) \arrow[d, "i_{F^*X}"] \\
\Omega^{k-1}(N) \arrow[r, "F^*"'] & \Omega^{k-1}(M)
\end{tikzcd}$$

Similarly, insertion operator i_X of a vector–field $X \in \mathcal{Y}^k(M)$ is natural with respect to the push–forward F_* of a diffeomorphism $F : M \to N$, i.e., the following diagram commutes:

$$\begin{tikzcd}
\Omega^k(M) \arrow[r, "F_*"] \arrow[d, "i_Y"'] & \Omega^k(N) \arrow[d, "i_{F_*Y}"] \\
\Omega^{k-1}(M) \arrow[r, "F_*"'] & \Omega^{k-1}(N)
\end{tikzcd}$$

In case of *Riemannian manifolds* there is another exterior operation. Let M be a smooth $n-$manifold with Riemannian metric $g = \langle , \rangle$ and the corresponding volume element μ (see section 2.5 below). The *Hodge star* operator $* : \Omega^k(M) \to \Omega^{n-k}(M)$ on M is defined as (see Introduction)

$$\alpha \wedge *\beta = \langle \alpha, \beta \rangle \, \mu \quad \text{for } \alpha, \beta \in \Omega^k(M).$$

The Hodge star operator satisfies the following properties for $\alpha, \beta \in \Omega^k(M)$ [AMR88]:

1. $\alpha \wedge *\beta = \langle \alpha, \beta \rangle \, \mu = \beta \wedge *\alpha$;
2. $*1 = \mu, \quad *\mu = (-1)^{Ind(g)}$;
3. $* * \alpha = (-1)^{Ind(g)}(-1)^{k(n-k)}\alpha$;
4. $\langle \alpha, \beta \rangle = (-1)^{Ind(g)} \, \langle *\alpha, *\beta \rangle$, where $Ind(g)$ is the *index* of the metric g.

Exterior Differential Systems

Here we give an informal introduction to *exterior differential systems* (EDS, for short), which are expressions involving differential forms related to any manifold M. Later, when we fully develop the necessary differential geometric as well as variational machinery (see (3.3.6) below), we will give a more precise definition of EDS.

Central in the language of EDS is the notion of *coframing*, which is a real finite–dimensional smooth manifold M with a given global *cobasis* and coordinates, but without requirement for a proper topological and differential structures. For example, $M = \mathbb{R}^3$ is a coframing with cobasis $\{dx, dy, dz\}$ and coordinates $\{x, y, z\}$. In addition to the cobasis and coordinates, a coframing can be given structure equations (2.5.2) and restrictions. For example, $M = \mathbb{R}^2 \backslash \{0\}$ is a coframing with cobasis $\{e^1, e^2\}$, a single coordinate $\{r\}$, structure equations $\{dr = e^1,\, de^1 = 0,\, de^2 = e^1 \wedge e^2/r\}$ and restrictions $\{r \neq 0\}$.

A *system* S on M in EDS terminology is a list of expressions including differential forms (e.g., $S = \{dz - y dx\}$).

Now, a simple EDS is a triple (S, Ω, M), where S is a system on M, and Ω is an *independence condition*: either a decomposable $k-$form or a system of $k-$forms on M. An EDS is a list of simple EDS objects where the various coframings are all disjoint.

An *integral element* of an exterior system (S, Ω, M) is a subspace $P \subset T_m M$ of the tangent space at some point $m \in M$ such that all forms in S vanish when evaluated on vectors from P. Alternatively, an integral element $P \subset T_m M$ can be represented by its annihilator $P^\perp \subset T_m^* M$, comprising those $1-$forms at m which annul every vector in P. For example, with $M = \mathbb{R}^3 = \{(x, y, z)\}$, $S = \{dx \wedge dz\}$ and $\Omega = \{dx, dz\}$, the integral element $P = \{\partial_x + \partial_z, \partial_y\}$ is equally determined by its annihilator $P^\perp = \{dz - dx\}$. Again, for $S = \{dz - y dx\}$ and $\Omega = \{dx\}$, the integral element $P = \{\partial_x + y \partial_z\}$ can be specified simply as $\{dy\}$.

Distributions and Nonholonomic Geometry

Let $TM = \cup_{x \in M} T_x M$, be the tangent bundle of a smooth nD biomechanical manifold M. A sub–bundle $V = \cup_{x \in M} V_x$, where V_x is a vector subspace of

T_xM, smoothly dependent on points $x \in M$, is called the *distribution*. If the manifold M is connected, $\dim V_x$ is called the dimension of the distribution. A vector–field X on M *belongs* to the distribution V if $X(x) \subset V_x$. A curve γ is *admissible relatively to* V, if the vector–field $\dot{\gamma}$ belongs to V. A *differential system* is a linear space of vector–fields having a structure of $C^\infty(M)$ – module. Vector–fields which belong to the distribution V form a differential system $N(V)$. A kD distribution V is *integrable* if the manifold M is foliated to kD sub-manifolds, having V_x as the tangent space at the point x. According to the Frobenius theorem, V is integrable iff the corresponding differential system $N(V)$ is *involutive*, i.e., if it is a Lie sub–algebra of the Lie algebra of vector–fields on M. The *flag* of a differential system N is a sequence of differential systems: $N_0 = N,\ N_1 = [N, N], \dots,\ N_l = [N_{l-1}, N], \dots$.

The differential systems N_i are not always differential systems of some distributions V_i, but if for every i, there exists V_i, such that $N_i = N(V_i)$, then there exists *a flag* of the distribution V: $V = V_0 \subset V_1 \dots$. Such distributions, which have flags, will be called *regular*. It is clear that the sequence $N(V_i)$ is going to stabilize, and there exists a number r such that $N(V_{r-1}) \subset N(V_r) = N(V_{r+1})$. If there exists a number r such that $V_r = TM$, the distribution V is called *completely nonholonomic*, and minimal such r is *the degree of nonholonomicity* of the distribution V.

Now, let us see the mechanical interpretation of these geometric objects. Consider a nonholonomic mechanical system corresponding to a Riemannian manifold (M, g), where g is a metric defined by the system's kinetic energy [DG03]. Suppose that the distribution V is defined by $(n - m)$ one–forms ω_α; in local coordinates $q = (q^1, ..., q^n)$ on M

$$\omega_\rho(q)(\dot{q}) = a_{\rho i}(q)\,\dot{q}^i = 0, \qquad (\rho = m + 1, \dots, n;\ i = 1, \dots, n).$$

A virtual displacement is a vector–field X on M, such that $\omega_\rho(X) = 0$, i.e., X belongs to the differential system $N(V)$.

Differential equations of motion of a given mechanical system follow from the *D'Alambert–Lagrange principle*: trajectory γ of the given system is a solution of the equation

$$\langle \nabla_{\dot{\gamma}}\dot{\gamma} - Q, X \rangle = 0, \tag{2.10}$$

where X is an arbitrary virtual displacement, Q a vector–field of internal forces, and ∇ is the affine Levi–Civita connection for the metric g.

The vector–field $R(x)$ on M, such that $R(x) \in V_x^\perp,\ V_x^\perp \oplus V_x = T_xM$, is called *reaction of ideal nonholonomic connections.* (2.10) can be rewritten as

$$\nabla_{\dot{\gamma}}\dot{\gamma} - Q = R, \qquad \omega_\alpha(\dot{\gamma}) = 0. \tag{2.11}$$

If the system is potential, by introducing $L = T - U$, where U is the potential energy of the system ($Q = -\operatorname{grad} U$), then in local coordinates q on M, equations (2.11) becomes the *forced Lagrangian equation*:

$$\frac{d}{dt}L_{\dot{q}} - L_q = \tilde{R}, \qquad \omega_\alpha(\dot{q}) = 0.$$

Now $\tilde{R}$ is a one–form in $(V^{\perp})$, and it can be represented as a linear combination of one–forms $\omega^{m+1}, \dots, \omega^n$ which define the distribution, $\tilde{R} = \lambda_\alpha \omega_\alpha$.

Suppose $e_1, \dots, e_n$ are the vector–fields on M, such that $e_1(x), \dots, e_n(x)$ form a base of the vector space $T_x M$ at every point $x \in M$, and $e_1, \dots, e_m$ generate the differential system $N(V)$. Express them through the coordinate vector–fields:

$$e_i = A_i^j(q)\partial_{q^j}, \quad (i, j = 1, \dots, n).$$

Denote by p a projection $p : TM \to V$ orthogonal according to the metric g. Corresponding homomorphism of $C^\infty-$modules of sections of TM and V is

$$p\partial_{q^i} = p_i^a e_a, \qquad (a = 1, \dots, m, \, i = 1, \dots, n).$$

Projecting by p the equations (2.11), from $R(x) \in V^{\perp}(x)$, we get $p(R) = 0$, and denoting $p(Q) = \tilde{Q}$ we get

$$\tilde{\nabla}_{\dot{\gamma}} \dot{\gamma} = \tilde{Q},$$

where $\tilde{\nabla}$ is the *projected connection* [DG03]. A relationship between standard Christoffel symbols Γ_{ij}^k and coefficients $\tilde{\Gamma}_{ab}^c$ of the connection $\tilde{\nabla}$, defined by

$$\tilde{\nabla}_{e_a} e_b = \tilde{\Gamma}_{ab}^c e_c, \qquad \text{is given by}$$
$$\tilde{\Gamma}_{ab}^c = \Gamma_{ij}^k A_a^i A_b^j p_k^c + A_a^i \, \partial_{q^i} A_b^j \, p_j^c.$$

If the motion takes place under the inertia $(Q = \tilde{Q} = 0)$, the trajectories of nonholonomic mechanical problem are the geodesics for $\tilde{\nabla}$.

Now, let V be a distribution on M. Denote a $C^\infty(M)-$ module of sections on V by $\Gamma(V)$. *A nonholonomic connection* on the sub–bundle V of TM is a map $\nabla : \Gamma(V) \times \Gamma(V) \to \Gamma(V)$ with the properties:

$$\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z, \qquad \nabla_X(f \cdot Y) = X(f)Y + f\nabla_X Y,$$
$$\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z, (X, Y, Z \in \Gamma(V); \, f, g \in C^\infty(M)).$$

Having a morphism of vector bundles $p_0 : TM \to V$, formed by the projection on V, denote by $q_0 = 1_{TM} - p_0$ the projection on W, $V \oplus W = TM$.

The tensor–field $T_\nabla : \Gamma(V) \times \Gamma(V) \to \Gamma(V)$ defined by

$$T_\nabla(X, Y) = \nabla_X Y - \nabla_Y X - p_0[X, Y] \quad ; \quad X, Y \in \Gamma(V)$$

is called the *torsion tensor* for the connection ∇.

Suppose there is a positively defined metric tensor $g = g_{ij}$ on V. Given a distribution V, with p_0 and g, there exists a unique nonholonomic connection ∇ with the properties [DG03]

$$\nabla_X g(Y, Z) = X(g(Y, Z)) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z) = 0, \qquad T_\nabla = 0.$$

These conditions can be rewritten in the form:

$$\nabla_X Y = \nabla_Y X + p_0[X,Y], \qquad Z(g(X,Y)) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y).$$

By cyclic permutation of X, Y, Z and summing we get:

$$g(\nabla_X Y, Z) = \frac{1}{2}\{X(g(Y,Z)) + Y(g(Z,X)) - Z(g(X,Y)) \tag{2.12}$$
$$+ g(Z, p_0[X,Y]) + g(Y, p_0[Z,X]) - g(X, p_0[Y,Z]\}.$$

Let q^i, $(i = 1, \dots, n)$ be local coordinates on M, such that the first m coordinate vector–fields ∂_{q^j} are projected by projection p_0 into vector–fields e_a, $(a = 1, \dots, m)$, generating the distribution V: $p_0 \partial_{q^j} = p_i^a(q) e_a$. Vector–fields e_a can be expressed in the basis ∂_{q^j} as $e_a = B_a^i \partial_{q^j}$, with $B_a^i p_i^b = \delta_a^b$. Now we give coordinate expressions for the coefficients of the connection Γ_{ab}^c, defined as $\nabla_{e_a} e_b = \Gamma_{ab}^c e_c$. From (2.12) we get

$$\Gamma_{ab}^c = \{{}^{c}_{ab}\} + g_{ae} g^{cd} \Omega_{bd}^e + g_{be} g^{cd} \Omega_{ad}^e - \Omega_{ab}^c,$$

where Ω is obtained from $p_0[e_a, e_b] = -2\Omega_{ab}^c e_c$ as

$$2\Omega_{ab}^c = p_i^c e_a(B_b^i) - p_i^c e_b(B_a^i),$$

and $\{{}^{c}_{ab}\} = \frac{1}{2} g^{ce}(e_a(g_{be}) + e_b(g_{ae}) - e_e(g_{ab}))$.

Exterior Derivative on M

The *exterior derivative* is an operation that takes $k-$forms to $(k+1)-$forms on a smooth manifold M. It defines a unique family of maps $d: \Omega^k(U) \to \Omega^{k+1}(U)$, U open in M, such that (see [AMR88]):

1. d is a $\wedge-$*antiderivation*; that is, d is $\mathbb{R}-$linear and for two forms $\alpha \in \Omega^k(U)$, $\beta \in \Omega^l(U)$,
$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta.$$
2. If $f \in C^k(U, \mathbb{R})$ is a function on M, then $df = \frac{\partial f}{\partial x^i} dx^i : M \to T^*M$ is the differential of f, such that $df(X) = i_X df = \mathcal{L}_X f - d i_X f = \mathcal{L}_X f = X[f]$ for any $X \in \mathcal{X}^k(M)$.
3. $d^2 = d \circ d = 0$ (that is, $d^{k+1}(U) \circ d^k(U) = 0$).
4. d is natural with respect to restrictions $|U$; that is, if $U \subset V \subset M$ are open and $\alpha \in \Omega^k(V)$, then $d(\alpha|U) = (d\alpha)|U$, or the following diagram commutes:

$$\begin{tikzcd}
\Omega^k(V) \arrow[r, "|U"] \arrow[d, "d"'] & \Omega^k(U) \arrow[d, "d"] \\
\Omega^{k+1}(V) \arrow[r, "|U"'] & \Omega^{k+1}(U)
\end{tikzcd}$$

5. d is natural with respect to the Lie derivative $\mathcal{L}_X$ (2.2.3) along any vector–field $X \in \mathcal{X}^k(M)$; that is, for $\omega \in \Omega^k(M)$ we have $\mathcal{L}_X\omega \in \Omega^k(M)$ and $d\mathcal{L}_X\omega = \mathcal{L}_X d\omega$, or the following diagram commutes:
$$\begin{tikzcd} \Omega^k(M) \arrow[r, "\mathcal{L}_X"] \arrow[d, "d"'] & \Omega^k(M) \arrow[d, "d"] \\ \Omega^{k+1}(M) \arrow[r, "\mathcal{L}_X"'] & \Omega^{k+1}(M) \end{tikzcd}$$
6. Let $\varphi : M \to N$ be a C^k map of manifolds. Then $\varphi^* : \Omega^k(N) \to \Omega^k(M)$ is a homomorphism of differential algebras (with $\wedge$ and d) and d is natural with respect to $\varphi^* = F^*$; that is, $\varphi^* d\omega = d\varphi^*\omega$, or the following diagram commutes:
$$\begin{tikzcd} \Omega^k(N) \arrow[r, "\varphi^*"] \arrow[d, "d"'] & \Omega^k(M) \arrow[d, "d"] \\ \Omega^{k+1}(N) \arrow[r, "\varphi^*"'] & \Omega^{k+1}(M) \end{tikzcd}$$
7. Analogously, d is natural with respect to diffeomorphism $\varphi_* = (F^*)^{-1}$; that is, $\varphi_* d\omega = d\varphi_*\omega$, or the following diagram commutes:
$$\begin{tikzcd} \Omega^k(N) \arrow[r, "\varphi_*"] \arrow[d, "d"'] & \Omega^k(M) \arrow[d, "d"] \\ \Omega^{k+1}(N) \arrow[r, "\varphi_*"'] & \Omega^{k+1}(M) \end{tikzcd}$$
8. $\mathcal{L}_X = i_X \circ d + d \circ i_X$ for any $X \in \mathcal{X}^k(M)$ (a 'magic' formula of Cartan).
9. $\mathcal{L}_X \circ d = d \circ \mathcal{L}_X$, i.e., $[\mathcal{L}_X, d] = 0$ for any $X \in \mathcal{X}^k(M)$.
10. $[\mathcal{L}_X, i_Y] = i_{[x,y]}$; in particular, $i_X \circ \mathcal{L}_X = \mathcal{L}_X \circ i_X$ for all $X, Y \in \mathcal{X}^k(M)$.

Given a k–form $\alpha = f_I\, dx^I \in \Omega^k(M)$, the exterior derivative is defined in local coordinates $(x^1, ..., x^n)$ of a point $m \in M$ as
$$d\alpha = d\left(f_I\, dx^I\right) = \frac{\partial f_I}{\partial x^{i_k}} dx^{i_k} \wedge dx^I = df_I \wedge dx^{i_1} \wedge ... \wedge dx^{i_k}.$$

In particular, the exterior derivative of a function $f \in C^k(M, \mathbb{R})$ is a 1–form $df \in \Omega^1(M)$, with the property that for any $m \in M$, and $X \in \mathcal{X}^k(M)$,
$$df_m(X) = X(f),$$
i.e., $df_m(X)$ is a Lie derivative of f at m in the direction of X. Therefore, in local coordinates $(x^1, ..., x^n)$ of a point $m \in M$ we have

$$df = \frac{\partial f}{\partial x^i} dx^i.$$

For any two functions $f, g \in C^k(M, \mathbb{R})$, exterior derivative obeys the *Leibniz rule*:

$$d(fg) = g\, df + f\, dg,$$

and the *chain rule*:

$$d\left(g(f)\right) = g'(f)\, df.$$

A $k-$form $\alpha \in \Omega^k(M)$ is called *closed form* if $d\alpha = 0$, and it is called *exact form* if there exists a $(k-1)-$form $\beta \in \Omega^{k-1}(M)$ such that $\alpha = d\beta$. Since $d^2 = 0$, every exact form is closed. The converse is only partially true (Poincaré Lemma): every closed form is *locally exact*. This means that given a closed $k-$form $\alpha \in \Omega^k(M)$ on an open set $U \subset M$, any point $m \in U$ has a neighborhood on which there exists a $(k-1)-$form $\beta \in \Omega^{k-1}(U)$ such that $d\beta = \alpha|_U$.

The Poincaré lemma is a generalization and unification of two well–known facts in vector calculus:

1. If $\operatorname{curl} F = 0$, then locally $F = \operatorname{grad} f$;
2. If $\operatorname{div} F = 0$, then locally $F = \operatorname{curl} G$.

Poincaré lemma for contractible manifolds: Any closed form on a smoothly contractible manifold is exact.

De Rham Complex and Homotopy Operators on M

Given a biomechanical manifold M, let $\Omega^p(M)$ denote the space of all smooth $p-$forms on M. The differential d, mapping $p-$forms to $(p+1)-$forms, serves to define the *De Rham complex* on M,

$$0 \to \Omega^0(M) \xrightarrow{d^0} \Omega^1(M) \xrightarrow{d^1} \ldots \xrightarrow{d^{n-1}} \Omega^n(M) \to 0. \qquad (2.13)$$

In general, a *complex* (see subsection (1.2.8) above) is defined as a sequence of vector spaces, and linear maps between successive spaces, with the property that the composition of any pair of successive maps is identically 0. In the case of the de Rham complex (2.13), this requirement is a restatement of the closure property for the exterior differential: $d \circ d = 0$.

In particular, for $n = 3$, the De Rham complex on a biomechanical manifold M reads

$$0 \to \Omega^0(M) \xrightarrow{d^0} \Omega^1(M) \xrightarrow{d^1} \Omega^2(M) \xrightarrow{d^2} \Omega^3(M) \to 0. \quad (2.14)$$

If $\omega \equiv f(x, y, z) \in \Omega^0(M)$, then

$$d^0\omega \equiv d^0 f = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = \operatorname{grad}\omega.$$

If $\omega \equiv f dx + g dy + h dz \in \Omega^1(M)$, then

$$d^1\omega \equiv \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx \wedge dy + \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right) dy \wedge dz + \left(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x}\right) dz \wedge dx = \operatorname{curl} \omega.$$

If $\omega \equiv F dy \wedge dz + G dz \wedge dx + H dx \wedge dy \in \Omega^2(M)$, then

$$d^2\omega \equiv \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z} = \operatorname{div} \omega.$$

Therefore, the De Rham complex (2.14) can be written as

$$0 \to \Omega^0(M) \xrightarrow{\text{grad}} \to \Omega^1(M) \xrightarrow{\text{curl}} \Omega^2(M) \xrightarrow{\text{div}} \Omega^3(M) \to 0.$$

Using the closure property for the exterior differential, $d \circ d = 0$, we get the standard identities from vector calculus

$$\operatorname{curl} \cdot \operatorname{grad} = 0 \qquad \text{and} \qquad \operatorname{div} \cdot \operatorname{curl} = 0.$$

The definition of the complex requires that the kernel of one of the linear maps contains the image of the preceding map. The complex is *exact* if this containment is *equality*. In the case of the De Rham complex (2.13), exactness means that a closed $p-$form ω, meaning that $d\omega = 0$, is necessarily an exact $p-$form, meaning that there exists a $(p-1)-$form θ such that $\omega = d\theta$. (For $p = 0$, it says that a smooth function f is closed, $df = 0$, iff it is constant). Clearly, any exact form is closed, but the converse need not hold. Thus the De Rham complex is *not* in general exact. The celebrated *De Rham theorem* states that the extent to which this complex fails to be exact measures purely topological information about the manifold M, its cohomology group.

On the local side, for special types of domains in Euclidean space $\mathbb{R}^m$, there is only trivial topology and we do have exactness of the De Rham complex (2.13). This result, known as the *Poincaré lemma*, holds for *star–shaped* domains $M \subset \mathbb{R}^m$: Let $M \subset \mathbb{R}^m$ be a star–shaped domain. Then the De Rham complex over M is exact.

The key to the proof of exactness of the De Rham complex lies in the construction of suitable *homotopy operators*. By definition, these are linear operators $h : \Omega^p \to \Omega^{p-1}$, taking differential $p-$forms into $(p-1)-$forms, and satisfying the basic identity [Olv86]

$$\omega = dh(\omega) + h(d\omega), \tag{2.15}$$

for all $p-$forms $\omega \in \Omega^p$. The discovery of such a set of operators immediately implies exactness of the complex. For if ω is closed, $d\omega = 0$, then (2.15) reduces to $\omega = d\theta$ where $\theta = h(\omega)$, so ω is exact.

Stokes Theorem and De Rham Cohomology of M

Stokes theorem states that if α is an $(n-1)-$form on an orientable $n-$manifold M, then the integral of $d\alpha$ over M equals the integral of α over ∂M, the boundary of M. The classical theorems of Gauss, Green, and Stokes are special cases of this result.

A *manifold with boundary* is a set M together with an atlas of charts (U,ϕ) with boundary on M. Define (see [AMR88]) the *interior* and *boundary* of M respectively as

$$\mathrm{Int}\, M = \bigcup_U \phi^{-1}\left(\mathrm{Int}\left(\phi(U)\right)\right), \qquad \text{and} \qquad \partial M = \bigcup_U \phi^{-1}\left(\partial\left(\phi(U)\right)\right).$$

If M is a manifold with boundary, then its interior $\mathrm{Int}\, M$ and its boundary ∂M are smooth manifolds without boundary. Moreover, if $f: M \to N$ is a diffeomorphism, N being another manifold with boundary, then f induces, by restriction, two diffeomorphisms

$$\mathrm{Int}\, f : \mathrm{Int}\, M \to \mathrm{Int}\, N, \qquad \text{and} \qquad \partial f : \partial M \to \partial N.$$

If $n = \dim M$, then $\dim(\mathrm{Int}\, M) = n$ and $\dim(\partial M) = n-1$.

To integrate a differential $n-$form over an $n-$manifold M, M must be oriented. If $\mathrm{Int}\, M$ is oriented, we want to choose an *orientation* on ∂M compatible with it. As for manifolds without boundary a volume form on an $n-$manifold with boundary M is a nowhere vanishing $n-$form on M. Fix an orientation on $\mathbb{R}^n$. Then a chart (U,ϕ) is called *positively oriented* if the map $T_m\phi : T_mM \to \mathbb{R}^n$ is orientation preserving for all $m \in U$.

Let M be a compact, oriented kD smooth manifold with boundary ∂M. Let α be a smooth $(k-1)-$form on M. Then the classical *Stokes formula* holds

$$\int_M d\alpha = \int_{\partial M} \alpha.$$

If $\partial M = \emptyset$ then $\int_M d\alpha = 0$.

The quotient space

$$H^k(M) = \frac{\mathrm{Ker}\left(d : \Omega^k(M) \to \Omega^{k+1}(M)\right)}{\mathrm{Im}\left(d : \Omega^{k-1}(M) \to \Omega^k(M)\right)}$$

is called the kth *De Rham cohomology group* of a manifold M. The *De Rham theorem* states that these Abelian groups are isomorphic to the so–called singular cohomology groups of M defined in algebraic topology in terms of simplices and that depend only on the topological structure of M and not on its differentiable structure. The isomorphism is provided by integration; the fact that the integration map drops to the preceding quotient is guaranteed by Stokes' theorem.

The exterior derivative commutes with the pull–back of differential forms. That means that the vector bundle Λ^kT^*M is in fact the value of a functor, which associates a bundle over M to each manifold M and a vector bundle homomorphism over φ to each (local) diffeomorphism φ between manifolds of the same dimension. This is a simple example of the concept of a natural bundle. The fact that the exterior derivative d transforms sections of $\Lambda^k T^*M$ into sections of $\Lambda^{k+1}T^*M$ for every manifold M can be expressed by saying that d is an operator from $\Lambda^k T^*M$ into $\Lambda^{k+1}T^*M$. That the exterior derivative d commutes with (local) diffeomorphisms now means, that d is a natural operator from the functor $\Lambda^k T^*$ into functor $\Lambda^{k+1}T^*$. If $k > 0$, one can show that d is the unique natural operator between these two natural bundles up to a constant. So even linearity is a consequence of naturality [KMS93].

Euler–Poincaré Characteristics of M

The *Euler–Poincaré characteristics* of a manifold M equals the sum of its *Betti numbers*

$$\chi(M) = \sum_{p=0}^{n} (-1)^p \, b_p.$$

In case of $2n$D oriented compact Riemannian manifold M (*Gauss–Bonnet theorem*) its Euler–Poincaré characteristics is equal

$$\chi(M) = \int_M \gamma,$$

where γ is a closed $2n$ form on M, given by

$$\gamma = \frac{(-1)^n}{(4\pi)^n n!} \epsilon^{1\dots 2n}_{i_1\dots i_{2n}} \, \Omega^{i_1}_{i_2} \wedge \Omega^{i_{2n-1}}_{i_{2n}},$$

where Ω^i_j is the curvature 2–form of a Riemannian connection on M (see Chapter 4 for more details).

Poincaré–Hopf theorem: The Euler–Poincaré characteristics $\chi(M)$ of a compact manifold M equals the sum of indices of zeros of any vector–field on M which has only isolated zeros.

Duality of Chains and Forms on M

In topology of finite–dimensional smooth (i.e., C^{p+1} with $p \geq 0$) manifolds, a fundamental notion is the *duality* between p–chains C and p–forms (i.e., p–cochains) ω on the smooth manifold M, or domains of integration and integrands – as an integral on M represents a bilinear functional (see [BM82, DP97])

$$\int_C \omega \equiv \langle C, \omega \rangle \,, \tag{2.16}$$

where the integral is called the *period* of ω. Period depends only on the cohomology class of ω and the homology class of C. A closed form (cocycle) is exact (coboundary) if all its periods vanish, i.e., $d\omega = 0$ implies $\omega = d\theta$. The duality (2.16) is based on the classical Stokes formula

$$\int_C d\omega = \int_{\partial C} \omega.$$

This is written in terms of scalar products on M as

$$\langle C, d\omega \rangle = \langle \partial C, \omega \rangle \,,$$

where ∂C is the boundary of the $p-$chain C oriented coherently with C. While the boundary operator ∂ is a global operator, the coboundary operator, that is, the exterior derivative d, is local, and thus more suitable for applications. The main property of the exterior differential,

$$d^2 = 0 \qquad \text{implies} \qquad \partial^2 = 0,$$

can be easily proved by the use of Stokes' formula

$$\left\langle \partial^2 C, \omega \right\rangle = \langle \partial C, d\omega \rangle = \left\langle C, d^2 \omega \right\rangle = 0.$$

The analysis of p–chains and p–forms on the finite–dimensional biomechanical manifold M is usually performed in (co)homology categories (see [DP97, Die88]) related to M.

Let $\mathcal{M}^\bullet$ denote the category of cochains, (i.e., p–forms) on the smooth manifold M. When $\mathcal{C} = \mathcal{M}^\bullet$, we have the category $\mathcal{S}^\bullet(\mathcal{M}^\bullet)$ of generalized cochain complexes $A^\bullet$ in $\mathcal{M}^\bullet$, and if $A' = 0$ for $n < 0$ we have a subcategory $\mathcal{S}^\bullet_{\mathcal{DR}}(\mathcal{M}^\bullet)$ of the De Rham differential complexes in $\mathcal{M}^\bullet$

$$\begin{aligned} A^\bullet_{DR} :\ 0 \to \Omega^0(M) \xrightarrow{\quad d \quad} \Omega^1(M) \xrightarrow{\quad d \quad} \Omega^2(M) \cdots \\ \cdots \xrightarrow{\quad d \quad} \Omega^n(M) \xrightarrow{\quad d \quad} \cdots . \end{aligned} \tag{2.17}$$

Here $A' = \Omega^n(M)$ is the vector space over $\mathbb{R}$ of all p–forms ω on M (for $p = 0$ the smooth functions on M) and $d_n = d : \Omega^{n-1}(M) \to \Omega^n(M)$ is the exterior differential. A form $\omega \in \Omega^n(M)$ such that $d\omega = 0$ is a closed form or n–cocycle. A form $\omega \in \Omega^n(M)$ such that $\omega = d\theta$, where $\theta \in \Omega^{n-1}(M)$, is an exact form or n–coboundary. Let $Z^n(M) = \mathrm{Ker}(d)$ (resp. $B^n(M) = \mathrm{Im}(d)$) denote a real vector space of cocycles (resp. coboundaries) of degree n. Since $d_{n+1}\, d_n = d^2 = 0$, we have $B^n(M) \subset Z^n(M)$. The quotient vector space

$$H^n_{DR}(M) = \mathrm{Ker}(d)/\,\mathrm{Im}(d) = Z^n(M)/B^n(M)$$

is the De Rham cohomology group. The elements of $H^n_{DR}(M)$ represent equivalence sets of cocycles. Two cocycles ω_1, ω_2 belong to the same equivalence set, or are cohomologous (written $\omega_1 \sim \omega_2$) iff they differ by a coboundary $\omega_1 - \omega_2 = d\theta$. The De Rham cohomology class of any form $\omega \in \Omega^n(M)$ is

$[\omega] \in H^n_{DR}(M)$. The De Rham differential complex (2.17) can be considered as a system of second–order ODEs $d^2\theta = 0$, $\theta \in \Omega^{n-1}(M)$ having a solution represented by $Z^n(M) = \mathrm{Ker}(d)$.

Analogously let $\mathcal{M}_\bullet$ denote the category of chains on the smooth manifold M. When $\mathcal{C} = \mathcal{M}_\bullet$, we have the category $\mathcal{S}_\bullet(\mathcal{M}_\bullet)$ of generalized chain complexes $A_\bullet$ in $\mathcal{M}_\bullet$, and if $A_n = 0$ for $n < 0$ we have a subcategory $\mathcal{S}^{\mathcal{C}}_\bullet(\mathcal{M}_\bullet)$ of chain complexes in $\mathcal{M}_\bullet$

$$A_\bullet : 0 \leftarrow C^0(M) \xleftarrow{\partial} C^1(M) \xleftarrow{\partial} C^2(M) \cdots \xleftarrow{\partial} C^n(M) \xleftarrow{\partial} \cdots .$$

Here $A_n = C^n(M)$ is the vector space over $\mathbb{R}$ of all finite chains C on the manifold M and $\partial_n = \partial : C^{n+1}(M) \to C^n(M)$. A finite chain C such that $\partial C = 0$ is an $n-$cycle. A finite chain C such that $C = \partial B$ is an $n-$boundary. Let $Z_n(M) = \mathrm{Ker}(\partial)$ (resp. $B_n(M) = \mathrm{Im}(\partial)$) denote a real vector space of cycles (resp. boundaries) of degree n. Since $\partial_{n+1}\partial_n = \partial^2 = 0$, we have $B_n(M) \subset Z_n(M)$. The quotient vector space

$$H^C_n(M) = \mathrm{Ker}(\partial)/\,\mathrm{Im}(\partial) = Z_n(M)/B_n(M)$$

is the $n-$homology group. The elements of $H^C_n(M)$ are equivalence sets of cycles. Two cycles C_1, C_2 belong to the same equivalence set, or are homologous (written $C_1 \sim C_2$), iff they differ by a boundary $C_1 - C_2 = \partial B$). The homology class of a finite chain $C \in C^n(M)$ is $[C] \in H^C_n(M)$.

The dimension of the $n-$cohomology (resp. $n-$homology) group equals the nth Betti number b^n (resp. b_n) of the manifold M. *Poincaré lemma* says that on an open set $U \in M$ diffeomorphic to $\mathbb{R}^N$, all closed forms (cycles) of degree $p \geq 1$ are exact (boundaries). That is, the Betti numbers satisfy $b^p = 0$ (resp. $b_p = 0$) for $p = 1, \ldots, n$.

The *De Rham theorem* states the following. The map Φ: $H_n \times H^n \to \mathbb{R}$ given by $([C], [\omega]) \to \langle C, \omega \rangle$ for $C \in Z_n, \omega \in Z^n$ is a bilinear nondegenerate map which establishes the duality of the groups (vector spaces) H_n and H^n and the equality $b_n = b^n$.

Other Exterior Operators on M

As the configuration manifold M is an oriented ND Riemannian manifold, we may select an orientation on all tangent spaces T_mM and all cotangent spaces T^*_mM, with the local coordinates $x^i = (q^i, p_i)$ at a point $m \in M$, in a consistent manner. The simplest way to do that is to choose the Euclidean orthonormal basis $\partial_1, ..., \partial_N$ of $\mathbb{R}^N$ as being positive.

Since the manifold M carries a Riemannian structure $g = \langle , \rangle$, we have a scalar product on each T^*_mM. So, we can define (as above) the linear *Hodge star* operator

$$* : \Lambda^p(T^*_mM) \to \Lambda^{N-p}(T^*_mM),$$

which is a base point preserving operator

$$* : \Omega^p(M) \to \Omega^{N-p}(M), \qquad (\Omega^p(M) = \Gamma(\Lambda^p(M)))$$

(here $\Lambda^p(V)$ denotes the p-fold exterior product of any vector space V, $\Omega^p(M)$ is a space of all $p-$forms on M, and $\Gamma(E)$ denotes the space of sections of the vector bundle E). Also,

$$** = (-1)^{p(N-p)} : \Lambda^p(T_x^* M) \to \Lambda^p(T_m^* M).$$

As the metric on $T_m^* M$ is given by $g^{ij}(x) = (g_{ij}(x))^{-1}$, we have the *volume form* defined in local coordinates as

$$*(1) = \sqrt{\det(g_{ij})} dx^1 \wedge ... \wedge dx^n,$$

and

$$\mathrm{vol}(M) = \int_M *(1).$$

For any to $p-$forms $\alpha, \beta \in \Omega^p(M)$ with compact support, we define the (bilinear and positive definite) L^2-product as

$$(\alpha, \beta) = \int_M \langle \alpha, \beta \rangle * (1) = \int_M \alpha \wedge *\beta.$$

We can extend the product $(\cdot, \cdot)$ to $L^2(\Omega^p(M))$; it remains bilinear and positive definite, because as usual, in the definition of L^2, functions that differ only on a set of measure zero are identified.

Using the Hodge star operator $*$, we can introduce the *codifferential* operator δ, which is formally adjoint to the exterior derivative $d : \Omega^p(M) \to \Omega^{p+1}(M)$ on $\oplus_{p=0}^N \Omega^p(M)$ w.r.t. $(\cdot, \cdot)$. This means that for $\alpha \in \Omega^{p-1}(M), \beta \in \Omega^p(M)$

$$(d\alpha, \beta) = (\alpha, \delta\beta).$$

Therefore, we have $\delta : \Omega^p(M) \to \Omega^{p-1}(M)$ and

$$\delta = (-1)^{N(p+1)+1} * d * .$$

Now, the Laplace–Beltrami operator (or, Hodge Laplacian, see subsection (4.3.1) below), Δ on $\Omega^p(M)$, is defined by relation similar to (2.15) above

$$\Delta = d\delta + \delta d : \Omega^p(M) \to \Omega^p(M) \tag{2.18}$$

and $\alpha \in \Omega^p(M)$ is called *harmonic* if $\Delta\alpha = 0$.

Let M be a compact, oriented Riemannian manifold, E a vector bundle with a bundle metric $\langle \cdot, \cdot \rangle$ over M,

$$D = d + A : \Omega^{p-1}(AdE) \to \Omega^p(AdE), \qquad \text{with} \quad A \in \Omega^1(AdE)$$

– a tensorial and $\mathbb{R}-$linear metric connection on E with curvature $F_D \in \Omega^2(AdE)$ (Here by $\Omega^p(AdE)$ we denote the space of those elements of $\Omega^p(EndE)$ for which the endomorphism of each fibre is skew symmetric; $EndE$ denotes the space of linear endomorphisms of the fibers of E).

2.4 Lie Categories in Human–Like Biomechanics

In this section we introduce *Lie categories in biomechanics*, as a unique framework for the concepts of Lie derivative, Lie groups and their associated Lie algebras, as well as more general Lie symmetries.

2.4.1 Lie Derivative in Biomechanics

Lie derivative is popularly called '*fisherman's derivative*'. In continuum mechanics it is called Liouville operator. This is a central differential operator in modern differential geometry and its physical and control applications.

Lie Derivative on Functions

To define how vector–fields operate on functions on an $m-$manifold M, we will use the *directional derivative* or *Lie derivative* (see (2.2.3)). Let $f : M \to \mathbb{R}$ so $Tf : TM \to T\mathbb{R} = \mathbb{R} \times \mathbb{R}$. Following [AMR88] we write Tf acting on a vector $v \in T_m M$ in the form

$$Tf \cdot v = (f(m), df(m) \cdot v) .$$

This defines, for each point $m \in M$, the element $df(m) \in T_m^* M$. Thus df is a section of the cotangent bundle T^*M, i.e., a $1-$form. The $1-$form $df : M \to T^*M$ defined this way is called the *differential* of f. If f is C^k, then df is C^{k-1}.

If $\phi : U \subset M \to V \subset E$ is a local chart for M, then the local representative of $f \in C^k(M, \mathbb{R})$ is the map $f : V \to \mathbb{R}$ defined by $f = f \circ \phi^{-1}$. The local representative of Tf is the tangent map for local manifolds,

$$Tf(x, v) = (f(x), Df(x) \cdot v) .$$

Thus the local representative of df is the derivative of the local representative of f. In particular, if $(x^1, ..., x^n)$ are local coordinates on M, then the local components of df are

$$(df)^i = \partial_{x^i} f.$$

The introduction of df leads to the following definition of the Lie derivative. The *directional* or *Lie derivative* $\mathcal{L}_X : C^k(M, \mathbb{R}) \to C^{k-1}(M, \mathbb{R})$ of a function $f \in C^k(M, \mathbb{R})$ along a vector–field X is defined by

$$\mathcal{L}_X f(m) = X[f](m) = df(m) \cdot X(m),$$

for any $m \in M$. Denote by $X[f] = df(X)$ the map $M \ni m \mapsto X[f](m) \in \mathbb{R}$. If f is $F-$valued, the same definition is used, but now $X[f]$ is $F-$valued.

If a local chart (U, ϕ) on an $n-$manifold M has local coordinates $(x^1, ..., x^n)$, the local representative of $X[f]$ is given by the function

$$\mathcal{L}_X f = X[f] = X^i \, \partial_{x^i} f.$$

Evidently if f is C^k and X is C^{k-1} then $X[f]$ is C^{k-1}.

Let $\varphi : M \to N$ be a diffeomorphism. Then $\mathcal{L}_X$ is natural with respect to push–forward by φ. That is, for each $f \in C^k(M, \mathbb{R})$,

$$\mathcal{L}_{\varphi_* X}(\varphi_* f) = \varphi_* \mathcal{L}_X f,$$

i.e., the following diagram commutes:

$$\begin{tikzcd} C^k(M,\mathbb{R}) \arrow[r, "\varphi_*"] \arrow[d, "\mathcal{L}_X"'] & C^k(N,\mathbb{R}) \arrow[d, "\mathcal{L}_{\varphi_* X}"] \\ C^k(M,\mathbb{R}) \arrow[r, "\varphi_*"'] & C^k(N,\mathbb{R}) \end{tikzcd}$$

Also, $\mathcal{L}_X$ is natural with respect to restrictions. That is, for U open in M and $f \in C^k(M, \mathbb{R})$,

$$\mathcal{L}_{X|U}(f|U) = (\mathcal{L}_X f)|U,$$

where $|U : C^k(M, \mathbb{R}) \to C^k(U, \mathbb{R})$ denotes restriction to U, i.e., the following diagram commutes:

$$\begin{tikzcd} C^k(M,\mathbb{R}) \arrow[r, "|U"] \arrow[d, "\mathcal{L}_X"'] & C^k(U,\mathbb{R}) \arrow[d, "\mathcal{L}_{X|U}"] \\ C^k(M,\mathbb{R}) \arrow[r, "|U"'] & C^k(U,\mathbb{R}) \end{tikzcd}$$

Since $\varphi^* = (\varphi^{-1})_*$ the Lie derivative is also natural with respect to pull–back by φ. This has a generalization to φ–related vector–fields as follows: Let $\varphi : M \to N$ be a C^k–map, $X \in \mathcal{X}^{k-1}(M)$ and $Y \in \mathcal{X}^{k-1}(N)$, $k \geq 1$. If $X \sim_\varphi Y$, then

$$\mathcal{L}_X(\varphi^* f) = \varphi^* \mathcal{L}_Y f$$

for all $f \in C^k(N, \mathbb{R})$, i.e., the following diagram commutes:

$$\begin{tikzcd} C^k(N,\mathbb{R}) \arrow[r, "\varphi^*"] \arrow[d, "\mathcal{L}_Y"'] & C^k(M,\mathbb{R}) \arrow[d, "\mathcal{L}_X"] \\ C^k(N,\mathbb{R}) \arrow[r, "\varphi^*"'] & C^k(M,\mathbb{R}) \end{tikzcd}$$

The Lie derivative map $\mathcal{L}_X : C^k(M, \mathbb{R}) \to C^{k-1}(M, \mathbb{R})$ is a *derivation*, i.e., for two functions $f, g \in C^k(M, \mathbb{R})$ the *Leibniz rule* is satisfied

$$\mathcal{L}_X(fg) = g\mathcal{L}_X f + f\mathcal{L}_X g;$$

Also, Lie derivative of a constant function is zero, $\mathcal{L}_X(\text{const}) = 0$.

The connection between the Lie derivative $\mathcal{L}_X f$ of a function $f \in C^k(M, \mathbb{R})$ and the flow F_t of a vector–field $X \in \mathcal{X}^{k-1}(M)$ is given as:

$$\frac{d}{dt}\left(F_t^* f\right) = F_t^*\left(\mathcal{L}_X f\right).$$

Lie Derivative of Vector Fields

If $X, Y \in \mathcal{X}^k(M)$, $k \geq 1$ are two vector–fields on M, then

$$[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_X \circ \mathcal{L}_Y - \mathcal{L}_Y \circ \mathcal{L}_X$$

is a derivation map from $C^{k+1}(M, \mathbb{R})$ to $C^{k-1}(M, \mathbb{R})$. Then there is a unique vector–field, $[X, Y] \in \mathcal{X}^k(M)$ of X and Y such that $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$ and $[X, Y](f) = X\left(Y(f)\right) - Y\left(X(f)\right)$ holds for all functions $f \in C^k(M, \mathbb{R})$. This vector–field is also denoted $\mathcal{L}_X Y$ and is called the Lie derivative (2.2.3) of Y with respect to X, or the *Lie bracket* of X and Y. In a local chart (U, ϕ) at a point $m \in M$ with coordinates $(x^1, ..., x^n)$, for $X|_U = X^i \partial_{x^i}$ and $Y|_U = Y^i \partial_{x^i}$ we have

$$\left[X^i \partial_{x^i}, Y^j \partial_{x^j}\right] = \left(X^i \left(\partial_{x^i} Y^j\right) - Y^i \left(\partial_{x^i} X^j\right)\right) \partial_{x^j},$$

since second partials commute. If, also X has flow F_t, then [AMR88]

$$\frac{d}{dt}\left(F_t^* Y\right) = F_t^*\left(\mathcal{L}_X Y\right).$$

In particular, if $t = 0$, this formula becomes

$$\frac{d}{dt}|_{t=0}\left(F_t^* Y\right) = \mathcal{L}_X Y.$$

Then the unique C^{k-1} vector–field $\mathcal{L}_X Y = [X, Y]$ on M defined by

$$[X, Y] = \frac{d}{dt}|_{t=0}\left(F_t^* Y\right),$$

is called the Lie derivative of Y with respect to X, or the Lie bracket of X and Y, and can be interpreted as the leading order term that results from the sequence of flows

$$F_t^{-Y} \circ F_t^{-X} \circ F_t^{Y} \circ F_t^{-X}(m) = \epsilon^2 [X, Y](m) + \mathcal{O}(\epsilon^3), \tag{2.19}$$

for some real $\epsilon > 0$. Therefore a Lie bracket can be interpreted as a 'new direction' in which the system can flow, by executing the sequence of flows (2.19).

Lie bracket satisfies the following property:

$$[X,Y][f] = X[Y[f]] - Y[X[f]],$$

for all $f \in C^{k+1}(U,\mathbb{R})$, where U is open in M.

An important relationship between flows of vector–fields is given by the *Campbell–Baker–Hausdorff* formula:

$$F_t^Y \circ F_t^X = F_t^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}([X,[X,Y]]-[Y,[X,Y]])+\dots} \tag{2.20}$$

Essentially, if given the composition of multiple flows along multiple vector–fields, this formula gives the one flow along one vector–field which results in the same net flow. One way to prove the Campbell–Baker–Hausdorf formula (2.20) is to expand the product of two formal exponentials and equate terms in the resulting formal power series.

Lie bracket is the $\mathbb{R}-$bilinear map $[,] : \mathcal{X}^k(M) \times \mathcal{X}^k(M) \to \mathcal{X}^k(M)$ with the following properties:

1. $[X,Y] = -[Y,X]$, i.e., $\mathcal{L}_X Y = -\mathcal{L}_Y X$ for all $X,Y \in \mathcal{X}^k(M)$ – skew–symmetry;
2. $[X,X] = 0$ for all $X \in \mathcal{X}^k(M)$;
3. $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ for all $X,Y,Z \in \mathcal{X}^k(M)$ – the Jacobi identity;
4. $[fX,Y] = f[X,Y] - (Yf)X$, i.e., $\mathcal{L}_{fX}(Y) = f(\mathcal{L}_X Y) - (\mathcal{L}_Y f)X$ for all $X,Y \in \mathcal{X}^k(M)$ and $f \in C^k(M,\mathbb{R})$;
5. $[X,fY] = f[X,Y] + (Xf)Y$, i.e., $\mathcal{L}_X(fY) = f(\mathcal{L}_X Y) + (\mathcal{L}_X f)Y$ for all $X,Y \in \mathcal{X}^k(M)$ and $f \in C^k(M,\mathbb{R})$;
6. $[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[x,y]}$ for all $X,Y \in \mathcal{X}^k(M)$.

The pair $(\mathcal{X}^k(M), [,])$ is the prototype of a *Lie algebra* [KMS93]. In more general case of a *general linear Lie algebra* $\mathfrak{gl}(n)$, which is the Lie algebra associated to the Lie group $GL(n)$, Lie bracket is given by a *matrix commutator*

$$[A,B] = AB - BA,$$

for any two matrices $A,B \in \mathfrak{gl}(n)$.

Let $\varphi : M \to N$ be a diffeomorphism. Then $\mathcal{L}_X : \mathcal{X}^k(M) \to \mathcal{X}^k(M)$ is natural with respect to push–forward by φ. That is, for each $f \in C^k(M,\mathbb{R})$,

$$\mathcal{L}_{\varphi_* X}(\varphi_* Y) = \varphi_* \mathcal{L}_X Y,$$

i.e., the following diagram commutes:

$$\begin{tikzcd}
\mathcal{X}^k(M) \arrow[r, "\varphi_*"] \arrow[d, "\mathcal{L}_X"'] & \mathcal{X}^k(N) \arrow[d, "\mathcal{L}_{\varphi_* X}"] \\
\mathcal{X}^k(M) \arrow[r, "\varphi_*"'] & \mathcal{X}^k(N)
\end{tikzcd}$$

Also, $\mathcal{L}_X$ is natural with respect to restrictions. That is, for U open in M and $f \in C^k(M, \mathbb{R})$,

$$[X|U, Y|U] = [X, Y]|U,$$

where $|U : C^k(M, \mathbb{R}) \to C^k(U, \mathbb{R})$ denotes restriction to U, i.e., the following diagram commutes [AMR88]:

$$\begin{tikzcd}
\mathcal{X}^k(M) \arrow[r, "|U"] \arrow[d, "\mathcal{L}_X"'] & \mathcal{X}^k(U) \arrow[d, "\mathcal{L}_{X|U}"] \\
\mathcal{X}^k(M) \arrow[r, "|U"'] & \mathcal{X}^k(U)
\end{tikzcd}$$

If a local chart (U, ϕ) on an $n-$manifold M has local coordinates $(x^1, ..., x^n)$, then the local components of a Lie bracket are

$$[X, Y]^j = X^i \, \partial_{x^i} Y^j - Y^i \, \partial_{x^i} X^j,$$

that is, $[X, Y] = (X \cdot \nabla)Y - (Y \cdot \nabla)X$.

Let $\varphi : M \to N$ be a C^k-map, $X \in \mathcal{X}^{k-1}(M)$ and $Y \in \mathcal{X}^{k-1}(N)$, $k \geq 1$. Then $X \sim_\varphi Y$, iff

$$(Y[f]) \circ \varphi = X[f \circ \varphi]$$

for all $f \in C^k(V, \mathbb{R})$, where V is open in N.

For every $X \in \mathcal{X}^k(M)$, the operator $\mathcal{L}_X$ is a derivation on $\big(C^k(M, \mathbb{R}), \mathcal{X}^k(M)\big)$, i.e., $\mathcal{L}_X$ is $\mathbb{R}-$linear.

For any two vector–fields $X \in \mathcal{X}^k(M)$ and $Y \in \mathcal{X}^k(N)$, $k \geq 1$ with flows F_t and G_t, respectively, if $[X, Y] = 0$ then $F_t^* Y = Y$ and $G_t^* X = X$.

Derivative of the Evolution Operator

Recall (2.3.1) that the time–dependent flow or evolution operator $F_{t,s}$ of a vector–field $X \in \mathcal{X}^k(M)$ is defined by the requirement that $t \mapsto F_{t,s}(m)$ be the integral curve of X starting at a point $m \in M$ at time $t = s$, i.e.,

$$\frac{d}{dt} F_{t,s}(m) = X\left(t, F_{t,s}(m)\right) \qquad \text{and} \qquad F_{t,t}(m) = m.$$

By uniqueness of integral curves we have $F_{t,s} \circ F_{s,r} = F_{t,r}$ (replacing the flow property $F_{t+s} = F_t + F_s$) and $F_{t,t} =$ identity.

Let $X_t \in \mathcal{X}^k(M)$, $k \geq 1$ for each t and suppose $X(t, m)$ is continuous in $(t, m) \in \mathbb{R} \times M$. Then $F_{t,s}$ is of class C^k and for $f \in C^{k+1}(M, \mathbb{R})$ [AMR88], and $Y \in \mathcal{X}^k(M)$, we have

1. $\frac{d}{dt} F_{t,s}^* f = F_{t,s}^* \left(\mathcal{L}_{X_t} f\right)$, and
2. $\frac{d}{dt} F_{t,s}^* f = F_{t,s}^* ([X_t, Y]) = F_{t,s}^* \left(\mathcal{L}_{X_t} Y\right)$.

From the above theorem, the following identity holds:

$$\frac{d}{dt} F_{t,s}^* f = -X_t \left[F_{t,s}^* f\right].$$

Lie Derivative of Differential Forms

Since $\mathcal{F} : M \Longrightarrow \Lambda^k T^* M$ is a vector bundle functor on $\mathcal{M}$, the Lie derivative (2.2.3) of a $k-$form $\alpha \in \Omega^k(M)$ along a vector–field $X \in \mathcal{X}^k(M)$ is defined by

$$\mathcal{L}_X \alpha = \frac{d}{dt}|_{t=0}\ F_t^* \alpha.$$

It has the following properties:

1. $\mathcal{L}_X(\alpha \wedge \beta) = \mathcal{L}_X\, \alpha \wedge \beta + \alpha \wedge \mathcal{L}_X\, \beta$, so $\mathcal{L}_X$ is a derivation.
2. $[\mathcal{L}_X, \mathcal{L}_Y]\ \alpha = \mathcal{L}_{[X,Y]}\, \alpha$.
3. $\frac{d}{dt} F_t^* \alpha = F_t^* \mathcal{L}_X \alpha = \mathcal{L}_X\,(F_t^* \alpha)$.

Formula (3) holds also for time–dependent vector–fields in the sense that $\frac{d}{dt} F_{t,s}^* \alpha = F_{t,s}^* \mathcal{L}_X \alpha = \mathcal{L}_X \left(F_{t,s}^* \alpha\right)$ and in the expression $\mathcal{L}_X \alpha$ the vector–field X is evaluated at time t.

Cartan magic formula (see [MR99]) states: the Lie derivative of a $k-$form $\alpha \in \Omega^k(M)$ along a vector–field $X \in \mathcal{X}^k(M)$ on a smooth manifold M is defined as

$$\mathcal{L}_X \alpha = di_X \alpha + i_X d\alpha = d(X \lrcorner \alpha) + X \lrcorner d\alpha.$$

Also, the following identities hold [MR99, KMS93]:

1. $\mathcal{L}_{fX} \alpha = f \mathcal{L}_X \alpha + df \wedge i_x \alpha$.
2. $\mathcal{L}_{[X,Y]} \alpha = \mathcal{L}_X \mathcal{L}_Y \alpha - \mathcal{L}_Y \mathcal{L}_X \alpha$.
3. $i_{[X,Y]} \alpha = \mathcal{L}_X i_Y \alpha - i_Y \mathcal{L}_X \alpha$.
4. $\mathcal{L}_X d\alpha = d\mathcal{L}_X \alpha$, i.e., $[\mathcal{L}_X, d] = 0$.
5. $\mathcal{L}_X i_X \alpha = i_X \mathcal{L}_X \alpha$, i.e., $[\mathcal{L}_X, i_X] = 0$.
6. $\mathcal{L}_X(\alpha \wedge \beta) = \mathcal{L}_X \alpha \wedge \beta + \alpha \wedge \mathcal{L}_X \beta$.

Lie Derivative of Various Tensor Fields

In this subsection, we use local coordinates x^i $(i = 1, ..., n)$ on a biomechanical $n-$manifold M, to calculate the Lie derivative $\mathcal{L}_{X^i}$ with respect to a generic vector–field X^i. (As always, $\partial_{x^i} \equiv \frac{\partial}{\partial x^i}$).

Lie Derivative of a Scalar Field

Given the scalar field ϕ, its Lie derivative $\mathcal{L}_{X^i} \phi$ is given as

$$\mathcal{L}_{X^i} \phi = X^i \partial_{x^i} \phi = X^1 \partial_{x^1} \phi + X^2 \partial_{x^2} \phi + ... + X^n \partial_{x^n} \phi.$$

Lie Derivative of Vector and Covector–Fields

Given a contravariant vector–field V^i, its Lie derivative $\mathcal{L}_{X^i} V^i$ is given as

$$\mathcal{L}_{X^i} V^i = X^k \partial_{x^k} V^i - V^k \partial_{x^k} X^i \equiv [X^i, V^i] - \text{the Lie bracket}.$$

Given a covariant vector–field (i.e., a one–form) ω_i, its Lie derivative $\mathcal{L}_{X^i} \omega_i$ is given as

$$\mathcal{L}_{X^i} \omega_i = X^k \partial_{x^k} \omega_i + \omega_k \partial_{x^i} X^k.$$

Lie Derivative of a Second–Order Tensor–Field

Given a $(2,0)$ tensor–field S^{ij}, its Lie derivative $\mathcal{L}_{X^i}S^{ij}$ is given as

$$\mathcal{L}_{X^i}S^{ij} = X^i\partial_{x^i}S^{ij} - S^{ij}\partial_{x^i}X^i - S^{ii}\partial_{x^i}X^j.$$

Given a $(1,1)$ tensor–field S^i_j, its Lie derivative $\mathcal{L}_{X^i}S^i_j$ is given as

$$\mathcal{L}_{X^i}S^i_j = X^i\partial_{x^i}S^i_j - S^i_j\partial_{x^i}X^i + S^i_i\partial_{x^j}X^i.$$

Given a $(0,2)$ tensor–field S_{ij}, its Lie derivative $\mathcal{L}_{X^i}S_{ij}$ is given as

$$\mathcal{L}_{X^i}S_{ij} = X^i\partial_{x^i}S_{ij} + S_{ij}\partial_{x^i}X^i + S_{ii}\partial_{x^j}X^i.$$

Lie Derivative of a Third–Order Tensor–Field

Given a $(3,0)$ tensor–field T^{ijk}, its Lie derivative $\mathcal{L}_{X^i}T^{ijk}$ is given as

$$\mathcal{L}_{X^i}T^{ijk} = X^i\partial_{x^i}T^{ijk} - T^{ijk}\partial_{x^i}X^i - T^{iik}\partial_{x^i}X^j - T^{iji}\partial_{x^i}X^k.$$

Given a $(2,1)$ tensor–field T^{ij}_k, its Lie derivative $\mathcal{L}_{X^i}T^{ij}_k$ is given as

$$\mathcal{L}_{X^i}T^{ij}_k = X^i\partial_{x^i}T^{ij}_k - T^{ij}_k\partial_{x^i}X^i + T^{ij}_i\partial_{x^k}X^i - T^{ii}_k\partial_{x^i}X^j.$$

Given a $(1,2)$ tensor–field T^i_{jk}, its Lie derivative $\mathcal{L}_{X^i}T^i_{jk}$ is given as

$$\mathcal{L}_{X^i}T^i_{jk} = X^i\partial_{x^i}T^i_{jk} - T^i_{jk}\partial_{x^i}X^i + T^i_{ik}\partial_{x^j}X^i + T^i_{ji}\partial_{x^k}X^i.$$

Given a $(0,3)$ tensor–field T_{ijk}, its Lie derivative $\mathcal{L}_{X^i}T_{ijk}$ is given as

$$\mathcal{L}_{X^i}T_{ijk} = X^i\partial_{x^i}T_{ijk} + T_{ijk}\partial_{x^i}X^i + T_{iik}\partial_{x^j}X^i + T_{iji}\partial_{x^k}X^i.$$

Lie Derivative of a Fourth–Order Tensor–Field

Given a $(4,0)$ tensor–field R^{ijkl}, its Lie derivative $\mathcal{L}_{X^i}R^{ijkl}$ is given as

$$\mathcal{L}_{X^i}R^{ijkl} = X^i\partial_{x^i}R^{ijkl} - R^{ijkl}\partial_{x^i}X^i - R^{iikl}\partial_{x^i}X^j - R^{ijil}\partial_{x^i}X^k - R^{ijki}\partial_{x^i}X^l.$$

Given a $(3,1)$ tensor–field R^{ijk}_l, its Lie derivative $\mathcal{L}_{X^i}R^{ijk}_l$ is given as

$$\mathcal{L}_{X^i}R^{ijk}_l = X^i\partial_{x^i}R^{ijk}_l - R^{ijk}_l\partial_{x^i}X^i + R^{ijk}_i\partial_{x^l}X^i - R^{iik}_l\partial_{x^i}X^j - R^{iji}_l\partial_{x^i}X^k.$$

Given a $(2,2)$ tensor–field R^{ij}_{kl}, its Lie derivative $\mathcal{L}_{X^i}R^{ij}_{kl}$ is given as

$$\mathcal{L}_{X^i}R^{ij}_{kl} = X^i\partial_{x^i}R^{ij}_{kl} - R^{ij}_{kl}\partial_{x^i}X^i + R^{ij}_{il}\partial_{x^k}X^i + R^{ij}_{ki}\partial_{x^l}X^i - R^{ii}_{kl}\partial_{x^i}X^j.$$

Given a $(1,3)$ tensor–field R^i_{jkl}, its Lie derivative $\mathcal{L}_{X^i}R^i_{jkl}$ is given as

$$\mathcal{L}_{X^i}R^i_{jkl} = X^i\partial_{x^i}R^i_{jkl} - R^i_{jkl}\partial_{x^i}X^i + R^i_{ikl}\partial_{x^j}X^i + R^i_{jil}\partial_{x^k}X^i + R^i_{jki}\partial_{x^l}X^i.$$

Given a $(0,4)$ tensor–field R_{ijkl}, its Lie derivative $\mathcal{L}_{X^i}R_{ijkl}$ is given as

$$\mathcal{L}_{X^i}R_{ijkl} = X^i\partial_{x^i}R_{ijkl} + R_{ijkl}\partial_{x^i}X^i + R_{iikl}\partial_{x^j}X^i + R_{ijil}\partial_{x^k}X^i + R_{ijki}\partial_{x^l}X^i.$$

Finally, recall that a *spinor* is a two–component complex column vector. Physically, spinors can describe both bosons and fermions, while tensors can describe only bosons. The Lie derivative of a spinor ϕ is defined by

$$\mathcal{L}_X\phi(x) = \lim_{t\to 0}\frac{\bar{\phi}_t(x) - \phi(x)}{t},$$

where $\bar{\phi}_t$ is the image of ϕ by a one–parameter group of isometries with X its generator. For a vector field X^a and a covariant derivative ∇_a, the Lie derivative of ϕ is given explicitly by

$$\mathcal{L}_X\phi = X^a\nabla_a\phi - \frac{1}{8}\left(\nabla_a X_b - \nabla_b X_a\right)\gamma^a\gamma^b\phi,$$

where γ^a and γ^b are *Dirac matrices* (see, e.g., [BM00]).

Lie Algebras

Recall from Introduction that an *algebra* A is a vector space with a product. The product must have the property that

$$a(uv) = (au)v = u(av),$$

for every $a \in \mathbb{R}$ and $u, v \in A$. A map $\phi : A \to A'$ between algebras is called an *algebra homomorphism* if $\phi(u \cdot v) = \phi(u)\cdot\phi(v)$. A vector subspace $\mathfrak{I}$ of an algebra A is called a *left ideal* (resp. *right ideal*) if it is closed under algebra multiplication and if $u \in A$ and $i \in \mathfrak{I}$ implies that $ui \in \mathfrak{I}$ (resp. $iu \in \mathfrak{I}$). A subspace $\mathfrak{I}$ is said to be a *two–sided ideal* if it is both a left and right ideal. An ideal may not be an algebra itself, but the quotient of an algebra by a two–sided ideal inherits an algebra structure from A.

A *Lie algebra* is an algebra A where the multiplication, i.e., the *Lie bracket* $(u, v) \mapsto [u, v]$, has the following properties:

LA 1. $[u, u] = 0$ for every $u \in A$, and

LA 2. $[u, [v, w]] + [w, [u, v]] + [v, w, u]] = 0$ for all $u, v, w \in A$.

The condition LA 2 is usually called *Jacobi identity*. A subspace $E \subset A$ of a Lie algebra is called a *Lie subalgebra* if $[u, v] \in E$ for every $u, v \in E$. A map $\phi : A \to A'$ between Lie algebras is called a *Lie algebra homomorphism* if $\phi([u, v]) = [\phi(u), \phi(v)]$ for each $u, v \in A$.

All Lie algebras (over a given field $\mathbb{K}$) and all smooth homomorphisms between them form the category $\mathcal{LAL}$, which is itself a complete subcategory of the category $\mathcal{AL}$ of all algebras and their homomorphisms.

2.4.2 Lie Groups in Human–Like Biomechanics

In the middle of the 19th century S. Lie made a far reaching discovery that techniques designed to solve particular unrelated types of ODEs, such as separable, homogeneous and exact equations, were in fact all special cases of a general form of integration procedure based on the invariance of the differential equation under a continuous group of symmetries. Roughly speaking a symmetry group of a system of differential equations is a group that transforms solutions of the system to other solutions. Once the symmetry group has been identified a number of techniques to solve and classify these differential equations becomes possible. In the classical framework of Lie, these groups were local groups and arose locally as groups of transformations on some Euclidean space. The passage from the local Lie group to the present day definition using manifolds was accomplished by E. Cartan at the end of the 19th century, whose work is a striking synthesis of Lie theory, classical geometry, differential geometry and topology.

These continuous groups, which originally appeared as symmetry groups of differential equations, have over the years had a profound impact on diverse areas such as algebraic topology, differential geometry, numerical analysis, control theory, classical mechanics, quantum mechanics etc. They are now universally known as Lie groups.

Lie Groups and Their Associated Lie Algebras

Recall that a *Lie group* is a smooth (Banach) manifold M that has at the same time a group $G-$structure consistent with its manifold $M-$structure in the sense that group multiplication

$$\mu : G \times G \to G, \qquad (g, h) \mapsto gh \tag{2.21}$$

and the inversion

$$\nu : G \to G, \qquad g \mapsto g^{-1} \tag{2.22}$$

are C^k-maps [Che55, AMR88, MR99, Put93]. A point $e \in G$ is called the group *identity* element.

For example, any finite–dimensional Banach vector space V is an Abelian Lie group with group operations $\mu : V \times V \to V$, $\mu(x, y) = x + y$, and $\nu : V \to V$, $\nu(x) = -x$. The identity is just the zero vector. We call such a Lie group a *vector group*.

Let G and H be two Lie groups. A map $G \to H$ is said to be a *morphism* of Lie groups (or their *smooth homomorphism*) if it is their homomorphism as abstract groups and their smooth map as manifolds [Pos86].

All Lie groups and all their morphisms form the category $\mathcal{LG}$ (more precisely, there is a countable family of categories $\mathcal{LG}$ depending on C^k-smoothness of the corresponding manifolds).

Similarly, a group G which is at the same time a topological space is said to be a *topological group* if maps (2.21–2.22) are continuous, i.e., C^0-maps for it. The homomorphism $G \to H$ of topological groups is said to be continuous if it is a continuous map. Topological groups and their continuous homomorphisms form the category $\mathcal{TG}$.

A topological group (as well as a smooth manifold) is not necessarily Hausdorff. A topological group G is Hausdorff iff its identity is closed. As a corollary we have that every Lie group is a Hausdorff topological group (see [Pos86]).

For every g in a Lie group G, the two maps,

$$\begin{aligned} L_g &: G \to G, \qquad h \mapsto gh, \qquad \text{and} \\ R_h &: G \to G, \qquad g \mapsto gh, \end{aligned}$$

are called *left* and *right translation* maps. Since $L_g \circ L_h = L_{gh}$, and $R_g \circ R_h = R_{gh}$, it follows that $(L_g)^{-1} = L_{g^{-1}}$ and $(R_g)^{-1} = R_{g^{-1}}$, so both L_g and R_g are diffeomorphisms. Moreover $L_g \circ R_h = R_h \circ L_g$, i.e., left and right translation commute.

A vector–field X on G is called *left invariant vector–field* if for every $g \in G$, $L_g^* X = X$, that is, if $(T_h L_g) X(h) = X(gh)$ for all $h \in G$, i.e., the following diagram commutes:

$$\begin{tikzcd} TG \arrow[r, "TL_g"] & TG \\ G \arrow[u, "X"] \arrow[r, "L_g"'] & G \arrow[u, "X"'] \end{tikzcd}$$

The correspondences $G \to TG$ and $L_g \to TL_g$ obviously define a functor $\mathcal{F} : \mathcal{LG} \Rightarrow \mathcal{LG}$ from the category G of Lie groups to itself. $\mathcal{F}$ is another special case of the vector bundle functor (2.2.3).

Let $\mathcal{X}_L(G)$ denote the set of left invariant vector–fields on G; it is a Lie subalgebra of $\mathcal{X}(G)$, the set of all vector–fields on G, since $L_g^*[X, Y] = [L_g^* X, L_g^* Y] = [X, Y]$, so the Lie bracket $[X, Y] \in \mathcal{X}_L(G)$.

Let e be the identity element of G. Then for each ξ on the tangent space $T_e G$ we define a vector–field X_ξ on G by

$$X_\xi(g) = T_e L_g(\xi).$$

$\mathcal{X}_L(G)$ and $T_e G$ are isomorphic as vector spaces. Define the Lie bracket on $T_e G$ by

$$[\xi, \eta] = [X_\xi, X_\eta](e),$$

for all $\xi, \eta \in T_e G$. This makes $T_e G$ into a Lie algebra. Also, by construction, we have

$$[X_\xi, X_\eta] = X_{[\xi, \eta]},$$

this defines a bracket in T_eG via *left extension*. The vector space T_eG with the above algebra structure is called the Lie algebra of the Lie group G and is denoted $\mathfrak{g}$.

For example, let V be a finite–dimensional vector space. Then $T_eV \simeq V$ and the left invariant vector–field defined by $\xi \in T_eV$ is the constant vector–field $X_\xi(\eta) = \xi$, for all $\eta \in V$. The Lie algebra of V is V itself.

Since any two elements of an Abelian Lie group G commute, it follows that all adjoint operators Ad_g, $g \in G$, equal the identity. Therefore, the Lie algebra g is Abelian; that is, $[\xi, \eta] = 0$ for all $\xi, \eta \in \mathfrak{g}$ [MR99].

Recall (2.4.1) that Lie algebras and their smooth homomorphisms form the category $\mathcal{LAL}$. We can now introduce the fundamental *Lie functor*, $\mathcal{F} : \mathcal{LG} \Rightarrow \mathcal{LAL}$, from the category of Lie groups to the category of Lie algebras [Pos86].

Let X_ξ be a left invariant vector–field on G corresponding to ξ in $\mathfrak{g}$. Then there is a unique integral curve $\gamma_\xi : \mathbb{R} \to G$ of X_ξ starting at e, i.e.,

$$\dot{\gamma}_\xi(t) = X_\xi\left(\gamma_\xi(t)\right), \qquad \gamma_\xi(0) = e.$$

$\gamma_\xi(t)$ is a smooth *one parameter subgroup* of G, i.e.,

$$\gamma_\xi(t+s) = \gamma_\xi(t) \cdot \gamma_\xi(s),$$

since, as functions of t both sides equal $\gamma_\xi(s)$ at $t = 0$ and both satisfy differential equation

$$\dot{\gamma}(t) = X_\xi\left(\gamma_\xi(t)\right)$$

by left invariance of X_ξ, so they are equal. Left invariance can be also used to show that $\gamma_\xi(t)$ is defined for all $t \in \mathbb{R}$. Moreover, if $\phi : \mathbb{R} \to G$ is a one parameter subgroup of G, i.e., a *smooth homomorphism* of the additive group $\mathbb{R}$ into G, then $\phi = \gamma_\xi$ with $\xi = \dot{\phi}(0)$, since taking derivative at $s = 0$ in the relation

$$\phi(t+s) = \phi(t) \cdot \phi(s) \qquad \text{gives} \qquad \dot{\phi}(t) = X_{\dot{\phi}(0)}\left(\phi(t)\right),$$

so $\phi = \gamma_\xi$ since both equal e at $t = 0$. Therefore, all one parameter subgroups of G are of the form $\gamma_\xi(t)$ for some $\xi \in \mathfrak{g}$.

The map $\exp : \mathfrak{g} \to G$, given by

$$\exp(\xi) = \gamma_\xi(1), \qquad \exp(0) = e, \tag{2.23}$$

is called the *exponential map* of the Lie algebra $\mathfrak{g}$ of G into G. exp is a C^k–map, similar to the projection π of tangent and cotangent bundles; exp is locally a diffeomorphism from a neighborhood of zero in $\mathfrak{g}$ onto a neighborhood of e in G; if $f : G \to H$ is a smooth homomorphism of Lie groups, then

$$f \circ \exp_G = \exp_H \circ T_e f\,.$$

Also, in this case (see [Che55, MR99, Pos86])

$$\exp(s\xi) = \gamma_\xi(s).$$

Indeed, for fixed $s \in \mathbb{R}$, the curve $t \mapsto \gamma_\xi(ts)$, which at $t = 0$ passes through e, satisfies the differential equation

$$\frac{d}{dt}\gamma_\xi(ts) = sX_\xi\left(\gamma_\xi(ts)\right) = X_{s\xi}\left(\gamma_\xi(ts)\right).$$

Since $\gamma_{s\xi}(t)$ satisfies the same differential equation and passes through e at $t = 0$, it follows that $\gamma_{s\xi}(t) = \gamma_\xi(st)$. Putting $t = 1$ yields $\exp(s\xi) = \gamma_\xi(s)$ [MR99].

Hence exp maps the line $s\xi$ in $\mathfrak{g}$ onto the one–parameter subgroup $\gamma_\xi(s)$ of G, which is tangent to ξ at e. It follows from left invariance that the flow F_t^ξ of X satisfies $F_t^\xi(g) = g\exp(s\xi)$.

Globally, the exponential map exp, as given by (2.23), is a natural operation, i.e., for any morphism $\varphi : G \to H$ of Lie groups G and H and a Lie functor $\mathcal{F}$, the following diagram commutes [Pos86]:

$$\begin{tikzcd}
\mathcal{F}(G) \arrow[r, "\mathcal{F}(\varphi)"] \arrow[d, "\exp"'] & \mathcal{F}(H) \arrow[d, "\exp"] \\
G \arrow[r, "\varphi"'] & H
\end{tikzcd}$$

Let G_1 and G_2 be Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$. Then $G_1 \times G_2$ is a Lie group with Lie algebra $\mathfrak{g}_1 \times \mathfrak{g}_2$, and the exponential map is given by [MR99].

$$\exp : \mathfrak{g}_1 \times \mathfrak{g}_2 \to G_1 \times G_2, \qquad (\xi_1, \xi_2) \mapsto (\exp_1(\xi_1), \exp_2(\xi_2)).$$

For example, in case of a finite–dimensional vector space, or infinite–dimensional Banach space, the exponential map is the identity.

The unit circle in the complex plane $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ is an Abelian Lie group under multiplication. The tangent space T_eS^1 is the imaginary axis, and we identify $\mathbb{R}$ with T_eS^1 by $t \mapsto 2\pi it$. With this identification, the exponential map $\exp : \mathbb{R} \to S^1$ is given by $\exp(t) = e^{2\pi it}$.

The nD torus $T^n = S^1 \times \cdots \times S^1$ (n times) is an Abelian Lie group. The exponential map $\exp : \mathbb{R}^n \to T^n$ is given by

$$\exp(t_1, ..., t_n) = (e^{2\pi it_1}, ..., e^{2\pi it_n}).$$

Since $S^1 = \mathbb{R}/\mathbb{Z}$, it follows that

$$T^n = \mathbb{R}^n/\mathbb{Z}^n,$$

the projection $\mathbb{R}^n \to T^n$ being given by the exp map (see [MR99, Pos86]).

For every $g \in G$, the map

$$Ad_g = T_e \left(R_{g^{-1}} \circ L_g \right) : \mathfrak{g} \to \mathfrak{g}$$

is called the *adjoint map* (or *operator*) associated with g.

For each $\xi \in \mathfrak{g}$ and $g \in G$ we have

$$\exp \left(Ad_g \xi \right) = g \left(\exp \xi \right) g^{-1}.$$

The relation between the adjoint map and the Lie bracket is the following: For all $\xi, \eta \in \mathfrak{g}$ we have

$$\left. \frac{d}{dt} \right|_{t=0} Ad_{\exp(t\xi)} \eta = [\xi, \eta].$$

A Lie subgroup H of G is a subgroup H of G which is also a submanifold of G. Then $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$ and moreover $\mathfrak{h} = \{\xi \in \mathfrak{g} | \exp(t\xi) \in H$, for all $t \in \mathbb{R}\}$.

One can characterize Lebesgue measure up to a multiplicative constant on $\mathbb{R}^n$ by its invariance under translations. Similarly, on a locally compact group there is a unique (up to a nonzero multiplicative constant) left–invariant measure, called *Haar measure*. For Lie groups the existence of such measures is especially simple [MR99]: Let G be a Lie group. Then there is a volume form $Ub5$, unique up to nonzero multiplicative constants, that is left invariant. If G is compact, $Ub5$ is right invariant as well.

Actions of Lie Groups on M

Let M be a smooth manifold. An action of a Lie group G (with the unit element e) on M is a smooth map $\phi : G \times M \to M$, such that for all $x \in M$ and $g, h \in G$, (i) $\phi(e, x) = x$ and (ii) $\phi\left(g, \phi(h, x)\right) = \phi(gh, x)$. In other words, letting $\phi_g : x \in M \mapsto \phi_g(x) = \phi(g, x) \in M$, we have (i') $\phi_e = id_M$ and (ii') $\phi_g \circ \phi_h = \phi_{gh}$. ϕ_g is a diffeomorphism, since $(\phi_g)^{-1} = \phi_{g^{-1}}$. We say that the map $g \in G \mapsto \phi_g \in Diff(M)$ is a homomorphism of G into the group of diffeomorphisms of M. In case that M is a vector space and each ϕ_g is a linear operator, the function of G on M is called a representation of G on M [Put93]

An action ϕ of G on M is said to be *transitive action*, if for every $x, y \in M$, there is $g \in G$ such that $\phi(g, x) = y$; *effective action*, if $\phi_g = id_M$ implies $g = e$, that is $g \mapsto \phi_g$ is one–to–one; and *free action*, if for each $x \in M$, $g \mapsto \phi_g(x)$ is one–to–one.

For example,

1. $G = \mathbb{R}$ acts on $M = \mathbb{R}$ by translations; explicitly,

$$\phi : G \times M \to M, \qquad \phi(s, x) = x + s.$$

Then for $x \in \mathbb{R}$, $O_x = \mathbb{R}$. Hence M/G is a single point, and the action is transitive and free.

2. A complete flow ϕ_t of a vector–field X on M gives an action of $\mathbb{R}$ on M, namely
$$(t, x) \in \mathbb{R} \times M \mapsto \phi_t(x) \in M.$$
3. Left translation $L_g : G \to G$ defines an effective action of G on itself. It is also transitive.
4. The coadjoint action of G on $\mathfrak{g}^*$ is given by
$$Ad^* : (g, \alpha) \in G \times \mathfrak{g}^* \mapsto Ad^*_{g^{-1}}(\alpha) = \left(T_e(R_{g^{-1}} \circ L_g)\right)^* \alpha \in \mathfrak{g}^*.$$

Let ϕ be an action of G on M. For $x \in M$ the *orbit* of x is defined by
$$O_x = \{\phi_g(x) | g \in G\} \subset M$$
and the *isotropy group* of ϕ at x is given by
$$G_x = \{g \in G | \phi(g, x) = x\} \subset G.$$

An action ϕ of G on a manifold M defines an equivalence relation on M by the relation belonging to the same orbit; explicitly, for $x, y \in M$, we write $x \sim y$ if there exists a $g \in G$ such that $\phi(g, x) = y$, that is, if $y \in O_x$. The set of all orbits M/G is called the *orbit space*.

For example, let $M = \mathbb{R}^2 \backslash \{0\}$, $G = SO(2)$, the group of rotations in plane, and the action of G on M given by
$$\left(\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, (x, y) \right) \longmapsto (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta).$$
The action is always free and effective, and the orbits are concentric circles, thus the orbit space is $M/G \simeq \mathbb{R}^*_+$.

A crucial concept in mechanics is the infinitesimal description of an action. Let $\phi : G \times M \to M$ be an action of a Lie group G on a smooth manifold M. For each $\xi \in \mathfrak{g}$,
$$\phi_\xi : \mathbb{R} \times M \to M, \qquad \phi_\xi(t, x) = \phi\left(\exp(t\xi), x\right)$$
is an $\mathbb{R}$–action on M. Therefore, $\phi_{\exp(t\xi)} : M \to M$ is a flow on M; the corresponding vector–field on M, given by
$$\xi_M(x) = \left.\frac{d}{dt}\right|_{t=0} \phi_{\exp(t\xi)}(x)$$
is called the infinitesimal generator of the action, corresponding to ξ in $\mathfrak{g}$.

The tangent space at x to an orbit O_x is given by
$$T_x O_x = \{\xi_M(x) | \xi \in \mathfrak{g}\}.$$

Let $\phi : G \times M \to M$ be a smooth G–action. For all $g \in G$, all $\xi, \eta \in \mathfrak{g}$ and all $\alpha, \beta \in \mathbb{R}$, we have:

$(Ad_g\xi)_M = \phi^*_{g^{-1}}\xi_M$, $[\xi_M, \eta_M] = -[\xi, \eta]_M$, and $(\alpha\xi+\beta\eta)_M = \alpha\xi_M + \beta\eta_M$.

Let M be a smooth manifold, G a Lie group and $\phi : G \times M \to M$ a G–action on M. We say that a smooth map $f : M \to M$ is with respect to this action if for all $g \in G$,

$$f \circ \phi_g = \phi_g \circ f.$$

Let $f : M \to M$ be an equivariant smooth map. Then for any $\xi \in \mathfrak{g}$ we have

$$Tf \circ \xi_M = \xi_M \circ f.$$

Cohomology of Lie Groups

E. Cartan only studied real cohomology, using the De Rham theorems (see Chapter 4). Let G be a compact Lie group, operating on the right on a C^k manifold M by a C^k-operation $(s, x) \mapsto x \cdot s$. Since there exists a measure ds on G, invariant by left and right translations and of total mass 1, Hurewicz's idea of taking *mean values* on G of an arbitrary exterior $p-$form α on M may be applied: for any point $x \in M$, the mean value $m(\alpha)$ of a $p-$form α takes the value

$$m(\alpha)(x) = \int_G \alpha(x \cdot s)\, ds.$$

Now $m(\alpha)$ is *invariant* under the action of G on M, and if α is closed (resp. exact), then $m(\alpha)$ is also closed (resp. exact). Furthermore, α and $m(\alpha)$ are *cohomologous* on M; if $H^\bullet_G(M)$ is the graded subspace of the real cohomology space $H^\bullet(M)$, consisting of the classes of the differential forms invariant under the action of G, this defines an *isomorphism*

$$m^* : H^\bullet(M) \xrightarrow{\sim} H^\bullet_G(M).$$

Cartan's interpretation of the real cohomology $H^\bullet(G)$ of a compact Lie group G is obtained as a corollary by consideration of the action $((s,t),x) \mapsto s^{-1}xt$ of $G \times G$ on G. A $p-$form is invariant under the action if it is *bi-invariant*, that is, invariant under both left and right translations in G. The Lie–Cartan theory implies that for such a form α, $d\alpha = 0$, so that when one computes $H^\bullet_{G\times G}(G)$, all cochains are cocycles and all coboundaries are 0. Hence the fundamental result that $H^\bullet(G)$ *is isomorphic to the graded algebra* $b^*(G)$ of all *bi–invariant differential forms*.

The explicit determination of $H^\bullet(G)$ is thus reduced to an algebraic problem. The group G operates on the dual $\mathfrak{g}^*$ of the Lie algebra $\mathfrak{g}$ by the coadjoint representation $s \mapsto {}^tAd(s)$; $b^*(G)$ is the sum of the 1D subspaces of $\mathfrak{g}^*$ stable for that representation; for a compact group, they can in principle be determined by Cartan's method of *highest weights*.

Basic Biomechanical Groups

Galilei Group

The *Galilei group* is the group of transformations in space and time that connect those Cartesian systems that are termed 'inertial frames' in Newtonian mechanics. The most general relationship between two such frames is the following. The origin of the time scale in the inertial frame S' may be shifted compared with that in S; the orientation of the Cartesian axes in S' may be different from that in S; the origin O of the Cartesian frame in S' may be moving relative to the origin O in S at a uniform velocity. The transition from S to S' involves ten parameters; thus the Galilei group is a ten parameter group. The basic assumption inherent in Galilei–Newtonian relativity is that there is an absolute time scale, so that the only way in which the time variables used by two different 'inertial observers' could possibly differ is that the zero of time for one of them may be shifted relative to the zero of time for the other.

Galilei space–time structure involves the following three elements:

1. *World*, as a 4D affine space A^4. The points of A^4 are called *world points* or *events*. The parallel transitions of the world A^4 form a linear (i.e., Euclidean) space $\mathbb{R}^4$.
2. *Time*, as a linear map $t : \mathbb{R}^4 \to \mathbb{R}$ of the linear space of the world parallel transitions onto the real 'time axes'. Time interval from the event $a \in A^4$ to $b \in A^4$ is called the number $t(b-a)$; if $t(b-a) = 0$ then the events a and b are called synchronous. The set of all mutually synchronous events consists a 3D affine space A^3, being a subspace of the world A^4. The kernel of the mapping t consists of the parallel transitions of A^4 translating arbitrary (and every) event to the synchronous one; it is a linear 3D subspace $\mathbb{R}^3$ of the space $\mathbb{R}^4$.
3. *Distance* (*metric*) between the synchronous events,
$$\rho(a, b) = \parallel a - b \parallel, \qquad \text{for all} \quad a, b \in A^3,$$
given by the scalar product in $\mathbb{R}^3$. The distance transforms arbitrary space of synchronous events into the well known 3D Euclidean space E^3.

The space A^4, with the Galilei space–time structure on it, is called Galilei space. Galilei group is the group of all possible transformations of the Galilei space, preserving its structure. The elements of the Galilei group are called Galilei transformations. Therefore, Galilei transformations are affine transformations of the world A^4 preserving the time intervals and distances between the synchronous events.

The direct product $\mathbb{R} \times \mathbb{R}^3$, of the time axes with the 3D linear space R3 with a fixed Euclidean structure, has a natural Galilei structure. It is called Galilei coordinate system.

General Linear Group

The group of linear isomorphisms of $\mathbb{R}^n$ to $\mathbb{R}^n$ is a Lie group of dimension n^2, called the general linear group and denoted $Gl(n,\mathbb{R})$. It is a smooth manifold, since it is a subset of the vector space $L(\mathbb{R}^n,\mathbb{R}^n)$ of all linear maps of $\mathbb{R}^n$ to $\mathbb{R}^n$, as $Gl(n,\mathbb{R})$ is the inverse image of $\mathbb{R}\backslash\{0\}$ under the continuous map $A \mapsto \det A$ of $L(\mathbb{R}^n,\mathbb{R}^n)$ to $\mathbb{R}$. The group operation is composition

$$(A,B) \in Gl(n,\mathbb{R}) \times Gl(n,\mathbb{R}) \mapsto A \circ B \in Gl(n,\mathbb{R})$$

and the inverse map is

$$A \in Gl(n,\mathbb{R}) \mapsto A^{-1} \in Gl(n,\mathbb{R}).$$

If we choose a basis in $\mathbb{R}^n$, we can represent each element $A \in Gl(n,\mathbb{R})$ by an invertible $(n \times n)$–matrix. The group operation is then matrix multiplication and the inversion is matrix inversion. The identity is the identity matrix I_n. The group operations are smooth since the formulas for the product and inverse of matrices are smooth in the matrix components.

The Lie algebra of $Gl(n,\mathbb{R})$ is $\mathfrak{gl}(n)$, the vector space $L(\mathbb{R}^n,\mathbb{R}^n)$ of all linear transformations of $\mathbb{R}^n$, with the commutator bracket

$$[A,B] = AB - BA.$$

For every $A \in L(\mathbb{R}^n,\mathbb{R}^n)$,

$$\gamma_A : t \in \mathbb{R} \mapsto \gamma_A(t) = \sum_{i=0}^{\infty} \frac{t^i}{i!} A^i \in Gl(n,\mathbb{R})$$

is a one parameter subgroup of $Gl(n,\mathbb{R})$, because

$$\gamma_A(0) = I, \qquad \text{and} \qquad \dot{\gamma}_A(t) = \sum_{i=0}^{\infty} \frac{t^{i-1}}{(i-1)!} A^i = \gamma_A(t)\, A.$$

Hence γ_A is an integral curve of the left invariant vector–field X_A. Therefore, the exponential map is given by

$$\exp : A \in L(\mathbb{R}^n,\mathbb{R}^n) \mapsto \exp(A) \equiv e^A = \gamma_A(1) = \sum_{i=0}^{\infty} \frac{A^i}{i!} \in Gl(n,\mathbb{R}).$$

For each $A \in Gl(n,\mathbb{R})$ the corresponding adjoint map

$$Ad_A : L(\mathbb{R}^n,\mathbb{R}^n) \to L(\mathbb{R}^n,\mathbb{R}^n)$$

is given by

$$Ad_A B = A \cdot B \cdot A^{-1}.$$

Groups of Joint Rotations

Local kinematics at each rotational robot or (synovial) human joint, is defined as a *group action* of an nD constrained rotational Lie group $SO(n)$ on the Euclidean space $\mathbb{R}^n$. In particular, there is an action of $SO(2)$–group in uniaxial human joints (cylindrical, or *hinge joints*, like knee and elbow) and an action of $SO(3)$–group in three–axial human joints (spherical, or *ball–and–socket joints*, like hip, shoulder, neck, wrist and ankle). In both cases, $SO(n)$ acts, with its operators of rotation, on the vector $x = \{x^\mu\}$, $(i = 1, 2, 3)$ of external, Cartesian coordinates of the parent body–segment, depending, at the same time, on the vector $q = \{q^s\}$, $(s = 1, \cdots, n)$ on n group–parameters, i.e., joint angles.

Each joint rotation $R \in SO(n)$ defines a map

$$R : x^\mu \mapsto \dot{x}^\mu, \qquad R(x^\mu, q^s) = R_{q^s} x^\mu,$$

where $R_{q^s} \in SO(n)$ are joint group operators. The vector $v = \{v_s\}$, $(s = 1, \cdots, n)$ of n infinitesimal generators of these rotations, i.e., joint angular velocities, given by

$$v_s = -[\frac{\partial R(x^\mu, q^s)}{\partial q^s}]_{q=0} \frac{\partial}{\partial x^\mu},$$

constitute an nD Lie algebra $\mathfrak{so}(n)$ corresponding to the joint rotation group $SO(n)$. Conversely, each joint group operator R_{q^s}, representing a one–parameter subgroup of $SO(n)$, is defined as the exponential map of the corresponding joint group generator v_s

$$R_{q^s} = \exp(q^s v_s). \tag{2.24}$$

The exponential map (2.24) represents a solution of the joint operator differential equation in the joint group–parameter space $\{q^s\}$

$$\frac{dR_{q^s}}{dq^s} = v_s R_{q^s}.$$

Uniaxial Group of Joint Rotations

The uniaxial joint rotation in a single Cartesian plane around a perpendicular axis, e.g., $xy-$plane about the z axis, by an internal joint angle θ, leads to the following transformation of the joint coordinates

$$\dot{x} = x \cos\theta - y \sin\theta, \qquad \dot{y} = x \sin\theta + y \cos\theta.$$

In this way, the joint $SO(2)-$group, given by

$$SO(2) = \left\{ R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} | \theta \in [0, 2\pi] \right\},$$

acts in a canonical way on the Euclidean plane $\mathbb{R}^2$ by

$$SO(2) = \left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \begin{pmatrix} x \\ w \end{pmatrix} \right) \longmapsto \begin{pmatrix} x\cos\theta & -y\sin\theta \\ x\sin\theta & y\cos\theta \end{pmatrix}.$$

Its associated Lie algebra $\mathfrak{so}(2)$ is given by

$$\mathfrak{so}(2) = \left\{ \begin{pmatrix} 0 & -t \\ t & 0 \end{pmatrix} | t \in \mathbb{R} \right\},$$

since the curve $\gamma_\theta \in SO(2)$ given by

$$\gamma_\theta : t \in \mathbb{R} \longmapsto \gamma_\theta(t) = \begin{pmatrix} \cos t\theta & -\sin t\theta \\ \sin t\theta & \cos t\theta \end{pmatrix} \in SO(2),$$

passes through the identity $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and then

$$\left.\frac{d}{dt}\right|_{t=0} \gamma_\theta(t) = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix},$$

so that I_2 is a basis of $\mathfrak{so}(2)$, since $\dim(SO(2)) = 1$.

The *exponential map* $\exp : \mathfrak{so}(2) \to SO(2)$ is given by

$$\exp \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} = \gamma_\theta(1) = \begin{pmatrix} \cos t\theta & -\sin t\theta \\ \sin t\theta & \cos t\theta \end{pmatrix}.$$

The *infinitesimal generator* of the action of $SO(2)$ on $\mathbb{R}^2$, i.e., joint angular velocity v, is given by

$$v = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y},$$

since

$$v_{\mathbb{R}^2}(x, y) = \left.\frac{d}{dt}\right|_{t=0} \exp(tv)(x, y) = \left.\frac{d}{dt}\right|_{t=0} \begin{pmatrix} \cos tv & -\sin tv \\ \sin tv & \cos tv \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

The *momentum map* (see subsection 2.6.3 below) $J : T^*\mathbb{R}^2 \to \mathbb{R}$ associated to the lifted action of $SO(2)$ on $T^*\mathbb{R}^2 \simeq \mathbb{R}^4$ is given by

$$J(x, y, p_1, p_2) = xp_y - yp_x, \qquad \text{since}$$
$$J(x, y, p_x, p_y)(\xi) = (p_x dx + p_y dy)(v_{\mathbb{R}^2}) = -vp_x y + -vp_y x.$$

The Lie group $SO(2)$ acts on the symplectic manifold $(\mathbb{R}^4, \omega = dp_x \wedge dx + dp_y \wedge dx)$ by

$$\phi\left(\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, (x, y, p_x, p_y) \right)$$
$$= (x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta,\ p_x\cos\theta - p_y\sin\theta,\ p_x\sin\theta + p_y\cos\theta).$$

Three–Axial Group of Joint Rotations

The three–axial $SO(3)-$group of human–like joint rotations depends on three parameters, Euler joint angles $q^i = (\varphi, \psi, \theta)$, defining the rotations about the Cartesian coordinate triedar (x, y, z) placed at the joint pivot point. Each of the Euler angles are defined in the constrained range $(-\pi, \pi)$, so the joint group space is a constrained sphere of radius π.

Let $G = SO(3) = \{A \in \mathcal{M}_{3\times 3}(\mathbb{R}) : A^t A = I_3, \det(A) = 1\}$ be the group of rotations in $\mathbb{R}^3$. It is a Lie group and $\dim(G) = 3$. Let us isolate its one–parameter joint subgroups, i.e., consider the three operators of the finite joint rotations $R_\varphi, R_\psi, R_\theta \in SO(3)$, given by

$$R_\varphi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{bmatrix}, \quad R_\psi = \begin{bmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{bmatrix}, \quad R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

corresponding respectively to rotations about $x-$axis by an angle φ, about $y-$axis by an angle ψ, and about $z-$axis by an angle θ.

The total three–axial joint rotation A is defined as the product of above one–parameter rotations $R_\varphi, R_\psi, R_\theta$, i.e., $A = R_\varphi \quad R_\psi \quad R_\theta$ is equal

$$A = \begin{bmatrix} \cos\psi\cos\varphi - \cos\theta\sin\varphi\sin\psi & \cos\psi\cos\varphi + \cos\theta\cos\varphi\sin\psi & \sin\theta\sin\psi \\ -\sin\psi\cos\varphi - \cos\theta\sin\varphi\sin\psi & -\sin\psi\sin\varphi + \cos\theta\cos\varphi\cos\psi & \sin\theta\cos\psi \\ \sin\theta\sin\varphi & -\sin\theta\cos\varphi & \cos\theta \end{bmatrix}.$$

R_φ, R_ψ andR_θ are curves in $SO(3)$ starting from $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Their derivatives in $\varphi = 0, \psi = 0$ and $\theta = 0$ belong to the associated *tangent Lie algebra* $\mathfrak{so}(3)$. That is the corresponding infinitesimal generators of joint rotations – joint angular velocities $v_\varphi, v_\psi, v_\theta \in \mathfrak{so}(3)$ – are respectively given by

$$v_\varphi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = -y\frac{\partial}{\partial z} + z\frac{\partial}{\partial y}, \qquad v_\psi = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} = -z\frac{\partial}{\partial x} + x\frac{\partial}{\partial z},$$

$$v_\theta = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = -x\frac{\partial}{\partial y} + y\frac{\partial}{\partial x}.$$

Moreover, the elements are linearly independent and so

$$\mathfrak{so}(3) = \left\{ \begin{bmatrix} 0 & -a & b \\ a & 0 & -\gamma \\ -b & \gamma & 0 \end{bmatrix} | a, b, \gamma \in \mathbb{R} \right\}.$$

the Lie algebra $\mathfrak{so}(3)$ is identified with $\mathbb{R}^3$ by associating to each $v = (v_\varphi, v_\psi, v_\theta) \in \mathbb{R}^3$ the matrix $v \in \mathfrak{so}(3)$ given by $v = \begin{bmatrix} 0 & -a & b \\ a & 0 & -\gamma \\ -b & \gamma & 0 \end{bmatrix}$. Then we have the following identities:

1. $\widehat{u \times v} = [\hat{u}, v]$; and
2. $\mathrm{u} \cdot v = -\frac{1}{2}\,\mathrm{Tr}(\hat{u} \cdot v)$.

The exponential map $\exp : \mathfrak{so}(3) \to SO(3)$ is given by *Rodrigues relation*

$$\exp(v) = I + \frac{\sin \|v\|}{\|v\|} v + \frac{1}{2} \left(\frac{\sin \frac{\|v\|}{2}}{\frac{\|v\|}{2}} \right)^2 v^2,$$

where the norm $\|v\|$ is given by

$$\|v\| = \sqrt{(v^1)^2 + (v^2)^2 + (v^3)^2}.$$

The the dual, *cotangent Lie algebra* $\mathfrak{so}(3)^*$, includes the three joint angular momenta $p_\varphi, p_\psi, p_\theta \in \mathfrak{so}(3)^*$, derived from the joint velocities v by multiplying them with corresponding moments of inertia.

Special Euclidean Groups of Total Joint Motions

Biomechanically realistic joint movement is predominantly rotational, plus restricted translational (translational motion in human joints is observed after reaching the limit of rotational amplitude). Gross translation in any human joint means joint dislocation, which is a severe injury. Obvious models for uniaxial and triaxial joint motions are *special Euclidean groups of rigid body motions*, $SE(2)$ and $SE(3)$, respectively.

Special Euclidean Group in the Plane

The motion in uniaxial human joints is naturally modelled by the *special Euclidean group in the plane*, $SE(2)$. It consists of all transformations of $\mathbb{R}^2$ of the form $Az + a$, where $z, a \in \mathbb{R}^2$, and

$$A \in SO(2) = \left\{ \text{matrices of the form } \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \right\}.$$

In other words [MR99], group $SE(2)$ consists of matrices of the form: $(R_\theta, a) = \begin{pmatrix} R_\theta & a \\ 0 & I \end{pmatrix}$, where $a \in \mathbb{R}^2$ and R_θ is the rotation matrix: $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, while I is the $3{\times}3$ identity matrix. The inverse $(R_\theta, a)^{-1}$ is given by

$$(R_\theta, a)^{-1} = \begin{pmatrix} R_\theta & a \\ 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} R_{-\theta} & -R_{-\theta} a \\ 0 & I \end{pmatrix}.$$

The Lie algebra $\mathfrak{se}(2)$ of $SE(2)$ consists of 3×3 block matrices of the form

$$\begin{pmatrix} -\xi J & v \\ 0 & 0 \end{pmatrix}, \qquad \text{where} \qquad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad (J^T = J^{-1} = -J),$$

with the usual commutator bracket. If we identify $\mathfrak{se}(2)$ with $\mathbb{R}^3$ by the isomorphism

$$\begin{pmatrix} -\xi J & v \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(2) \longmapsto (\xi, v) \in \mathbb{R}^3,$$

then the expression for the Lie algebra bracket becomes

$$[(\xi, v_1, v_2), (\zeta, w_1, w_2)] = (0, \zeta v_2 - \xi w_2, \xi w_1 - \zeta v1) = (0, \xi J^T w - \zeta J^T v),$$

where $v = (v_1, v_2)$ and $w = (w_1, w_2)$.

The *adjoint group action* of

$$(R_\theta, a) \begin{pmatrix} R_\theta & a \\ 0 & I \end{pmatrix} \qquad \text{on} \qquad (\xi, v) = \begin{pmatrix} -\xi J & v \\ 0 & 0 \end{pmatrix}$$

is given by the *group conjugation*,

$$\begin{pmatrix} R_\theta & a \\ 0 & I \end{pmatrix} \begin{pmatrix} -\xi J & v \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R_{-\theta} & -R_{-\theta} a \\ 0 & I \end{pmatrix} = \begin{pmatrix} -\xi J & \xi J a + R_\theta v \\ 0 & 0 \end{pmatrix},$$

or, in coordinates [MR99],

$$Ad_{(R_\theta, a)}(\xi, v) = (\xi, \xi J a + R_\theta v). \tag{2.25}$$

In proving (2.25) we used the identity $R_\theta J = J R_\theta$. Identify the dual algebra, $\mathfrak{se}(2)^*$, with matrices of the form $\begin{pmatrix} \frac{\mu}{2} J & 0 \\ \alpha & 0 \end{pmatrix}$, via the nondegenerate pairing given by the trace of the product. Thus, $\mathfrak{se}(2)^*$ is isomorphic to $\mathbb{R}^3$ via

$$\begin{pmatrix} \frac{\mu}{2} J & 0 \\ \alpha & 0 \end{pmatrix} \in \mathfrak{se}(2)^* \longmapsto (\mu, \alpha) \in \mathbb{R}^3,$$

so that in these coordinates, the pairing between $\mathfrak{se}(2)^*$ and $\mathfrak{se}(2)$ becomes

$$\langle (\mu, \alpha), (\xi, v) \rangle = \mu \xi + \alpha \cdot v,$$

that is, the usual dot product in $\mathbb{R}^3$. The *coadjoint group action* is thus given by

$$Ad^*_{(R_\theta, a)^{-1}}(\mu, \alpha) = (\mu - R_\theta \alpha \cdot J a + R_\theta \alpha). \tag{2.26}$$

Formula (2.26) shows that the coadjoint orbits are the cylinders $T^* S^1_\alpha = \{(\mu, \alpha) | \, \|\alpha\| = \text{const}\}$ if $\alpha \neq 0$ together with the points are on the $\mu-$axis. The canonical cotangent bundle projection $\pi : T^* S^1_\alpha \to S^1_\alpha$ is defined as $\pi(\mu, \alpha) = \alpha$.

Special Euclidean Group in the 3D Space

The most common group structure in human–like biomechanics is the *special Euclidean group in 3D space*, $SE(3)$. It is defined as a semidirect (noncommutative) product of 3D rotations and 3D translations, $SO(3) \rhd \mathbb{R}^3$.

The Heavy Top

As a starting point consider a rigid body (see (3.2.1) below) moving with a fixed point but under the influence of gravity. This problem still has a configuration space $SO(3)$, but the symmetry group is only the circle group S^1, consisting of rotations about the direction of gravity. One says that gravity has broken the symmetry from $SO(3)$ to S^1. This time, eliminating the S^1 symmetry mysteriously leads one to the larger Euclidean group $SE(3)$ of rigid motion of $\mathbb{R}^3$. Conversely, we can start with $SE(3)$ as the configuration space for the rigid–body and 'reduce out' translations to arrive at $SO(3)$ as the configuration space (see [MR99]).

The equations of motion for a rigid body with a fixed point in a gravitational field provide an interesting example of a system that is Hamiltonian (see (3.2.1)) relative to a *Lie–Poisson bracket* (see (3.2.3)). The underlying Lie algebra consists of the algebra of infinitesimal Euclidean motions in $\mathbb{R}^3$.

The basic phase–space we start with is again $T^*SO(3)$, parameterized by Euler angles and their conjugate momenta. In these variables, the equations are in canonical Hamiltonian form; however, the presence of gravity breaks the symmetry, and the system is no longer $SO(3)$ invariant, so it cannot be written entirely in terms of the body angular momentum p. One also needs to keep track of Γ, the 'direction of gravity' as seen from the body. This is defined by $\Gamma = A^{-1}k$, where k points upward and A is the element of $SO(3)$ describing the current configuration of the body. The equations of motion are

$$\begin{aligned}
\dot{p}_1 &= \frac{I_2 - I_3}{I_2 I_3} p_2 p_3 + Mgl(\Gamma^2\chi^3 - \Gamma^3\chi^2),\\
\dot{p}_2 &= \frac{I_3 - I_1}{I_3 I_1} p_3 p_1 + Mgl(\Gamma^3\chi^1 - \Gamma^1\chi^3),\\
\dot{p}_3 &= \frac{I_1 - I_2}{I_1 I_2} p_1 p_2 + Mgl(\Gamma^1\chi^2 - \Gamma^2\chi^1),
\end{aligned}$$

and

$$\dot{\Gamma} = \Gamma \times \Omega,$$

where Ω is the body angular velocity vector, I_1, I_2, I_3 are the body's principal moments of inertia, M is the body's mass, g is the acceleration of gravity, χ is the body fixed unit vector on the line segment connecting the fixed point with the body's center of mass, and l is the length of this segment.

The Euclidean Group and Its Lie Algebra

An element of $SE(3)$ is a pair (A, a) where $A \in SO(3)$ and $a \in \mathbb{R}^3$. The action of $SE(3)$ on $\mathbb{R}^3$ is the rotation A followed by translation by the vector a and has the expression

$$(A, a) \cdot x = Ax + a.$$

Using this formula, one sees that multiplication and inversion in $SE(3)$ are given by

$$(A, a)(B, b) = (AB, Ab + a) \qquad \text{and} \qquad (A, a)^{-1} = (A^{-1}, -A^{-1}a),$$

for $A, B \in SO(3)$ and $a, b \in \mathbb{R}^3$. The identity element is $(l, 0)$.

The Lie algebra of the Euclidean group $SE(3)$ is $\mathfrak{se}(3) = \mathbb{R}^3 \times \mathbb{R}^3$ with the Lie bracket

$$[(\xi, u), (\eta, v)] = (\xi \times \eta, \xi \times v - \eta \times u). \tag{2.27}$$

The Lie algebra of the Euclidean group has a structure that is a special case of what is called a *semidirect product*. Here it is the *product of the group of rotations with the corresponding group of translations*. It turns out that semidirect products occur under rather general circumstances when the symmetry in T^*G is broken.

The dual Lie algebra of the Euclidean group $SE(3)$ is $\mathfrak{se}(3)^* = \mathbb{R}^3 \times \mathbb{R}^3$ with the same Lie bracket (2.27). For the further details on adjoint orbits in $\mathfrak{se}(3)$ as well as coadjoint orbits in $\mathfrak{se}(3)^*$ see [MR99].

Symplectic Group in Hamiltonian Mechanics

Let $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$, with I the $n \times n$ identity matrix. Now, $A \in L(\mathbb{R}^{2n}, \mathbb{R}^{2n})$ is called a *symplectic matrix* if $A^T J\, A = J$. Let $Sp(2n, \mathbb{R})$ be the set of $2n \times 2n$ symplectic matrices. Taking determinants of the condition $A^T J\, A = J$ gives $\det A = \pm 1$, and so $A \in GL(2n, \mathbb{R})$. Furthermore, if $A, B \in Sp(2n, R)$, then $(AB)^T J(AB) = B^T A^T JAB = J$. Hence, $AB \in Sp(2n, \mathbb{R})$, and if $A^T J\, A = J$, then $JA = (A^T)^{-1} J = (A^{-1})^T J$, so $J = (A - 1)^T JA^{-1}$, or $A^{-1} \in Sp(2n, \mathbb{R})$. Thus, $Sp(2n, \mathbb{R})$ is a group [MR99].

The *symplectic Lie group*

$$Sp(2n, \mathbb{R}) = \left\{ A \in GL(2n, \mathbb{R}) : A^T J\, A = J \right\}$$

is a noncompact, connected Lie group of dimension $2n^2 + n$. Its Lie algebra

$$\mathfrak{sp}(2n, \mathbb{R}) = \left\{ A \in L(\mathbb{R}^{2n}, \mathbb{R}^{2n}) : A^T J\, A = J = 0 \right\},$$

called the *symplectic Lie algebra*, consists of the $2n \times 2n$ matrices A satisfying $A^T J\, A = 0$ [MR99].

Consider a particle of mass m moving in a potential $V(q)$, where $q^i = (q^1, q^2, q^3) \in \mathbb{R}^3$. Newtonian second law states that the particle moves along a curve $q(t)$ in $\mathbb{R}^3$ in such a way that $m\ddot{q}^i = -\operatorname{grad} V(q^i)$. Introduce the momentum $p_i = m\dot{q}^i$, and the energy

$$H(q, p) = \frac{1}{2m} \sum_{i=1}^{3} p_i^2 + V(q).$$

Then

$$\frac{\partial H}{\partial q^i} = \frac{\partial V}{\partial q^i} = -m\ddot{q}^i = -\dot{p}_i, \quad \text{and}$$
$$\frac{\partial H}{\partial p_i} = \frac{1}{m} p_i = \dot{q}^i, \qquad (i = 1, 2, 3),$$

and hence Newtonian law $F = m\ddot{q}^i$ is equivalent to Hamilton's equations

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}.$$

Now, writing $z = (q^i, p_i)$ [MR99],

$$J \,\mathrm{grad}\, H(z) = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H}{\partial q^i} \\ \frac{\partial H}{\partial p_i} \end{pmatrix} = \left(\dot{q}^i, \dot{p}_i\right) = \dot{z},$$

so Hamilton's equations read

$$\dot{z} = J \,\mathrm{grad}\, H(z). \tag{2.28}$$

Now let $f : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \times \mathbb{R}^3$ and write $w = f(z)$. If $z(t)$ satisfies Hamilton's equations (2.28) then $w(t) = f(z(t))$ satisfies $\dot{w} = A^T \dot{z}$, where $A^T = [\partial w^i / \partial z^j]$ is the Jacobian matrix of f. By the chain rule,

$$\dot{w} = A^T J \,\underset{z}{\mathrm{grad}}\, H(z) = A^T J\, A \,\underset{w}{\mathrm{grad}}\, H(z(w)).$$

Thus, the equations for $w(t)$ have the form of Hamilton's equations with energy $K(w) = H(z(w))$ iff $A^T J\, A = J$, that is, iff A is symplectic. A nonlinear transformation f is canonical iff its Jacobian matrix is symplectic. $Sp(2n, \mathbb{R})$ is the linear invariance group of classical mechanics [MR99].

Now, before giving our main biomechanical applications of Lie groups, we give an interesting application in the realm of dynamical games.

2.4.3 Dynamical Games on Lie Groups

In this section we propose a general approach to modelling *conflict resolution manoeuvres* for land, sea and airborne *vehicles*, using dynamical games on Lie groups. We use the generic name 'vehicle' to represent all planar vehicles, namely land and sea vehicles, as well as fixed altitude motion of aircrafts (see, e.g., [LGS, TPS98]). First, we elaborate on the two–vehicle conflict resolution manoeuvres, and after that discuss the multi–vehicle manoeuvres.

We explore special features of the dynamical games solution when the underlying dynamics correspond to left–invariant control systems on Lie groups. We show that the 2D (i.e., planar) motion of a vehicle may be modelled as a control system on the Lie group $SE(2)$. The proposed algorithm surrounds each vehicle with a circular protected zone, while the simplification in the derivation of *saddle* and *Nash strategies* follows from the use of symplectic reduction techniques [MR99]. To model the two–vehicle conflict resolution, we construct the safe subset of the state space for one of the vehicles using zero–sum non–cooperative dynamic game theory [BO95] which we specialize to the $SE(2)$ group. If the underlying continuous dynamics are left–invariant control systems, reduction techniques can be used in the computation of safe sets.

Configuration Models for Planar Vehicles

The configuration of each individual vehicle is described by an element of the *Lie group* $SE(2)$ of rigid–body motions in $\mathbb{R}^2$. Let $g_i \in SE(2)$ denote the configurations of vehicles labelled i, with $i = 1, 2$. The motion of each vehicle may be modelled as a left–invariant vector–field on $SE(2)$:

$$\dot{g}_i = g_i X_i, \tag{2.29}$$

where the vector–fields X_i belong to the vector space $\mathfrak{se}(2)$, the *Lie algebra* associated with the group $SE(2)$.

Each $g_i \in SE(2)$ can be represented in standard local coordinates (x_i, y_i, θ_i) as

$$g_i = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & x_i \\ \sin\theta_i & \cos\theta_i & y_i \\ 0 & 0 & 1 \end{bmatrix},$$

where x_i, y_i is the position of vehicle i and θ_i is its orientation, or heading. The associated Lie algebra is $\mathfrak{se}(2)$, with $X_i \in \mathfrak{se}(2)$ represented as

$$X_i = \begin{bmatrix} 0 & -\omega_i & v_i \\ \omega_i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where v_i and ω_i represent the *translational* (linear) and *rotational* (angular) *velocities*, respectively.

Now, to determine *dynamics of the relative configuration* of two vehicles, we perform a change (transformation) of coordinates, to place the identity element of the group $SE(2)$ on vehicle 1. If $g^{rel} \in SE(2)$ denotes the relative configuration of vehicle 2 with respect to vehicle 1, then

$$g_2 = g_1 g^{rel} \Longrightarrow g^{rel} = g_1^{-1} g_2.$$

Differentiation with respect to time yields the dynamics of the relative configuration:

$$\dot{g}^{rel} = g^{rel} X_2 - X_1 g^{rel},$$

which expands into:

$$\begin{aligned} \dot{x}^{rel} &= -v_1 + v_2 \cos\theta^{rel} + \omega_1 y^{rel}, \\ \dot{y}^{rel} &= v_2 \sin\theta^{rel} - \omega_1 x^{rel}, \\ \dot{\theta}^{rel} &= \omega_2 - \omega_1. \end{aligned}$$

Two–Vehicles Conflict Resolution Manoeuvres

Next, we seek *control strategies* for each vehicle, which are *safe* under (possible) *uncertainty* in the actions of neighbouring vehicle. For this, we expand the dynamics of two vehicles (2.29),

$$\dot{g}_1 = g_1 X_1, \qquad \dot{g}_2 = g_2 X_2,$$

and write it in the matrix form as

$$\dot{g} = gX, \tag{2.30}$$

with

$$g = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix}, \qquad X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix},$$

in which g is an element in the configuration manifold $M = SE(2) \times SE(2)$, while the vector–fields $X_i \in \mathfrak{se}(2)\times\mathfrak{se}(2)$ are linearly parameterised by velocity inputs $(\omega_1, v_1) \in \mathbb{R}^2$ and $(\omega_2, v_2) \in \mathbb{R}^2$.

The goal of each vehicle is to maintain *safe operation*, meaning that

(i) the vehicles remain outside of a specified *target set* T with boundary ∂T, defined by

$$T = \{g \in M | l(g) < 0\},$$

where $l(g)$ is a differentiable circular function,

$$l(g) = (x_2 - x_1)^2 + (y_2 - y_1)^2 - \rho^2$$

(with ρ denoting the *radius of a circular protected zone*) defines the minimum allowable lateral separation between vehicles; and

(ii)

$$dl(g) \neq 0 \quad \text{on} \quad \partial T = \{g \in M | l(g) = 0\},$$

where d represents the *exterior derivative* (a unique generalization of the gradient, divergence and curl).

Now, due to possible uncertainty in the actions of vehicle 2, the safest possible strategy of vehicle 1 is to drive along a *trajectory* which guarantees that the minimum allowable separation with vehicle 2 is maintained regardless of the actions of vehicle 2. We therefore formulate this problem as a *zero–sum dynamical game* with two players: control vs. disturbance. The *control* is the action of vehicle 1,

$$u = (\omega_1, v_1) \in U,$$

and the *disturbance* is the action of vehicle 2,

$$d = (\omega_2, v_2) \in D.$$

Here the control and disturbance sets, U and D, are defined as

$$U = ([\omega_1^{\min}, \omega_1^{\max}], [v_1^{\min}, v_1^{\max}]),$$
$$D = ([\omega_2^{\min}, \omega_2^{\max}], [v_2^{\min}, v_2^{\max}])$$

and the corresponding control and disturbance functional spaces, $\mathcal{U}$ and $\mathcal{D}$ are defined as:

$$\mathcal{U} = \{u(\cdot) \in PC^0(\mathbb{R}^2) | u(t) \in U, t \in \mathbb{R}\},$$
$$\mathcal{D} = \{d(\cdot) \in PC^0(\mathbb{R}^2) | d(t) \in U, t \in \mathbb{R}\},$$

where $PC^0(\mathbb{R}^2)$ is the space of piecewise continuous functions over $\mathbb{R}^2$.

We define the *cost of a trajectory* $g(t)$ which starts at state g at initial time $t \leq 0$, evolves according to (2.30) with input $(u(\cdot), d(\cdot))$, and ends at the final state $g(0)$ as:

$$\begin{aligned} & J(g, u(\cdot), d(\cdot), t) \ : \ SE(2) \times SE(2) \times \mathcal{U} \times \mathcal{D} \times \mathbb{R}_- \to \mathbb{R}, \\ \text{such that} \quad & J(g, u(\cdot), d(\cdot), t) = l(g(0)), \end{aligned} \tag{2.31}$$

where 0 is the final time (without loss of generality). Thus the cost depends only on the final state $g(0)$ (the Lagrangian, or running cost, is identically zero). The game is won by vehicle 1 if the terminal state $g(0)$ is either outside T or on ∂T (i.e., $J(g, 0) \geq 0$), and is won by vehicle 2 otherwise.

This two–player zero–sum dynamical game on $SE(2)$ is defined as follows. Consider the matrix system (2.30), $\dot{g} = gX$, over the time interval $[t, 0]$ where $t < 0$ with the cost function $J(g, u(\cdot), d(\cdot), t)$ defined by (2.31) As vehicle 1 attempts to maximize this cost assuming that vehicle 2 is acting blindly, the optimal control action and worst disturbance actions are calculated as

$$u^* = \arg \max_{u \in \mathcal{U}} \min_{d \in \mathcal{D}} J(g, u(\cdot), d(\cdot), t), \qquad d^* = \arg \min_{d \in \mathcal{D}} \max_{u \in \mathcal{U}} J(g, u(\cdot), d(\cdot), t).$$

The game is said to have a *saddle solution* (u^*,d^*) if the resulting optimal cost $J^*(g, t)$ does not depend on the order of play, i.e., on the order in which the maximization and minimization is performed:

$$J^*(g, t) = \max_{u \in \mathcal{U}} \min_{d \in \mathcal{D}} J(g, u(\cdot), d(\cdot), t) = \min_{d \in \mathcal{D}} \max_{u \in \mathcal{U}} J(g, u(\cdot), d(\cdot), t).$$

Using this saddle solution we calculate the 'losing states' for vehicle 1, called the *predecessor* $Pre_t(T)$ of the target set T,

$$Pre_t(T) = \{g \in M | J(g, u^*(\cdot), d(\cdot), t) < 0\}.$$

Symplectic Reduction and Dynamical Games on $SE(2)$

Since vehicles 1 and 2 have dynamics given by left–invariant control systems on the Lie group $SE(2)$, we have

$$X_1 = \xi^1 \omega_1 + \xi^2 v_1, \qquad X_2 = \xi^1 \omega_2 + \xi^2 v_2,$$

with ξ^1, ξ^2 being two of the three basis elements for the tangent Lie algebra $\mathfrak{se}(2)$ given by

$$\xi^1 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \xi^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \xi^3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

If p_1 (resp. p_2) is a cotangent vector–field to $SE(2)$ at g_1 (resp. g_2), belonging to the cotangent (dual) Lie algebra $\mathfrak{se}(2)^*$, we can define the *momentum functions* for both vehicles:

$$\begin{aligned} P_1^1 &= < p_1, g_1\xi^1 >, P_1^2 =< p_1, g_1\xi^2 >, P_1^3 =< p_1, g_1\xi^3 >, \\ P_2^1 &= < p_2, g_2\xi^1 >, P_2^2 =< p_2, g_2\xi^2 >, P_2^3 =< p_2, g_2\xi^3 >, \end{aligned}$$

which can be compactly written as

$$P_i^j =< p_i, g_i\xi^j > .$$

Defining $p = (p_1, p_2) \in \mathfrak{se}(2)^* \times \mathfrak{se}(2)^*$, the *optimal cost* for the two-player, zero-sum dynamical game is given by

$$J^*(g,t) = \max_{u\in\mathcal{U}} \min_{d\in\mathcal{D}} J(g, u(\cdot), d(\cdot), t) = \max_{u\in\mathcal{U}} \min_{d\in\mathcal{D}} l(g(0)).$$

The *Hamiltonian* $H(g, p, u, d)$ is given by

$$H(g,p,u,d) = P_1^1\omega_1 + P_1^2 v_1 + P_2^1\omega_1 + P_2^2 v_1$$

for control and disturbance inputs $(\omega_1, v_1) \in U$ and $(\omega_2, v_2) \in D$ as defined above. It follows that the *optimal Hamiltonian* $H^*(g, p)$, defined on the cotangent bundle $T^*SE(2)$, is given by

$$\begin{aligned} H^*(g,p) = {} & P_1^1 \frac{\omega_1^{\max} + \omega_1^{\min}}{2} + P_2^1 \frac{\omega_2^{\max} + \omega_2^{\min}}{2} + |P_1^1| \frac{\omega_1^{\max} - \omega_1^{\min}}{2} \\ & - |P_1^1| \frac{\omega_2^{\max} - \omega_2^{\min}}{2} + P_1^2 \frac{v_1^{\max} + v_1^{\min}}{2} + P_2^2 \frac{v_2^{\max} + v_2^{\min}}{2} \\ & + |P_1^2| \frac{v_1^{\max} - v_1^{\min}}{2} - |P_1^2| \frac{v_2^{\max} - v_2^{\min}}{2} \end{aligned}$$

and the *saddle solution* (u^*, d^*) is given by

$$u^* = \arg\max_{u\in\mathcal{U}} \min_{d\in\mathcal{D}} H(g,p,u,d), \qquad d^* = \arg\min_{d\in\mathcal{D}} \max_{u\in\mathcal{U}} H(g,p,u,d). \tag{2.32}$$

Note that $H(g, p, u, d)$ and $H^*(g, p)$ do not depend on the state g and costate p directly, rather through the momentum functions P_1^j, P_2^j. This is because the dynamics are determined by left–invariant vector fields on the Lie group and the Lagrangian is state independent [MR99].

The optimal Hamiltonian $H^*(g,p)$ determines a 12D Hamiltonian vector–field X_{H^*} on the symplectic manifold $T^*M = SE(2)\times SE(2)\times \mathfrak{se}(2)^*\times \mathfrak{se}(2)^*$ (which is the cotangent bundle of the configuration manifold M), defined by Hamilton's equations

$$X_{H^*}: \dot{g} = \frac{\partial H^*(g,p)}{\partial p}, \qquad \dot{p} = -\frac{\partial H^*(g,p)}{\partial g},$$

with initial condition at time t being $g(t) = g$ and final condition at time 0 being $p(0) = dl(g(0))$. In general, to solve for the saddle solution (2.32), one needs to solve the ODE system for all states. However since the original system on $M = SE(2)\times SE(2)$ is left–invariant, it induces generic symmetries in the Hamiltonian dynamics on T^*M , referred to as Marsden–Weinstein reduction of Hamiltonian systems on symplectic manifolds, see [MR99]. In general for such systems one only needs to solve an ODE system with half of the dimensions of the underlying symplectic manifold.

For the two-vehicle case we only need to solve an ODE system with 6 states. That is exactly given by the dynamics of the 6 momentum functions

$$\dot{P}_i^j = L_{X_{H^*}} P_i^j = \{P_i^j, H^*(g,p)\}, \tag{2.33}$$

for $i,j = 1,2$, which is the *Lie derivative* of P_i^j with respect to the Hamiltonian vector–field X_{H^*}. In the equation (2.33), the bracket $\{\cdot,\cdot\}$ is the *Poisson bracket* [IP01a], giving the commutation relations:

$$\begin{aligned} &\{P_1^1, P_1^2\} = P_1^3, \qquad \{P_1^2, P_1^3\} = 0, \qquad \{P_1^3, P_1^1\} = P_1^2, \\ &\{P_2^1, P_2^2\} = P_2^3, \qquad \{P_2^2, P_2^3\} = 0, \qquad \{P_2^3, P_2^1\} = P_2^2. \end{aligned}$$

Using these commutation relations, equation (2.33) can be written explicitly:

$$\begin{aligned}
\dot{P}_1^1 &= P_1^3\left(\frac{v_1^{\max}+v_1^{\min}}{2} + sign(P_1^2)\frac{v_1^{\max}+v_1^{\min}}{2}\right),\\
\dot{P}_1^2 &= P_1^3\left(-\frac{\omega_1^{\max}+\omega_1^{\min}}{2} - sign(P_1^1)\frac{\omega_1^{\max}-\omega_1^{\min}}{2}\right),\\
\dot{P}_1^3 &= P_1^2\left(\frac{\omega_1^{\max}+\omega_1^{\min}}{2} + sign(P_1^1)\frac{\omega_1^{\max}-\omega_1^{\min}}{2}\right),\\
\dot{P}_2^1 &= P_2^3\left(\frac{v_2^{\max}+v_2^{\min}}{2} + sign(P_2^2)\frac{v_2^{\max}+v_2^{\min}}{2}\right),\\
\dot{P}_2^2 &= P_2^3\left(-\frac{\omega_2^{\max}+\omega_2^{\min}}{2} - sign(P_2^1)\frac{\omega_2^{\max}-\omega_2^{\min}}{2}\right),\\
\dot{P}_2^3 &= P_2^2\left(\frac{\omega_2^{\max}+\omega_2^{\min}}{2} + sign(P_2^1)\frac{\omega_2^{\max}-\omega_2^{\min}}{2}\right).
\end{aligned}$$

The final conditions for the variables $P_1^j(t)$ and $P_2^j(t)$ are obtained from the boundary of the safe set as

$$P_1^j(0) =< d_1 l(g), g_1\xi^j >, \qquad P_2^j(0) =< d_2 l(g), g_2\xi^j >,$$

where d_1 is the derivative of l taken with respect to its first argument g_1 only (and similarly for d_2). In this way, $P_1^j(t)$ and $P_2^j(t)$ are obtained for $t \leq 0$. Once this has been calculated, the optimal input $u^*(t)$ and the worst disturbance $d^*(t)$ are given respectively as

$$u^*(t) = \begin{cases} \omega_1^*(t) = \begin{cases} \omega_1^{\max} & \text{if} \quad P_1^1(t) > 0 \\ \omega_1^{\min} & \text{if} \quad P_1^1(t) < 0 \end{cases} \\ v_1^*(t) = \begin{cases} v_1^{\max} & \text{if} \quad P_1^2(t) > 0 \\ v_1^{\min} & \text{if} \quad P_1^2(t) < 0 \end{cases} \end{cases}$$
$$d^*(t) = \begin{cases} \omega_2^*(t) = \begin{cases} \omega_2^{\max} & \text{if} \quad P_2^1(t) > 0 \\ \omega_2^{\min} & \text{if} \quad P_2^1(t) < 0 \end{cases} \\ v_2^*(t) = \begin{cases} v_2^{\max} & \text{if} \quad P_2^2(t) > 0 \\ v_2^{\min} & \text{if} \quad P_2^2(t) < 0 \end{cases} \end{cases} .$$

Nash Solutions for Multi–Vehicle Manoeuvres

The methodology introduced in the previous sections can be generalized to find conflict–resolutions for multi–vehicle manoeuvres. Consider the three–vehicle dynamics:

$$\dot{g} = gX, \tag{2.34}$$

with

$$g = \begin{bmatrix} g_1 & 0 & 0 \\ 0 & g_2 & 0 \\ 0 & 0 & g_3 \end{bmatrix}, \qquad X = \begin{bmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_3 \end{bmatrix},$$

where g is an element in the configuration space $M = SE(2) \times SE(2) \times SE(2)$ and $X \in \mathfrak{se}(2) \times \mathfrak{se}(2) \times \mathfrak{se}(2)$ is linearly parameterised by inputs (ω_1, v_1), (ω_2, v_2) and (ω_3, v_3).

Now, the target set T is defined as

$$T = \{g \in M | l_1(g) < 0 \vee l_2(g) < 0 \vee l_3(g) < 0\},$$

where

$$\begin{aligned} l_1(g) &= \min\{(x_2 - x_1)^2 + (y_2 - y_1)^2 - \rho^2, \quad (x_3 - x_1)^2 + (y_3 - y_1)^2 - \rho^2\}, \\ l_2(g) &= \min\{(x_3 - x_2)^2 + (y_3 - y_2)^2 - \rho^2, \quad (x_1 - x_2)^2 + (y_1 - y_2)^2 - \rho^2\}, \\ l_3(g) &= \min\{(x_2 - x_3)^2 + (y_2 - y_3)^2 - \rho^2, \quad (x_1 - x_3)^2 + (y_1 - y_3)^2 - \rho^2\}. \end{aligned}$$

The control inputs $u = (u_1, u_2, u_3)$ are the actions of vehicle 1, 2 and 3:

$$u_i = (\omega_i, v_i) \in U_i,$$

where U_i are defined as

$$U_i = ([\omega_i^{\min}, \omega_i^{\max}], [v_i^{\min}, v_i^{\max}]).$$

Clearly, this can be generalized to N vehicles.

The cost functions $J_i(g, \{u_i(\cdot)\}, t)$ are defined as

$$J_i(g, \{u_i(\cdot)\}, t) : \prod_{i=1}^{N} SE_i(2) \times \prod_{i=1}^{N} \mathcal{U}_i \times \mathbb{R}_- \to \mathbb{R},$$

such that $J_i(g, \{u_i(\cdot)\}, t) = l_i(g(0))$.

The simplest non–cooperative solution strategy is a so–called *non–cooperative Nash equilibrium* (see e.g., [BO95]). A set of controls u_i^*, $(i = 1, ..., N)$ is said to be a *Nash strategy*, if for each player modification of that strategy under the assumption that the others play their Nash strategies results in a decrease in his payoff, that is for $i = 1, ..., N$, and $\forall u_i(\cdot)$,

$$J_i(u_1, ..., u_i, ..., u_N) \leq J_i(u_1^*, ..., u_i^*, ..., u_N^*), \qquad (u \neq u^*).$$

(Note that Nash equilibria may not be unique. It is also easy to see that for the two–player zero–sum game, a Nash equilibrium is a saddle solution with $J = J_1 = -J_2$.)

For N vehicles, the momentum functions are defined as in the two–vehicle case:

$$P_i^j =< p_i, g_i \xi^j >,$$

with $p_i \in \mathfrak{se}(2)^*$ for $i = 1, ..., N$ and ξ^j defined as above.

Then the Hamiltonian $H(g, p, u_1, ...u_N)$ is given by

$$H(g, p, u_1, ...u_N) = P_i^1 \omega_i + P_i^2 v_i.$$

The first case we consider is one in which all the vehicles are *cooperating*, meaning that each tries to avoid conflict assuming the others are doing the same. In this case, the optimal Hamiltonian $H^*(g, p)$ is

$$H^*(g, p) = \max_{u_i \in U_i} H(g, p, u_1, ...u_N).$$

For example, if $N = 3$, one may solve for (u_1^*, u_2^*, u_3^*), on the 9D quotient space T^*M/M, so that the optimal control inputs are given as

$$u_i^*(t) = \begin{cases} \omega_i^*(t) = \begin{cases} \omega_i^{\max} & \text{if} \quad P_i^1(t) > 0 \\ \omega_i^{\min} & \text{if} \quad P_i^1(t) < 0 \end{cases} \\ v_i^*(t) = \begin{cases} v_i^{\max} & \text{if} \quad P_i^2(t) > 0 \\ v_i^{\min} & \text{if} \quad P_i^2(t) < 0 \end{cases} \end{cases}.$$

One possibility for the optimal Hamiltonian corresponding to the *non–cooperative case* is

$$H^*(g, p) = \max_{u_1 \in U_1} \max_{u_2 \in U_2} \max_{u_3 \in U_3} H(g, p, u_1, u_2, u_3).$$

2.4.4 Group Structure of the Biomechanical Manifold M

Purely Rotational Biomechanical Manifold

Kinematics of an n–segment human–body chain (like arm, leg or spine) is usually defined as a map between *external coordinates* (usually, end–effector coordinates) x^r $(r = 1, \dots, n)$ and *internal joint coordinates* q^i $(i = 1, \dots, N)$ (see [IS01, Iva02, IP01b, IP01b, Iva05]). The *forward kinematics* are defined as a nonlinear map $x^r = x^r(q^i)$ with a corresponding linear vector functions $dx^r = \partial x^r/\partial q^i \, dq^i$ of differentials: and $\dot{x}^r = \partial x^r/\partial q^i \, \dot{q}^i$ of velocities. When the rank of the configuration–dependent Jacobian matrix $J \equiv \partial x^r/\partial q^i$ is less than n the *kinematic singularities* occur; the onset of this condition could be detected by the *manipulability measure*. The *inverse kinematics* are defined conversely by a nonlinear map $q^i = q^i(x^r)$ with a corresponding linear vector functions $dq^i = \partial q^i/\partial x^r \, dx^r$ of differentials and $\dot{q}^i = \partial q^i/\partial x^r \, \dot{x}^r$ of velocities. Again, in the case of *redundancy* $(n < N)$, the inverse kinematic problem admits infinite solutions; often the *pseudo–inverse* configuration–control is used instead: $\dot{q}^i = J^* \, \dot{x}^r$, where $J^* = J^T (J \, J^T)^{-1}$ denotes the Moore–Penrose pseudo–inverse of the Jacobian matrix J.

Humanoid joints, that is, internal coordinates q^i $(i = 1, \dots, N)$, constitute a smooth configuration manifold M, described as follows. Uniaxial, 'hinge' joints represent constrained, rotational Lie groups $SO(2)^i_{cnstr}$, parameterized by constrained angles $q^i_{cnstr} \equiv q^i \in [q^i_{\min}, \, q^i_{\max}]$. Three–axial, 'ball–and–socket' joints represent constrained rotational Lie groups $SO(3)^i_{cnstr}$, parameterized by constrained Euler angles $q^i = q^{\phi_i}_{cnstr}$ (in the following text, the subscript 'cnstr' will be omitted, for the sake of simplicity, and always assumed in relation to internal coordinates q^i).

All $SO(n)$–joints are Hausdorff C^k–manifolds with atlases $(U_\alpha, \, u_\alpha)$; in other words, they are paracompact and metrizable smooth manifolds, admitting Riemannian metric.

Let A and B be two smooth manifolds described by smooth atlases (U_α, u_α) and (V_β, v_β), respectively. Then the family $(U_\alpha \times V_\beta, u_\alpha \times v_\beta : U_\alpha \times V_\beta \to \mathbb{R}^m \times \mathbb{R}^n)_(\alpha, \beta) \in A \times B$ is a smooth atlas for the direct product $A \times B$. Now, if A and B are two Lie groups (say, $SO(n)$), then their *direct product* $G = A \times B$ is at the same time their direct product as smooth manifolds and their direct product as algebraic groups, with the product law

$$(a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2), \qquad (a_{1,2} \in A, \; b_{1,2} \in B).$$

Generalizing the direct product to N rotational joint groups, we can draw an *anthropomorphic product–tree* (see Figure 2.1) using a line segment '–' to represent direct products of human $SO(n)$–joints. This is our basic model of the biomechanical configuration manifold M (see (3.2.1) below).

Let $T_q M$ be a tangent space to M at the point q. The *tangent bundle* TM represents a union $\cup_{q \in M} T_q M$, together with the standard topology on TM and a natural smooth manifold structure, the dimension of which is twice the

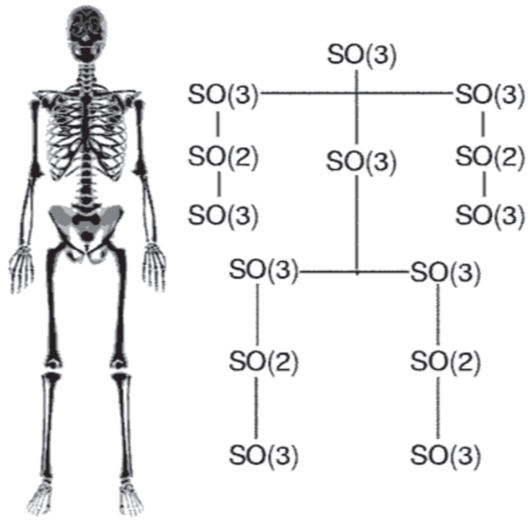


Fig. 2.1. Purely rotational, whole–body biomechanical manifold, with a single $SO(3)-$joint representing the whole spinal movability.

dimension of M. A vector–field X on M represents a section $X : M \to TM$ of the tangent bundle TM.

Analogously let T_q^*M be a cotangent space to M at q, the dual to its tangent space T_qM. The *cotangent bundle* T^*M represents a union $\cup_{q \in M} T_q^*M$, together with the standard topology on T^*M and a natural smooth manifold structure, the dimension of which is twice the dimension of M. A $1-$form θ on M represents a section $\theta : M \to T^*M$ of the cotangent bundle T^*M.

We refer to the tangent bundle TM of biomechanical configuration manifold M as the *velocity phase–space* manifold, and to its cotangent bundle T^*M as the *momentum phase–space* manifold.

Reduction of the Rotational Biomechanical Manifold

The biomechanical configuration manifold M (Figure 2.1) can be (for the sake of the brain–like motor control) reduced to N–torus T^N, in three steps, as follows.

First, a single three–axial $SO(3)-$joint can be reduced to the direct product of three uniaxial $SO(2)-$joints, in the sense that three hinge joints can produce any orientation in space, just as a ball–joint can. Algebraically, this means reduction (using symbol '$\gtrsim$') of each of the three $SO(3)$ rotation matrices to the corresponding $SO(2)$ rotation matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} \gtrsim \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

$$\begin{pmatrix} \cos\psi & 0 & \sin\psi \\ 0 & 1 & 0 \\ -\sin\psi & 0 & \cos\psi \end{pmatrix} \gtrsim \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix}$$

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \gtrsim \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

In this way we can set the *reduction equivalence relation* $SO(3) \gtrsim SO(2) \triangleright SO(2) \triangleright SO(2)$, where '$\triangleright$' denotes the noncommutative *semidirect product* (see (2.4.2) above).

Second, we have a homeomorphism: $SO(2) \sim S^1$, where S^1 denotes the constrained unit circle in the complex plane, which is an Abelian Lie group.

Third, let I^N be the unit cube $[0,1]^N$ in $\mathbb{R}^N$ and '$\sim$' an equivalence relation on $\mathbb{R}^N$ obtained by 'gluing' together the opposite sides of I^N, preserving their orientation. The manifold of human–body configurations (Figure 2.1) can be represented as the *quotient space* of $\mathbb{R}^N$ by the space of the integral lattice points in $\mathbb{R}^N$, that is a constrained ND torus T^N (4.2),

$$\mathbb{R}^N/Z^N = I^N/\sim \cong \prod_{i=1}^{N} S_i^1 \equiv \{(q^i, i = 1, \dots, N) : \text{mod } 2\pi\} = T^N. \qquad (2.35)$$

Since S^1 is an Abelian Lie group, its N–fold tensor product T^N is also an Abelian Lie group, the toral group, of all nondegenerate diagonal $N \times N$ matrices. As a Lie group, the biomechanical configuration space $M \equiv T^N$ has a natural Banach manifold structure with local internal coordinates $q^i \in U$, U being an open set (chart) in T^N.

Conversely by 'ungluing' the configuration space we get the primary unit cube. Let '$\sim^*$' denote an equivalent decomposition or 'ungluing' relation. By the Tychonoff product–topology theorem, for every such quotient space there exists a 'selector' such that their quotient models are homeomorphic, that is, $T^N/\sim^* \approx A^N/\sim^*$. Therefore I^N represents a 'selector' for the configuration torus T^N and can be used as an N–directional 'command–space' for the topological control of human motion. Any subset of DOF on the configuration torus T^N representing the joints included in human motion has its simple, rectangular image in the command space – selector I^N. Operationally, this resembles what the *brain–motor–controller*, the *cerebellum*, actually performs on the highest level of human motor control (see Chapter 5).

The Complete Biomechanical Manifold

The full kinematics of a whole human–like body can be split down into *five kinematic chains*: one for each leg and arm, plus one for spine with the head. In all five chains internal joint coordinates, namely n_1 constrained rotations

x^k_{rt} together with n_2 of even more constrained translations x^j_{tr} (see Figure 2.2), constitute a smooth nD anthropomorphic configuration manifold M, with local coordinates x^i, $(i = 1,\ldots,n)$. That is, the motion space in each joint is defined as a semidirect (noncommutative) product of the Lie group $SO(n)$ of constrained rotations and a corresponding Lie group $\mathbb{R}^n$ of even more restricted translations. More precisely, in each movable human–like joint we have an action of the constrained special Euclidean $SE(3)$ group (see (2.4.2) above). The joints themselves are linked by direct (commutative) products.

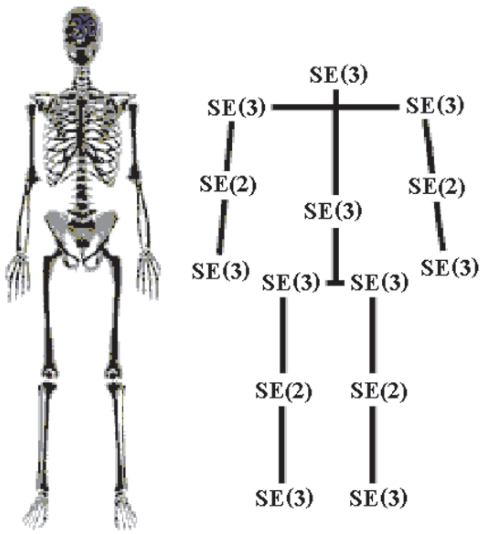


Fig. 2.2. A medium–resolution, whole–body biomechanical manifold, with just a single $SE(3)-$joint representing the spinal movability.

Realistic Human Spine Manifold

The high–resolution human spine manifold is a dynamical chain consisting of 25 constrained $SE(3)-$ joints. Each movable spinal joint has 6 DOF: 3 dominant rotations, (performed first in any free spinal movement), restricted to about 7 angular degrees and 3 secondary translations (performed after reaching the limit of rotational amplitude), restricted to about 5 mm (see Figure 2.3).

Now, $SE(3) = SO(3) \rhd \mathbb{R}^3$ is a non–compact group, so there is no any natural metric given by the kinetic energy on $SE(3)$, and consequently, no natural controls in the sense of geodesics on $SE(3)$. However, both of its subgroups, $SO(3)$ and $\mathbb{R}^3$, are compact with quadratic metric forms defined by standard line element $g_{ij}dq^idq^j$, and therefore admit optimal muscular–like controls in the sense of geodesics (see section 2.5.1 below).

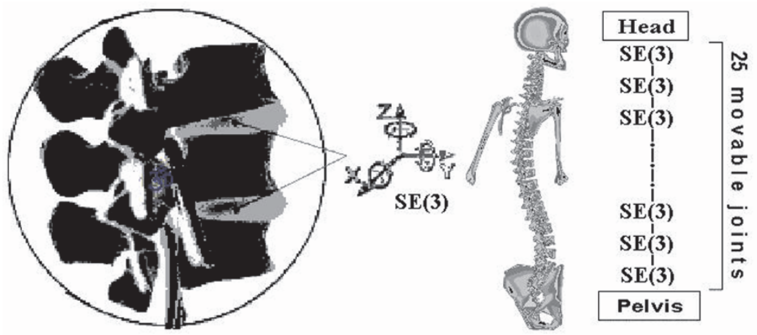


Fig. 2.3. The high–resolution human spine manifold is a dynamical chain consisting of 25 constrained $SE(3)-$joints.

2.4.5 Lie Symmetries in Biomechanics

Lie Symmetry Groups

Exponentiation of Vector Fields on M

Let $x = (x^1, ..., x^r)$ be local coordinates at a point m on a smooth $n-$manifold M. Recall that the flow generated by the vector–field

$$v = \xi^i(x)\,\partial_{x^i} \in M,$$

is a solution of the system of ODEs

$$\frac{dx^i}{d\varepsilon} = \xi^i(x^1, ..., x^m), \qquad (i = 1, ..., r).$$

The computation of the flow, or one–parameter group of diffeomorphisms, generated by a given vector–field v (i.e., solving the system of ODEs) is often referred to as *exponentiation* of the vector–field, denoted by $\exp(\varepsilon v)\,x$ (see [Olv86]).

If $v, w \in M$ are two vectors defined by

$$v = \xi^i(x)\,\partial_{x^i} \qquad \text{and} \qquad w = \eta^i(x)\,\partial_{x^i},$$

then

$$\exp(\varepsilon v)\,\exp(\theta w)\,x = \exp(\theta w)\exp(\varepsilon v)\,x,$$

for all $\epsilon, \theta \in \mathbb{R}, x \in M$, such that both sides are defined, iff they commute, i.e., $[v, w] = 0$ everywhere [Olv86].

A system of vector–fields $\{v_1, ..., v_r\}$ on a smooth manifold M is in *involution* if there exist smooth real–valued functions $h^k_{ij}(x)$, $x \in M$, $i, j, k = 1, ..., r$, such that for each i, j,

$$[v_i, v_j] = h^k_{ij} \cdot v_k.$$

Let $v \neq 0$ be a right–invariant vector–field on a Lie group G. Then the flow generated by v through the identity e, namely

$$g_\varepsilon = \exp(\varepsilon v)\, e \equiv \exp(\varepsilon v),$$

is defined for all $\varepsilon \in \mathbb{R}$ and forms a one–parameter subgroup of G, with

$$g_{\varepsilon+\delta} = g_\varepsilon \cdot g_\delta, \qquad g_0 = e, \qquad g_\varepsilon^{-1} = g_{-\varepsilon},$$

isomorphic to either $\mathbb{R}$ itself or the circle group $SO(2)$. Conversely, any connected 1D subgroup of G is generated by such a right–invariant vector–field in the above manner [Olv86].

For example, let $G = GL(n)$ with Lie algebra $\mathfrak{gl}(n)$, the space of all $n \times n$ matrices with commutator as the Lie bracket. If $A \in \mathfrak{gl}(n)$, then the corresponding right–invariant vector–field v_A on $GL(n)$ has the expression [Olv86]

$$v_A = a^i_k x^k_j \, \partial_{x^i_j}.$$

The one–parameter subgroup $\exp(\varepsilon v_A)\, e$ is found by integrating the system of n^2 ordinary differential equations

$$\frac{dx^i_j}{d\varepsilon} = a^i_k x^k_j, \qquad x^i_j(0) = \delta^i_j, \qquad (i,j = 1,...,n),$$

involving matrix entries of A. The solution is just the matrix exponential $X(\varepsilon) = e^{\varepsilon A}$, which is the one–parameter subgroup of $GL(n)$ generated by a matrix A in $\mathfrak{gl}(n)$.

Recall that the *exponential map* $\exp : \mathfrak{g} \to G$ is obtained by setting $\varepsilon = 1$ in the one–parameter subgroup generated by vector–field v :

$$\exp(v) \equiv \exp(v)\, e.$$

Its differential at 0,

$$d\exp : T\mathfrak{g}|_0 \simeq \mathfrak{g} \to TG|_e \simeq \mathfrak{g}$$

is the identity map.

Lie Symmetry Groups and General Differential Equations

Consider a system $\mathcal{S}$ of general differential equations (DEs, to be distinguished from ODEs) involving p independent variables $x = (x^1, ..., x^p)$, and q dependent variables $u = (u^1, ..., u^q)$. The solution of the system will be of the form $u = f(x)$, or, in components, $u^\alpha = f^\alpha(x^1, ..., x^p)$, $\alpha = 1, ..., q$ (so that Latin indices refer to independent variables while Greek indices refer to dependent variables). Let $X = \mathbb{R}^p$, with coordinates $x = (x^1, ..., x^p)$, be the space representing the independent variables, and let $U = \mathbb{R}^q$, with coordinates $u = (u^1, ..., u^q)$, represent dependent variables. A Lie symmetry group G of the system $\mathcal{S}$ will be a local group of transformations acting on some open subset $M \subset X \times U$ in such way that G transforms solutions of $\mathcal{S}$ to other solutions of $\mathcal{S}$ [Olv86].

More precisely, we need to explain exactly how a given transformation $g \in G$, where G is a Lie group, transforms a function $u = f(x)$. We firstly identify the function $u = f(x)$ with its graph

$$\Gamma_f \equiv \{(x, f(x)) : x \in \operatorname{dom} f \equiv \Omega\} \subset X \times U,$$

where Γ_f is a submanifold of $X \times U$. If $\Gamma_f \subset M_g \equiv \operatorname{dom} g$, then the transform of Γ_f by g is defined as

$$g \cdot \Gamma_f = \{(\tilde{x}, \tilde{u}) = g \cdot (x, u) : (x, u) \in \Gamma_f\}.$$

We write $\tilde{f} = g \cdot f$ and call the function $\tilde{f}$ the *transform* of f by g.

For example, let $p = 1$ and $q = 1$, so $X = \mathbb{R}$ with a single independent variable x, and $U = \mathbb{R}$ with a single dependent variable u, so we have a single ODE involving a single function $u = f(x)$. Let $G = SO(2)$ be the rotation group acting on $X \times U \simeq \mathbb{R}^2$. The transformations in G are given by

$$(\tilde{x}, \tilde{u}) = \theta \cdot (x, u) = (x \cos \theta - u \sin \theta,\ x \sin \theta + u \cos \theta).$$

Let $u = f(x)$ be a function whose graph is a subset $\Gamma_f \subset X \times U$. The group $SO(2)$ acts on f by rotating its graph.

In general, the procedure for finding the transformed function $\tilde{f} = g \cdot f$ is given by [Olv86]:

$$g \cdot f = [\Phi_g \circ (1 \times f)] \circ [\Xi_g \circ (1 \times f)]^{-1}, \tag{2.36}$$

where $\Xi_g = \Xi_g(x, u)$, $\Phi_g = \Phi_g(x, u)$ are smooth functions such that

$$(\tilde{x}, \tilde{u}) = g \cdot (x, u) = (\Xi_g(x, u),\ \Phi_g(x, u)),$$

while 1 denotes the identity function of X, so $1(x) = x$. Formula (2.36) holds whenever the second factor is invertible.

Let $\mathcal{S}$ be a system of DEs. A *symmetry group* of the system $\mathcal{S}$ is a local Lie group of transformations G acting on an open subset $M \subset X \times U$ of the space $X \times U$ of independent and dependent variables of the system with the property that whenever $u = f(x)$ is a solution of $\mathcal{S}$, and whenever $g \cdot f$ is defined for $g \in G$, then $u = g \cdot f(x)$ is also a solution of the system.

For example, in the case of the ODE $u_{xx} = 0$, the rotation group $SO(2)$ is obviously a symmetry group, since the solutions are all linear functions and $SO(2)$ takes any linear function to another linear function. Another easy example is given by the classical *heat equation* $u_t = u_{xx}$. Here the group of translations

$$(x, t, u) \mapsto (x + \varepsilon a,\ t + \varepsilon b, u), \qquad \varepsilon \in \mathbb{R},$$

is a symmetry group since $u = f(x - \varepsilon a,\ t - \varepsilon b)$ is a solution to the heat equation whenever $u = f(x, t)$ is.

Prolongations

Prolongations of Functions

Given a smooth real–valued function $u = f(x) = f(x^1, ..., x^p)$ of p independent variables, there is an induced function $u^{(n)} = \mathbf{pr}^{(n)} f(x)$, called the nth *prolongation* of f [Olv86], which is defined by the equations

$$u_J = \partial_J f(x) = \frac{\partial^k f(x)}{\partial x^{j_1} \partial x^{j_2 \dots} \partial x^{j_k}},$$

where the multi–index $J = (j_1, ..., j_k)$ is an unordered $k-$tuple of integers, with entries $1 \leq j_k \leq p$ indicating which derivatives are being taken. More generally, if $f : X \to U$ is a smooth function from $X \simeq \mathbb{R}^p$ to $U \simeq \mathbb{R}^q$, so $u = f(x) = f(f^1(x), ..., f^q(x))$, there are $q \cdot p_k$ numbers

$$u_J^\alpha = \partial_J f^\alpha(x) = \frac{\partial^k f^\alpha(x)}{\partial x^{j_1} \partial x^{j_2 \dots} \partial x^{j_k}},$$

needed to represent all the different kth order derivatives of the components of f at a point x. Thus $\mathbf{pr}^{(n)} f : X \to U^{(n)}$ is a function from X to the space $U^{(n)}$, and for each $x \in X$, $\mathbf{pr}^{(n)} f(x)$ is a vector whose $q \cdot p^{(n)}$ entries represent the values of f and al its derivatives up to order n at the point x.

For example, in the case $p = 2$, $q = 1$ we have $X \simeq \mathbb{R}^2$ with coordinates $(x^1, x^2) = (x, y)$, and $U \simeq \mathbb{R}$ with the single coordinate $u = f(x, y)$. The second prolongation $u^{(2)} = \mathbf{pr}^{(2)} f(x, y)$ is given by [Olv86]

$$(u; u_x, u_y; u_{xx}, u_{xy}, u_{yy}) = \left(f; \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}; \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y^2} \right), \qquad (2.37)$$

all evaluated at (x, y).

The nth prolongation $\mathbf{pr}^{(n)} f(x)$ is also known as the $n-$jet of f. In other words, the nth prolongation $\mathbf{pr}^{(n)} f(x)$ represents the Taylor polynomial of degree n for f at the point x, since the derivatives of order $\leq n$ determine the Taylor polynomial and vice versa.

Prolongations of Differential Equations

A system $\mathcal{S}$ of nth order DEs in p independent and q dependent variables is given as a system of equations [Olv86]

$$\Delta_r(x, u^{(n)}) = 0, \qquad (r = 1, ..., l), \qquad (2.38)$$

involving $x = (x^1, ..., x^p)$, $u = (u^1, ..., u^q)$ and the derivatives of u with respect to x up to order n. The functions $\Delta(x, u^{(n)}) = (\Delta_1(x, u^{(n)}), ..., \Delta_l(x, u^{(n)}))$ are assumed to be smooth in their arguments, so $\Delta : X \times U^{(n)} \to \mathbb{R}^l$ represents a smooth map from the *jet space* $X \times U^{(n)}$ to some lD Euclidean space. The DEs themselves tell where the given map Δ vanishes on the jet space $X \times U^{(n)}$, and thus determine a submanifold

$$\mathcal{S}_\Delta = \left\{(x, u^{(n)}) : \Delta(x, u^{(n)}) = 0\right\} \subset X \times U^{(n)} \tag{2.39}$$

of the total the jet space $X \times U^{(n)}$.

We can identify the system of DEs (2.38) with its corresponding submanifold $\mathcal{S}_\Delta$ (2.39). From this point of view, a smooth *solution* of the given system of DEs is a smooth function $u = f(x)$ such that [Olv86]

$$\Delta_r(x, \mathbf{pr}^{(n)} f(x)) = 0, \qquad (r = 1, ..., l),$$

whenever x lies in the domain of f. This is just a restatement of the fact that the derivatives $\partial_J f^\alpha(x)$ of f must satisfy the algebraic constraints imposed by the system of DEs. This condition is equivalent to the statement that the graph of the prolongation $\mathbf{pr}^{(n)} f(x)$ must lie entirely within the submanifold $\mathcal{S}_\Delta$ determined by the system:

$$\Gamma_f^{(n)} \equiv \left\{(x, \mathbf{pr}^{(n)} f(x))\right\} \subset \mathcal{S}_\Delta = \left\{\Delta(x, u^{(n)}) = 0\right\}.$$

We can thus take an nth order system of DEs to be a submanifold $\mathcal{S}_\Delta$ in the $n-$jet space $X \times U^{(n)}$ and a solution to be a function $u = f(x)$ such that the graph of the nth prolongation $\mathbf{pr}^{(n)} f(x)$ is contained in the submanifold $\mathcal{S}_\Delta$.

For example, consider the case of *Laplace equation* in the plane

$$u_{xx} + u_{yy} = 0 \qquad \text{(remember,} \quad u_x \equiv \partial_x u).$$

Here $p = 2$ since there are two independent variables x and y, and $q = 1$ since there is one dependent variable u. Also $n = 2$ since the equation is second order, so $\mathcal{S}_\Delta \subset X \times U^{(2)}$ is given by (2.37). A solution $u = f(x, y)$ must satisfy

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

for all (x, y). This is the same as requiring that the graph of the second prolongation $\mathbf{pr}^{(2)} f$ lie in $\mathcal{S}_\Delta$.

Prolongations of Group Actions

Let G be a local group of transformations acting on an open subset $M \subset X \times U$ of the space of independent and dependent variables. There is an induced local action of G on the $n-$jet space $M^{(n)}$, called the nth prolongation $\mathbf{pr}^{(n)} G$ of the action of G on M. This prolongation is defined so that it transforms the derivatives of functions $u = f(x)$ into the corresponding derivatives of the transformed function $\tilde{u} = \tilde{f}(\tilde{x})$ [Olv86].

More precisely, suppose $(x_0, u_0^{(n)})$ is a given point in $M^{(n)}$. Choose any smooth function $u = f(x)$ defined in a neighborhood of x_0, whose graph Γ_f lies in M, and has the given derivatives at x_0 :

$$u_0^{(n)} = \mathbf{pr}^{(n)} f(x_0), \qquad \text{i.e.,} \qquad u_{J0}^{\alpha} = \partial_J f^{\alpha}(x_0).$$

If g is an element of G sufficiently near the identity, the transformed function $g \cdot f$ as given by (2.36) is defined in a neighborhood of the corresponding point $(\tilde{x}_0, \tilde{u}_0) = g \cdot (x_0, u_0)$, with $u_0 = f(x_0)$ being the zeroth order components of $u_0^{(n)}$. We then determine the action of the prolonged group of transformations $\mathbf{pr}^{(n)} g$ on the point $(x_0, u_0^{(n)})$ by evaluating the derivatives of the transformed function $g \cdot f$ at $\tilde{x}_0$; explicitly [Olv86]

$$\mathbf{pr}^{(n)} g \cdot (x_0, u_0^{(n)}) = (\tilde{x}_0, \tilde{u}_0^{(n)}),$$

where

$$\tilde{u}_0^{(n)} \equiv \mathbf{pr}^{(n)} (g \cdot f)(\tilde{x}_0).$$

For example, let $p = q = 1$, so $X \times U \simeq \mathbb{R}^2$, and consider the action of the rotation group $SO(2)$. To calculate its first prolongation $\mathbf{pr}^{(1)} SO(2)$, first note that $X \times U^{(1)} \simeq \mathbb{R}^3$, with coordinates (x, u, u_x). given a function $u = f(x)$, the first prolongation is [Olv86]

$$\mathbf{pr}^{(1)} f(x) = (f(x), f'(x)).$$

Now, given a point $(x^0, u^0, u_x^0) \in X \times U^{(1)}$, and a rotation in $SO(2)$ characterized by the angle θ as given above, the corresponding transformed point

$$\mathbf{pr}^{(1)} \theta \cdot (x^0, u^0, u_x^0) = (\tilde{x}^0, \tilde{u}^0, \tilde{u}_x^0)$$

(provided it exists). As for the first order derivative, we find

$$\tilde{u}_x^0 = \frac{\sin\theta + u_x \cos\theta}{\cos\theta - u_x \sin\theta}.$$

Now, applying the group transformations given above, and dropping the 0−indices, we find that the prolonged action $\mathbf{pr}^{(1)} SO(2)$ on $X \times U^{(1)}$ is given by

$$\mathbf{pr}^{(1)} \theta \cdot (x, u, u_x) = \left(x\cos\theta - u\sin\theta,\, x\sin\theta + u\cos\theta, \frac{\sin\theta + u_x \cos\theta}{\cos\theta - u_x \sin\theta} \right),$$

which is defined for $|\theta| < |\operatorname{arccot} u_x|$. Note that even though $SO(2)$ is a linear, globally defined group of transformations, its first prolongation $\mathbf{pr}^{(1)} SO(2)$ is both nonlinear and only locally defined. This fact demonstrates the complexity of the operation of prolonging a group of transformations.

In general, for any Lie group G, the first prolongation $\mathbf{pr}^{(1)} G$ acts on the original variables (x, u) exactly the same way that G itself does; only the action on the derivative u_x provides an new information. Therefore, $\mathbf{pr}^{(0)} G$ agrees with G itself, acting on $M^{(0)} = M$.

Prolongations of Vector Fields

Prolongation of the infinitesimal generators of the group action turn out to be the *infinitesimal generators* of the *prolonged group action* [Olv86]. Let $M \subset X \times U$ be open and suppose v is a vector–field on M, with corresponding local one–parameter group $\exp(\varepsilon v)$. The nth prolongation of v, denoted $\mathbf{pr}^{(n)} v$, will be a vector–field on the $n-$jet space $M^{(n)}$, and is defined to be the infinitesimal generator of the corresponding prolonged on–parameter group $\mathbf{pr}^{(n)}[\exp(\varepsilon v)]$. In other words,

$$\mathbf{pr}^{(n)} v|_{(x,u^{(n)})} = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \mathbf{pr}^{(n)}[\exp(\varepsilon v)](x, u^{(n)}) \tag{2.40}$$

for any $(x, u^{(n)}) \in M^{(n)}$.

For a vector–field v on M, given by

$$v = \xi^i(x,u)\frac{\partial}{\partial x^i} + \phi^\alpha(x,u)\frac{\partial}{\partial u^\alpha}, \qquad (i = 1, ..., p,\ \alpha = 1, ..., q),$$

the nth prolongation $\mathbf{pr}^{(n)} v$ is given by [Olv86]

$$\mathbf{pr}^{(n)} v = \xi^i(x,u)\frac{\partial}{\partial x^i} + \phi^\alpha_J(x, u^{(n)})\frac{\partial}{\partial u^\alpha_J},$$

with $\phi^\alpha_0 = \phi^\alpha$, and J a multiindex defined above.

For example, in the case of $SO(2)$ group, the corresponding infinitesimal generator is

$$v = -u\frac{\partial}{\partial x} + x\frac{\partial}{\partial u},$$

with

$$\exp(\varepsilon v)(x,u) = (x\cos\varepsilon - u\sin\varepsilon,\ x\sin\varepsilon + u\cos\varepsilon),$$

being the rotation through angle ε. The first prolongation takes the form

$$\mathbf{pr}^{(1)}[\exp(\varepsilon v)](x,u,u_x) = \left(x\cos\varepsilon - u\sin\varepsilon,\ x\sin\varepsilon + u\cos\varepsilon,\ \frac{\sin\varepsilon + u_x\cos\varepsilon}{\cos\varepsilon - u_x\sin\varepsilon}\right).$$

According to (2.40), the first prolongation of v is obtained by differentiating these expressions with respect to ε and setting $\varepsilon = 0$, which gives

$$\mathbf{pr}^{(1)} v = -u\frac{\partial}{\partial x} + x\frac{\partial}{\partial u} + (1 + u_x^2)\frac{\partial}{\partial u_x}.$$

General Prolongation Formula

Let

$$v = \xi^i(x,u)\frac{\partial}{\partial x^i} + \phi^\alpha(x,u)\frac{\partial}{\partial u^\alpha}, \qquad (i = 1, ..., p,\ \alpha = 1, ..., q), \tag{2.41}$$

be a vector–field defined on an open subset $M \subset X \times U$. The nth prolongation of v is the vector–field [Olv86]

$$\mathbf{pr}^{(n)} v = v + \phi_J^\alpha (x, u^{(n)}) \frac{\partial}{\partial u_J^\alpha}, \tag{2.42}$$

defined on the corresponding jet space $M^{(n)} \subset X \times U^{(n)}$. The coefficient functions ϕ_J^α are given by the following formula:

$$\phi_J^\alpha = D_J \left(\phi^\alpha - \xi^i u_i^\alpha \right) + \xi^i u_{J,i}^\alpha , \tag{2.43}$$

where $u_i^\alpha = \partial u^\alpha / \partial x^i$, and $u_{J,i}^\alpha = \partial u_J^\alpha / \partial x^i$. D_J is the *total derivative* with respect to the multiindex J, i.e.,

$$D_J = D_{j_1} D_{j_2} ... D_{j_k},$$

while the total derivative with respect to the ordinary index, D_i, is defined as follows. Let $P(x, u^{(n)})$ be a smooth function of x, u and derivatives of u up to order n, defined on an open subset $M^{(n)} \subset X \times U^{(n)}$. the total derivative of P with respect to x^i is the unique smooth function $D_i P(x, u^{(n)})$ defined on $M^{(n+1)}$ and depending on derivatives of u up to order $n+1$, with the *recursive property* that if $u = f(x)$ is any smooth function then

$$D_i P(x, \mathbf{pr}^{(n+1)} f(x)) = \partial_{x^i} \{ P(x, \mathbf{pr}^{(n)} f(x)) \}.$$

For example, in the case of $SO(2)$ group, with the infinitesimal generator

$$v = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u},$$

the first prolongation is (as calculated above)

$$\mathbf{pr}^{(1)} v = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u} + \phi^x \frac{\partial}{\partial u_x},$$

where

$$\phi^x = D_x (\phi - \xi u_x) + \xi u_{xx} = 1 + u_x^2.$$

Also,

$$\phi^{xx} = D_x \phi^x - u_{xx} D_x \xi = 3 u_x u_{xx},$$

thus the infinitesimal generator of the second prolongation $\mathbf{pr}^{(2)} SO(2)$ acting on $X \times U^{(2)}$ is

$$\mathbf{pr}^{(2)} v = -u \frac{\partial}{\partial x} + x \frac{\partial}{\partial u} + (1 + u_x^2) \frac{\partial}{\partial u_x} + 3 u_x u_{xx} \frac{\partial}{\partial u_{xx}}.$$

Let v and w be two smooth vector–fields on $M \subset X \times U$. Then their nth prolongations, $\mathbf{pr}^{(n)} v$ and $\mathbf{pr}^{(n)} w$ respectively, have the *linearity property*

$$\mathbf{pr}^{(n)} (c_1 v + c_2 w) = c_1 \mathbf{pr}^{(n)} v + c_2 \mathbf{pr}^{(n)} w, \qquad (c_1, c_2 - \text{constant}),$$

and the *Lie bracket property*

$$\mathbf{pr}^{(n)} [v, w] = [\mathbf{pr}^{(n)} v, \mathbf{pr}^{(n)} w].$$

Lie Symmetries of Special Biomechanical Equations

Here we consider two most important equations for human–like biomechanics:

1. The heat equation, which has been analyzed in muscular mechanics since the early works of A.V. Hill ([Hil38]); and
2. The Korteweg–De Vries equation, the basic equation for solitary models of muscular excitation–contraction dynamics (see subsection (3.2.3) below).

Suppose

$$\mathcal{S} : \varDelta_r(x, u^{(n)}) = 0, \qquad (r = 1, ..., l),$$

is a system of DEs of maximal rank defined over $M \subset X \times U$. If G is a local group of transformations acting on M, and

$$\mathbf{pr}^{(n)} v[\varDelta_r(x, u^{(n)})] = 0, \qquad \text{whenever} \qquad \varDelta(x, u^{(n)}) = 0, \tag{2.44}$$

(with $r = 1, ..., l$) for every infinitesimal generator v of G, then G is a symmetry group of the system $\mathcal{S}$ [Olv86].

The Heat Equation

The $(1+1)$D *heat equation* (with the thermal diffusivity normalized to unity)

$$u_t = u_{xx} \tag{2.45}$$

has two independent variables x and t, and one dependent variable u, so $p = 2$ and $q = 1$. Equation (A.30) has the second order, $n = 2$, and can be identified with the linear submanifold $M^{(2)} \subset X \times U^{(2)}$ determined by the vanishing of $\varDelta(x, t, u^{(2)}) = u_t - u_{xx}$.

Let

$$v = \xi(x, t, u)\frac{\partial}{\partial x} + \tau(x, t, u)\frac{\partial}{\partial t} + \phi(x, t, u)\frac{\partial}{\partial u}$$

be a vector–field on $X \times U$. According to (2.44) we need to now the second prolongation

$$\mathbf{pr}^{(2)} v = v + \phi^x \frac{\partial}{\partial u_x} + \phi^t \frac{\partial}{\partial u_t} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xt} \frac{\partial}{\partial u_{xt}} + \phi^{tt} \frac{\partial}{\partial u_{tt}}$$

of v. Applying $\mathbf{pr}^{(2)} v$ to (A.30) we find the infinitesimal criterion (2.44) to be

$$\phi^t = \phi^{xx},$$

which must be satisfied whenever $u_t = u_{xx}$.

The Korteweg–De Vries Equation

The Korteweg–De Vries equation

$$u_t + u_{xxx} + uu_x = 0 \tag{2.46}$$

arises in physical systems in which both nonlinear and dispersive effects are relevant. A vector–field

$$v = \xi(x,t,u)\frac{\partial}{\partial x} + \tau(x,t,u)\frac{\partial}{\partial t} + \phi(x,t,u)\frac{\partial}{\partial u}$$

generates a one–parameter symmetry group iff

$$\phi^t + \phi^{xxx} + u\phi^x + u_x\phi = 0,$$

whenever u satisfies (2.46), etc.

Generalized Lie Symmetries

Consider a vector–field (2.41) defined on an open subset $M \subset X \times U$. Provided the coefficient functions ξ^i and ϕ^α depend only on x and u, v will generate a (local) one–parameter group of transformations $\exp(\varepsilon v)$ acting pointwise on the underlying space M. A significant generalization of the notion of symmetry group is obtained by relaxing this geometric assumption, and allowing the coefficient functions ξ^i and ϕ^α to also depend on derivatives of u [Olv86].

A *generalized vector–field* is a (formal) expression

$$v = \xi^i[u]\frac{\partial}{\partial x^i} + \phi^\alpha[u]\frac{\partial}{\partial u^\alpha}, \qquad (i = 1,...,p,\ \alpha = 1,...,q), \tag{2.47}$$

in which ξ^i and ϕ^α are smooth functions. For example,

$$v = xu_x\frac{\partial}{\partial x} + u_{xx}\frac{\partial}{\partial u}$$

is a generalized vector in the case $p = q = 1$.

According to the general prolongation formula (2.42), we can define the *prolonged generalized vector–field*

$$\mathbf{pr}^{(n)}v = v + \phi_J^\alpha[u]\frac{\partial}{\partial u_J^\alpha},$$

whose coefficients are as before determined by the formula (2.43). Thus, in our previous example [Olv86],

$$\mathbf{pr}^{(n)}v = xu_x\frac{\partial}{\partial x} + u_{xx}\frac{\partial}{\partial u} + [u_{xxx} - (xu_{xx} + u_x)u_x]\frac{\partial}{\partial u_x}.$$

Given a generalized vector–field v, its *infinite prolongation* (including all the derivatives) is the formal expression

$$\mathbf{pr}\, v = \xi^i \frac{\partial}{\partial x^i} + \phi^\alpha_J \frac{\partial}{\partial u^\alpha_J}.$$

Now, a generalized vector–field v is a *generalized infinitesimal symmetry* of a system $\mathcal{S}$ of differential equations

$$\Delta_r[u] = \Delta_r(x, u^{(n)}) = 0, \qquad (r = 1, ..., l),$$

iff

$$\mathbf{pr}\, v[\Delta_r] = 0$$

for every smooth solution m $u = f(x)$ [Olv86].

For example, consider the heat equation

$$\Delta[u] = u_t - u_{xx} = 0.$$

The generalized vector–field $v = u_x \frac{\partial}{\partial u}$ has prolongation

$$\mathbf{pr}\, v = u_x \frac{\partial}{\partial u} + u_{xx} \frac{\partial}{\partial u_x} + u_{xt} \frac{\partial}{\partial u_t} + u_{xxx} \frac{\partial}{\partial u_{xx}} + ...$$

Thus

$$\mathbf{pr}\, v(\Delta) = u_{xt} - u_{xxx} = D_x(u_t - u_{xx}) = D_x \Delta,$$

and hence v is a generalized symmetry of the heat equation.

Noether Symmetries

Here we present some results about *Noether symmetries*, in particular for the first order Lagrangians $L(q, \dot{q})$ (see [BGG89, PSS96]). We start with a *Noether Lagrangian symmetry*,

$$\delta L = \dot{F},$$

and we will investigate the conversion of this symmetry to the Hamiltonian formalism. Defining

$$G = (\partial L/\partial \dot{q}^i)\, \delta q^i - F,$$

we can write

$$\delta_i L\, \delta q^i + \dot{G} = 0, \tag{2.48}$$

where $\delta_i L$ is the *Euler–Lagrange functional derivative* of L,

$$\delta_i L = \alpha_i - W_{ik}\, \ddot{q}^k,$$

where

$$W_{ik} \equiv \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^k} \qquad \text{and} \qquad \alpha_i \equiv -\frac{\partial^2 L}{\partial \dot{q}^i \partial q^k}\, \dot{q}^k + \frac{\partial L}{\partial q^i}.$$

We consider the general case where the mass matrix, or *Hessian* (W_{ij}), may be a singular matrix. In this case there exists a kernel for the pull–back $\mathbb{F}L^*$ of the *Legendre map*, i.e., *fiber–derivative* $\mathbb{F}L$, from the velocity phase–space TM (tangent bundle of the biomechanical manifold M) to the momentum phase–space T^*M (cotangent bundle of M). This kernel is spanned by the vector–fields

$$\Gamma_\mu = \gamma^i_\mu \frac{\partial}{\partial \dot{q}^i},$$

where γ^i_μ are a basis for the null vectors of W_{ij}. The Lagrangian time–evolution differential operator can therefore be expressed as:

$$X = \partial_t + \dot{q}^k \frac{\partial}{\partial q^k} + a^k(q, \dot{q}) \frac{\partial}{\partial \dot{q}^k} + \lambda^\mu \Gamma_\mu \equiv X_o + \lambda^\mu \Gamma_\mu,$$

where a^k are functions which are determined by the formalism, and λ^μ are arbitrary functions. It is not necessary to use the Hamiltonian technique to find the Γ_μ, but it does facilitate the calculation:

$$\gamma^i_\mu = \mathbb{F}L^* \left(\frac{\partial \phi_\mu}{\partial p_i} \right), \tag{2.49}$$

where the ϕ_μ are the Hamiltonian primary first class constraints.

Notice that the highest derivative in (2.48), $\ddot{q}^i$, appears linearly. Because δL is a symmetry, (2.48) is identically satisfied, and therefore the coefficient of $\ddot{q}^i$ vanishes:

$$W_{ik} \delta q^k - \frac{\partial G}{\partial \dot{q}^i} = 0. \tag{2.50}$$

We contract with a null vector γ^i_μ to find that

$$\Gamma_\mu G = 0.$$

It follows that G is projectable to a function G_{H} in T^*Q; that is, it is the pull–back of a function (not necessarily unique) in T^*Q:

$$G = \mathbb{F}L^*(G_{\mathrm{H}}).$$

This important property is valid for any conserved quantity associated with a Noether symmetry. Observe that G_{H} is determined up to the addition of linear combinations of the primary constraints. Substitution of this result in (2.50) gives

$$W_{ik} \left[\delta q^k - \mathbb{F}L^* \left(\frac{\partial G_{\mathrm{H}}}{\partial p_k} \right) \right] = 0,$$

and so the brackets enclose a null vector of W_{ik}:

$$\delta q^i - \mathbb{F}L^* \left(\frac{\partial G_{\mathrm{H}}}{\partial p_i} \right) = r^\mu \gamma^i_\mu, \tag{2.51}$$

for some $r^\mu(t, q, \dot{q})$.

We shall investigate the projectability of variations generated by diffeomorphisms in the following section. Assume that an infinitesimal transformation δq^i is projectable:

$$\Gamma_\mu \delta q^i = 0.$$

If δq^i is projectable, so must be r^μ, so that $r^\mu = \mathbb{F}L^*(r^\mu_{\mathrm{H}})$. Then, using (2.49) and (2.51), we see that

$$\delta q^i = \mathbb{F}L^* \left(\frac{\partial (G_H + r^\mu_{\mathrm{H}} \phi_\mu)}{\partial p_i} \right).$$

We now redefine G_{H} to absorb the piece $r^\mu_{\mathrm{H}} \phi_\mu$, and from now on we will have

$$\delta q^i = \mathbb{F}L^* \left(\frac{\partial G_{\mathrm{H}}}{\partial p_i} \right).$$

Define

$$\hat{p}_i = \frac{\partial L}{\partial \dot{q}^i};$$

after eliminating (2.50) times $\ddot{q}^i$ from (2.48), we get

$$\left(\frac{\partial L}{\partial q^i} - \dot{q}^k \frac{\partial \hat{p}_i}{\partial q^k} \right) \mathbb{F}L^* (\frac{\partial G_{\mathrm{H}}}{\partial p_i}) + \dot{q}^i \frac{\partial}{\partial q^i} \mathbb{F}L^*(G_{\mathrm{H}}) + \mathbb{F}L^* \partial_t G_{\mathrm{H}} = 0,$$

which simplifies to

$$\frac{\partial L}{\partial q^i} \mathbb{F}L^* (\frac{\partial G_{\mathrm{H}}}{\partial p_i}) + \dot{q}^i \mathbb{F}L^* (\frac{\partial G_{\mathrm{H}}}{\partial q^i}) + \mathbb{F}L^* \partial_t G_{\mathrm{H}} = 0. \tag{2.52}$$

Now let us invoke two identities [BGG89] that are at the core of the connection between the Lagrangian and the Hamiltonian equations of motion. They are

$$\dot{q}^i = \mathbb{F}L^* (\frac{\partial H}{\partial p_i}) + v^\mu(q, \dot{q}) \mathbb{F}L^* (\frac{\partial \phi_\mu}{\partial p_i}),$$

and

$$\frac{\partial L}{\partial q^i} = -\mathbb{F}L^* (\frac{\partial H}{\partial q^i}) - v^\mu(q, \dot{q}) \mathbb{F}L^* (\frac{\partial \phi_\mu}{\partial q^i});$$

where H is any canonical Hamiltonian, so that $\mathbb{F}L^*(H) = \dot{q}^i(\partial L/\partial \dot{q}^i) - L = \hat{E}$, the Lagrangian energy, and the functions v^μ are determined so as to render the first relation an identity. Notice the important relation

$$\Gamma_\mu v^\nu = \delta^\nu_\mu,$$

which stems from applying Γ_μ to the first identity and taking into account that

$$\Gamma_\mu \circ \mathbb{F}L^* = 0.$$

Substitution of these two identities into (2.52) yields (where { , } denotes the *Poisson bracket*)

$$\mathbb{F}L^*\{G_{\rm H}, H\} + v^\mu \mathbb{F}L^*\{G_{\rm H}, \phi_\mu\} + \mathbb{F}L^*\partial_t G_{\rm H} = 0.$$

This result can be split through the action of Γ_μ into

$$\mathbb{F}L^*\{G_{\rm H}, H\} + \mathbb{F}L^*\partial_t G_{\rm H} = 0,$$

and

$$\mathbb{F}L^*\{G_{\rm H}, \phi_\mu\} = 0;$$

or equivalently,

$$\{G_{\rm H}, H\} + \partial_t G_{\rm H} = pc,$$

and

$$\{G_{\rm H}, \phi_\mu\} = pc,$$

where pc stands for any linear combination of primary constraints. In this way, we have arrived at a neat characterization for a generator $G_{\rm H}$ of Noether transformations in the canonical formalism.

Lie–Invariant Differential Forms

Robot Kinematics

Recall that a typical *motion planning* problem in robotics consists in a collection of objects moving around obstacles from an initial to a final configuration (see [BL92, Pry96]). This may include in particular, solving the *collision detection* problem.

When a solid object undergoes a rigid motion, the totality of points through which it passed constitutes a region in space called the *swept volume*. To describe the geometric structure of the swept volume we pose this problem as one of geometric study of some manifold swept by surface points using powerful tools from both modern differential geometry and nonlinear dynamical systems theory [Ric93, LP94, Pry96, GJ94] on manifolds. For some special cases of the Euclidean motion in the space $\mathbb{R}^3$ one can construct a very rich hydrodynamic system [BL92] modelling a sweep flow, which appears to be a *completely integrable Hamiltonian system* having a special *Lax type representation*. To describe in detail these and other properties of swept volume dynamical systems, we develop *Cartan's theory of Lie–invariant geometric objects* generated by closed ideals in the Grassmann algebra, following [BPS98].

Let a Lie group G act on an analytical manifold Y in the transitive way, that is the action $G \times Y \xrightarrow{\rho} Y$ generates some nonlinear exact representation of the Lie group G on the manifold Y. In the frame of the Cartan's theory, the representation $G \times Y \xrightarrow{\rho} Y$ can be described by means of a system of differential 1–forms (see section 3.3.6 below)

$$\bar{\beta}^j = dy^j + \xi^j_i \bar{\omega}^i(a, da) \qquad (2.53)$$

in the Grassmann algebra $\Lambda(Y \times G)$ on the product $Y \times G$, where $\bar{\omega}^i(a, da) \in T^*_a(G)$, $i = 1, ..., r = \dim G$ is a basis of left invariant Cartan's forms of the Lie group G at a point $a \in G$, $y = \{y^j : j = 1, ..., n = \dim Y\} \in Y$ and $\xi^j_i : Y \times G \to \mathbb{R}$ are some smooth real valued functions.

The following *Cartan theorem* is basic in describing a geometric object invariant with respect to the mentioned above group action $G \times Y \xrightarrow{\rho} Y$: The system of differential forms (2.53) is a system of an invariant geometric object iff the following conditions are fulfilled:

1. The coefficients $\xi^j_i \in C^k(Y; R)$ for all $i = 1, ..., r$, $j = 1, ..., n$, are some *analytical functions* on Y; and
2. The differential system (2.53) is completely integrable within the *Frobenius–Cartan criterion*.

The Cartan's theorem actually says that the differential system (2.53) can be written down as

$$\bar{\beta}^j = dy^j + \xi^j_i(y) \bar{\omega}^i(a, da), \qquad (2.54)$$

where 1–forms $\{\bar{\omega}^i(a, da) : i = 1, ..., r\}$ satisfy the standard *Maurer–Cartan equations*

$$\bar{\Omega}^j = d\bar{\omega}^j + \frac{1}{2} c^j_{ik} \bar{\omega}^i \wedge \bar{\omega}^k = 0, \qquad (2.55)$$

for all $j = 1, ..., r$ on G, coefficients $c^j_{ik} \in \mathbb{R}$, $i, j, k = 1, ..., r$, being the corresponding structure constants of the Lie algebra $\mathfrak{g}$ of the Lie group G.

Maurer–Cartan 1–Forms

Consider a Lie group G with the Lie algebra $\mathfrak{g} \simeq T_e(G)$, whose basis is a set $\{A_i \in \mathfrak{g} : i = 1, ..., r\}$, where $r = \dim\, G \equiv \dim\, \mathfrak{g}$. Let also a set $U_0 \subset \{a^i \in \mathbb{R} : i = 1, ..., r\}$ be some open neighborhood of the zero point in $\mathbb{R}^r$. The exponential mapping $\exp : U_0 \to G_0$, where by definition [BPS98]

$$\mathbb{R}^r \supset U_0 \ni (a^1, \ldots, a^r) : \xrightarrow{\exp} \exp\left(a^i A_i\right) = a \in G_0 \subset G, \qquad (2.56)$$

is an analytical mapping of the whole U_0 on some open neighborhood G_0 of the unity element $e \in G$. From (2.56) it is easy to find that $T_e(G) = T_e(G_0) \simeq \mathfrak{g}$, where $e = \exp(0) \in G$. Define now the following left invariant $\mathfrak{g}-$valued differential 1–form on $G_0 \subset G$:

$$\bar{\omega}(a, da) = a^{-1} da = \bar{\omega}^j(a, da) A_j, \qquad (2.57)$$

where $A_j \in \mathfrak{g}$, $\bar{\omega}^j(a, da) \in T^*_a(G)$, $a \in G_0$, $j = 1, ..., r$. To build effectively the unknown forms $\{\bar{\omega}^j(a, da) : j = 1, ..., r\}$, let us consider the following analytical one–parameter 1–form $\bar{\omega}_t(a, da) = \bar{\omega}(a_t; da_t)$ on G_0, where

$a_t = \exp\left(ta^i A_i\right)$, $t \in [0,1]$, and differentiate this form with respect to the parameter $t \in [0,1]$. We will get [BPS98]

$$d\bar{\omega}_t/dt = -a^j A_j a_t^{-1} da_t + a_t^{-1} a_t da^j A_j + a_t^{-1} da_t a^j A_j = -a^j [A_j, \bar{\omega}_t] + A_j da_j. \tag{2.58}$$

Having used the Lie identity, $[A_j, A_k] = c^i_{jk} A_i$, $j,k = 1,...,r$, and the r.h.s of (2.57) in form

$$\bar{\omega}^j(a, da) = \bar{\omega}^j_k(a) da^k, \tag{2.59}$$

we finally get

$$\frac{d}{dt}(t\bar{\omega}^j_i(ta)) = \mathcal{A}^j_k t\bar{\omega}^k_i(ta) + \delta^j_i, \tag{2.60}$$

with

$$\mathcal{A}^k_i = c^k_{ij} a^j. \tag{2.61}$$

The series solution of (2.60) is [BPS98]

$$\bar{\omega}^j_k(a) = W^j_k(t)\Big|_{t=1} = \sum_{n=1}^{\infty} \frac{1}{n!} \mathcal{A}^{n-1}. \tag{2.62}$$

General Structure of Integrable One–Forms

Given 2−forms generating a closed ideal $\mathfrak{I}(\alpha)$ in the Grassmann algebra $\Lambda(M)$, we will denote by $\mathfrak{I}(\alpha,\beta)$ an augmented ideal in $\Lambda(M;Y)$, where the manifold Y will be called in further the *representation space* of some adjoint Lie group G action: $G \times Y \xrightarrow{\rho} Y$. In this way, we can define the set of 1–forms $\{\beta\}$ and 2–forms $\{\alpha\}$

$$\begin{aligned} \{\alpha\} &= \{\alpha^i \in \Lambda^2(M) : \, i = 1,...,m_\alpha\}, \\ \{\beta\} &= \{\beta^j \in \Lambda^1(M \times Y) : \, j = 1,...,n = \dim Y\}, \end{aligned} \tag{2.63}$$

satisfying [BPS98]:

$$d\alpha^i = a^i_k(\alpha) \wedge \alpha^k, \qquad d\beta^j = f^j_k \alpha^k + \omega^j_s \wedge \beta^s, \tag{2.64}$$

where $a^i_k(\alpha) \in \Lambda^1(M)$, $f^j_k \in \Lambda^0(M \times Y)$ and $\omega^j_s \in \Lambda^1(M \times Y)$ for all $i,k = 1,...,m_\alpha$, $j,s = 1,...,n$. Since the identity $d^2\beta^j \equiv 0$ takes place for all $j = 1,...,n$, from (2.64) it follows that

$$\left(d\omega^j_k + \omega^j_s \wedge \omega^s_k\right) \wedge \beta^k + \left(df^j_s + \omega^j_k f^k_s + f^j_l a^l_s(\alpha)\right) \wedge \alpha^s \equiv 0. \tag{2.65}$$

From (2.65) we get [BPS98]

$$d\omega^j_k + \omega^j_s \wedge \omega^s_k \in \mathfrak{I}(\alpha,\beta), \qquad df^j_s + \omega^j_k f^k_s + f^j_l a^l_s(\alpha) \in \mathfrak{I}(\alpha,\beta) \tag{2.66}$$

for all $j,k = 1,...,n$, $s = 1,...,m_\alpha$. The second inclusion in (2.66) gives a possibility to define the 1–forms $\theta^j_s = f^j_l a^l_s(\alpha)$ satisfying the inclusion

$$d\theta_s^j + \omega_k^j \wedge \theta_s^k \in \mathfrak{I}(\alpha, \beta) \oplus f_l^j c_s^l(\alpha), \tag{2.67}$$

which we got using the identities $d^2\alpha^j \equiv 0$, $j = 1, ..., m_\alpha$, in the form $c_s^j(\alpha) \wedge \alpha^s \equiv 0$,

$$c_s^j(\alpha) = da_s^j(\alpha) + a_l^j(\alpha) \wedge a_s^l(\alpha), \tag{2.68}$$

following from (2.64). Further, if $s = s_0$ the 2–forms $c_{s_0}^j(\alpha) \equiv 0$ for all $j = 1, ..., m_\alpha$, then as $s = s_0$, we can define a set of 1–forms $\theta^j = \theta_{s_0}^j \in \Lambda^1(M \times Y)$, $j = 1, ..., n$, satisfying the exact inclusions

$$d\theta^j + \omega_k^j \wedge \theta^k = \Theta^j \in \mathfrak{I}(\alpha, \beta), \tag{2.69}$$

together with a set of inclusions for 1–forms $\omega_k^j \in \Lambda^1(M \times Y)$

$$d\omega_k^j + \omega_s^j \wedge \omega_k^s = \Omega_k^j \in \mathfrak{I}(\alpha, \beta). \tag{2.70}$$

Using the *general theory of connections on the fibered frame space* $P(M; GL(n))$ over a base manifold M (see [SW72]), we can interpret the equations (2.70) as defining the curvature 2–forms $\Omega_k^j \in \Lambda^2(P)$, and (2.69) as defining the torsion 2–forms $\Theta^j \in \Lambda^2(P)$. Since $\mathfrak{I}(\alpha) = 0 = \mathfrak{I}(\alpha, \beta)$ upon the integral submanifold $\bar{M} \subset M$, the reduced fibered frame space $P(\bar{M}; GL(n))$ will have the flat curvature and be torsion free, being as a result, completely trivialized on $\bar{M} \subset M$.

Lax Integrable Dynamical Systems

Consider some set $\{\beta\}$ defining a Cartan's Lie group G invariant object on a manifold $M \times Y$:

$$\beta^j = dy^j + \xi_k^j(y) b^k(z), \tag{2.71}$$

where $i = 1, ..., n = \dim Y$, $r = \dim G$. (2.71) defines a set $\{\xi\}$ of vector–fields on Y, giving a representation $\rho : \mathfrak{g} \to \{\xi\}$ of a given Lie algebra $\mathfrak{g}$. In other words, for the vector–fields $\xi_s = \xi_s^j(y) \frac{\partial}{\partial y^j} \in \{\xi\}$, $s = 1, ..., r$the following Lie algebra $\mathfrak{g}$ relationships are valid

$$[\xi_s, \xi_l] = c_{sl}^k \xi_k, \qquad (s, l, k = 1, ..., r). \tag{2.72}$$

Now, we can compute the differentials $d\beta^j \in \Lambda^2(M \times Y)$, $j = 1, ..., n$, using (2.71) and (2.72) as [BPS98]:

$$d\beta^j = \frac{\partial \xi_k^j(y)}{\partial y^l}(\beta^l - \xi_s^l(y) b^s(z)) \wedge b^k(z) + \xi_k^j(y) db^k(z), \tag{2.73}$$

which is equal to

$$\frac{\partial \xi_k^j(y)}{\partial y^l}\beta^l \wedge b_k(z) + \xi_l^j(db^l(z) + \frac{1}{2}c_{ks}^l db^k(z) \wedge db^s(z)),$$

where $\{\alpha\} \subset \Lambda^2(M)$ is some *a'priori* given integrable system of 2–forms on M, vanishing upon the integral submanifold $\bar{M} \subset M$.

Example: Burgers Dynamical System

Consider the *Burgers dynamical system* on a *functional manifold* $M \subset C^k(\mathbb{R};\mathbb{R})$:

$$u_t = uu_x + u_{xx}, \tag{2.74}$$

where $u \in M$ and $t \in \mathbb{R}$ is an evolution (time) parameter. The flow of (2.74) on M can be recast into a set of 2–forms $\{\alpha\} \subset \Lambda^2(J(\mathbb{R}^2;\mathbb{R}))$ upon the adjoint jet–manifold $J(\mathbb{R}^2;\mathbb{R})$ (see section 3.3.6 below) as follows [BPS98]:

$$\begin{aligned} &\{\alpha\} = \big\{du^{(0)} \wedge dt - u^{(1)}dx \wedge dt = \alpha^1,\ du^{(0)} \wedge dx + u^{(0)}du^{(0)} \wedge dt \\ &+du^{(1)} \wedge dt = \alpha^2 :\ \big(x, t; u^{(0)}, u^{(1)}\big)^\tau \in M^4 \subset J^1(\mathbb{R}^2;\mathbb{R})\big\}, \end{aligned} \tag{2.75}$$

where M^4 is some finite–dimensional submanifold in $J^1(\mathbb{R}^2;\mathbb{R}))$ with coordinates $(x, t, u^{(0)} = u, u^{(1)} = u_x)$. The set of 2–forms (2.75) generates the closed ideal $\mathfrak{I}(\alpha)$, since

$$d\alpha^1 = dx \wedge \alpha^2 - u^{(0)}dx \wedge \alpha^1, \qquad d\alpha^2 = 0, \tag{2.76}$$

the integral submanifold $\bar{M} = \{x, t \in \mathbb{R}\} \subset M^4$ being defined by the condition $\mathfrak{I}(\alpha) = 0$. We now look for a reduced 'curvature' 1–form $\Gamma \in \Lambda^1(M^4) \otimes \mathfrak{g}$, belonging to some (not yet determined) Lie algebra $\mathfrak{g}$. This 1–form can be represented using (2.75), as follows:

$$\Gamma = b^{(x)}(u^{(0)}, u^{(1)})dx + b^{(t)}(u^{(0)}, u^{(1)})dt, \tag{2.77}$$

where elements $b^{(x)}, b^{(t)} \in \mathfrak{g}$ satisfy [BPS98]

$$\begin{aligned} &\tfrac{\partial b^{(x)}}{\partial u^{(0)}} = g_2, \qquad \tfrac{\partial b^{(x)}}{\partial u^{(1)}} = 0, \qquad \tfrac{\partial b^{(t)}}{\partial u^{(0)}} = g_1 + g_2 u^{(0)}, \\ &\tfrac{\partial b^{(t)}}{\partial u^{(1)}} = g_2, \qquad [b^{(x)}, b^{(t)}] = -u^{(1)}g_1. \end{aligned} \tag{2.78}$$

The set (2.78) has the following unique solution

$$b^{(x)} = A_0 + A_1 u^{(0)}, \qquad b^{(t)} = u^{(1)}A_1 + \frac{{u^{(0)}}^2}{2}A_1 + [A_1, A_0]u^{(0)} + A_2, \tag{2.79}$$

where $A_j \in \mathfrak{g}$, $j = \overline{0, 2}$, are some constant elements on M of a Lie algebra $\mathfrak{g}$ under search, satisfying the next Lie structure equations:

$$\begin{aligned} &[A_0, A_2] = 0, \\ &[A_0, [A_1, A_0]] + [A_1, A_2] = 0, \\ &[A_1, [A_1, A_0]] + \tfrac{1}{2}[A_0, A_1] = 0. \end{aligned} \tag{2.80}$$

From (2.78) one can see that the curvature 2–form $\Omega \in \operatorname{span}_{\mathbb{R}}\{A_1, [A_0, A_1] : A_j \in \mathfrak{g}, j = 0, 1\}$. Therefore, reducing via the *Ambrose–Singer theorem* the associated principal fibered frame space $P(M; G = GL(n))$ to the principal fibre bundle $P(M; G(h))$, where $G(h) \subset G$ is the corresponding holonomy Lie group of the connection Γ on P, we need to satisfy the following conditions for the set $\mathfrak{g}(h) \subset \mathfrak{g}$ to be a Lie subalgebra in $\mathfrak{g} : \nabla_x^m \nabla_t^n \Omega \in \mathfrak{g}(h)$ for all $m, n \in \mathbb{Z}_+$.

Let us try now to close the above procedure requiring that [BPS98]

$$\mathfrak{g}(h) = \mathfrak{g}(h)_0 = \mathrm{span}_{\mathbb{R}}\{\nabla_x^m \nabla_x^n \Omega \in \mathfrak{g} : \ m + n = 0\} \tag{2.81}$$

This means that

$$\mathfrak{g}(h)_0 = \mathrm{span}_{\mathbb{R}}\{A_1, A_3 = [A_0, A_1]\}. \tag{2.82}$$

To satisfy the set of relations (2.80) we need to use expansions over the basis (2.82) of the external elements $A_0, A_2 \in \mathfrak{g}(h)$:

$$A_0 = q_{01}A_1 + q_{13}A_3, \qquad A_2 = q_{21}A_1 + q_{23}A_3. \tag{2.83}$$

Substituting expansions (2.83) into (2.80), we get that $q_{01} = q_{23} = \lambda$, $q_{21} = -\lambda^2/2$ and $q_{03} = -2$ for some arbitrary real parameter $\lambda \in \mathbb{R}$, that is $\mathfrak{g}(h) = \mathrm{span}_{\mathbb{R}}\{A_1, A_3\}$, where

$$[A_1, A_3] = A_3/2; \qquad A_0 = \lambda A_1 - 2A_3, \qquad A_2 = -\lambda^2 A_1/2 + \lambda A_3. \tag{2.84}$$

As a result of (2.84) we can state that the holonomy Lie algebra $\mathfrak{g}(h)$ is a real 2D one, assuming the following $(2 \times 2)-$matrix representation [BPS98]:

$$\begin{aligned} A_1 &= \begin{pmatrix} 1/4 & 0 \\ 0 & -1/4 \end{pmatrix}, \qquad & A_3 &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ A_0 &= \begin{pmatrix} \lambda/4 & -2 \\ 0 & -\lambda/4 \end{pmatrix}, \qquad & A_2 &= \begin{pmatrix} -\lambda^2/8 & \lambda \\ 0 & \lambda^2/8 \end{pmatrix}. \end{aligned} \tag{2.85}$$

Thereby from (2.77), (2.79) and (2.85) we get the *reduced curvature 1–form* $\Gamma \in \Lambda^1(M) \otimes \mathfrak{g}$,

$$\Gamma = (A_0 + uA_1)dx + ((u_x + u^2/2)A_1 - uA_3 + A_2)dt, \tag{2.86}$$

generating *parallel transport* of vectors from the representation space Y of the holonomy Lie algebra $\mathfrak{g}(h)$:

$$dy + \Gamma y = 0, \tag{2.87}$$

upon the integral submanifold $\bar{M} \subset M^4$ of the ideal $\mathcal{I}(\alpha)$, generated by the set of 2–forms (2.75). The result (2.87) means also that the Burgers dynamical system (2.74) is endowed with the standard Lax type representation, having the spectral parameter $\lambda \in \mathbb{R}$ necessary for its integrability in quadratures.

2.5 Riemannian Geometry in Human–Like Biomechanics

In this section we develop the basic techniques of Riemannian geometry on the biomechanical manifold M, from both local and global perspective. We start with the local Riemannian notions of metric, geodesics and curvature on M, including the first variation formula and parallel transport along the vector–fields on M. After that we move to the global Riemannian notions on M, including the second variation and Gauss–Bonnet formulae, as well as the global Ricci flow on M. The last part of the section presents the structure equations on M, the basics of Morse theory (as a preparation for the next Chapter), and the basics of (co)bordism theory.

2.5.1 Local Riemannian Geometry on M

An important class of problems in Riemannian geometry is to understand the interaction between the curvature and topology on a differentiable manifold [CC99]. A prime example of this interaction is the *Gauss–Bonnet formula* on a closed surface M^2, which says

$$\int_M K\, dA = 2\pi\, \chi(M), \tag{2.88}$$

where dA is the area element of a metric g on M, K is the Gaussian curvature of g, and $\chi(M)$ is the Euler characteristic of M.

To study the geometry of a differentiable manifold we need an additional structure: the *Riemannian metric*. The metric is an inner product on each of the tangent spaces and tells us how to measure angles and distances infinitesimally. In local coordinates $(x^1, x^2, \cdots, x^n)$, the metric g is given by $g_{ij}(x)\, dx^i \otimes dx^j$, where $(g_{ij}(x))$ is a positive definite symmetric matrix at each point x. For a differentiable manifold one can differentiate functions. A Riemannian metric defines a natural way of differentiating vector–fields: *covariant differentiation*. In Euclidean space, one can change the order of differentiation. On a Riemannian manifold the commutator of twice covariant differentiating vector–fields is in general nonzero and is called the *Riemann curvature tensor*, which is a $4-$tensor–field on the manifold.

For surfaces, the Riemann curvature tensor is equivalent to the *Gaussian curvature* K, a scalar function. In dimensions 3 or more, the Riemann curvature tensor is inherently a tensor–field. In local coordinates, it is denoted by R_{ijkl}, which is anti-symmetric in i and k and in j and l, and symmetric in the pairs $\{ij\}$ and $\{kl\}$. Thus, it can be considered as a bilinear form on $2-$forms which is called the *curvature operator*. We now describe heuristically the various curvatures associated to the *Riemann curvature tensor*. Given a point $x \in M^n$ and 2-plane Π in the tangent space of M at x, we can define a surface S in M to be the union of all geodesics passing through x and tangent to Π. In a neighborhood of x, S is a smooth 2D submanifold of M. We define *the sectional curvature* $K(\Pi)$ of the $2-$plane to be the Gauss curvature of S at x:

$$K(\Pi) = K_S(x).$$

Thus the sectional curvature K of a Riemannian manifold associates to each 2-plane in a tangent space a real number. Given a line L in a tangent space, we can average the sectional curvatures of all planes through L to get the *Ricci curvature* $Rc(L)$. Likewise, given a point $x \in M$, we can average the Ricci curvatures of all lines in the tangent space of x to get the *scalar curvature* $R(x)$. In local coordinates, the *Ricci tensor* is given by $R_{ik} = g^{jl}R_{ijkl}$ and the scalar curvature is given by $R = g^{ik}R_{ik}$, where $(g^{ij}) = (g_{ij})^{-1}$ is the inverse of the metric tensor (g_{ij}).

Riemannian Metric on M

In this subsection we mainly follow [Pet99, Pet98].

Riemann in 1854 observed that around each point $m \in M$ one can pick a *special* coordinate system $(x^1, \dots, x^n)$ such that there is a symmetric $(0, 2)-$tensor–field $g_{ij}(m)$ called the *metric tensor* defined as

$$g_{ij}(m) = g(\partial_{x^i}, \partial_{x^j}) = \delta_{ij}, \qquad \partial_{x^k} g_{ij}(m) = 0.$$

Thus the metric, at the specified point $m \in M$, in the coordinates $(x^1, \dots, x^n)$ looks like the Euclidean metric on $\mathbb{R}^n$. We emphasize that these conditions only hold at the specified point $m \in M$. When passing to different points it is necessary to pick different coordinates. If a curve γ passes through m, say, $\gamma(0) = m$, then the acceleration at 0 is simply defined by firstly, writing the curve out in our special coordinates

$$\gamma(t) = (\gamma^1(t), \dots, \gamma^n(t)),$$

secondly, defining the tangent, *velocity* vector–field, as

$$\dot{\gamma} = \dot{\gamma}^i(t) \cdot \partial_{x^i},$$

and finally, the *acceleration* vector–field as

$$\ddot{\gamma}(0) = \ddot{\gamma}^i(0) \cdot \partial_{x^i}.$$

Here, the background idea is that we have a *connection*.

A vector–field X along a parameterized curve $\alpha : I \to M$ in M is *tangent to M along* α if $X(t) \in M_{\alpha(t)}$ for all for $t \in I \subset \mathbb{R}$. The derivative $\dot{X}$ of such a vector–field is, however, generally not tangent to M. We can, nevertheless, get a vector–field tangent to M by projecting $\dot{X}(t)$ orthogonally onto $M_{\alpha(t)}$ for each $t \in I$. This process of differentiating and then projecting onto the tangent space to M defines an operation with the same properties as differentiation, except that now differentiation of vector–fields tangent to M yields vector–fields tangent to M. This operation is called *covariant differentiation*.

Let $\gamma : I \to M$ be a parameterized curve in M, and let X be a smooth vector–field tangent to M along α. The *absolute covariant derivative* of X is the vector–field $\dot{\bar{X}}$ tangent to M along α, defined by $\dot{\bar{X}} = \dot{X}(t) - [\dot{X}(t) \cdot N(\alpha(t))]\, N(\alpha(t))$, where N is an orientation on M. Note that $\dot{\bar{X}}$ is independent of the choice of N since replacing N by $-N$ has no effect on the above formula.

Lie bracket (2.4.1) defines a *symmetric affine connection* ∇ on any manifold M:

$$[X, Y] = \nabla_X Y - \nabla_Y X.$$

In case of a Riemannian manifold M, the connection ∇ is also compatible with the Riemannian metrics g on M and is called the *Levi–Civita connection* on TM.

For a function $f \in C^k(M, \mathbb{R})$ and a vector a vector–field $X \in \mathcal{X}^k(M)$ we always have the Lie derivative (2.4.1)

$$\mathcal{L}_X f = \nabla_X f = df(X).$$

But there is no natural definition for $\nabla_X Y$, where $Y \in \mathcal{X}^k(M)$, unless one also has a Riemannian metric. Given the tangent field $\dot{\gamma}$, the acceleration can then be computed by using a Leibniz rule on the r.h.s, if we can make sense of the derivative of ∂_{x^i} in the direction of $\dot{\gamma}$. This is exactly what the *covariant derivative* $\nabla_X Y$ does. If $Y \in T_m M$ then we can simply write $Y = a^i \partial_{x^i}$, and therefore

$$\nabla_X Y = \mathcal{L}_X a^i \partial_{x^i}. \tag{2.89}$$

Since there are several ways of choosing these coordinates, one must check that the definition does not depend on the choice. Note that for two vector–fields we define $(\nabla_Y X)(m) = \nabla_{Y(m)} X$. In the end we get a *connection*

$$\nabla : \mathcal{X}^k(M) \times \mathcal{X}^k(M) \to \mathcal{X}^k(M),$$

which satisfies (for all $f \in C^k(M, \mathbb{R})$ and $X, Y, Z \in \mathcal{X}^k(M)$):

1. $Y \to \nabla_Y X$ is tensorial, i.e., linear and $\nabla_{fY} X = f \nabla_Y X$.
2. $X \to \nabla_Y X$ is linear.
3. $\nabla_X (fY) = (\nabla_X f) Y(m) + f(m) \nabla_X Y$.
4. $\nabla_X Y - \nabla_Y X = [X, Y]$.
5. $\mathcal{L}_X g(Z, Y) = g(\nabla_X Z, Y) + g(Z, \nabla_X Y)$.

A semicolon is commonly used to denote covariant differentiation with respect to a natural basis vector. If $X = \partial_{x^i}$, then the components of $\nabla_X Y$ in (2.89) are denoted

$$Y^k_{;i} = \partial_{x^i} Y^k + \Gamma^k_{ij} Y^j, \tag{2.90}$$

where Γ^k_{ij} are *Christoffel symbols* defined in (2.91) below. Similar relations hold for higher–order tensor–fields (with as many terms with Christoffel symbols as is the tensor valence).

Therefore, no matter which coordinates we use, we can now define the acceleration of a curve in the following way:

$$\begin{aligned}
\gamma(t) &= (\gamma^1(t), \dots, \gamma^n(t)), \\
\dot{\gamma}(t) &= \dot{\gamma}^i(t) \partial_{x^i}, \\
\ddot{\gamma}(t) &= \ddot{\gamma}^i(t) \partial_{x^i} + \dot{\gamma}^i(t) \nabla_{\dot{\gamma}(t)} \partial_{x^i}.
\end{aligned}$$

We call γ a *geodesic* if $\gamma(t) = 0$. This is a second order nonlinear ODE in a fixed coordinate system $(x^1, \dots, x^n)$ at the specified point $m \in M$. Thus we see that given any tangent vector $X \in T_m M$, there is a unique geodesic $\gamma_X(t)$ with $\dot\gamma_X(0) = X$. If the manifold M is closed, the geodesic must exist for all time, but in case the manifold M is open this might not be so. To see this, simply take as M any open subset of Euclidean space with the induced metric.

Given an arbitrary vector–field $Y(t)$ along γ, i.e., $Y(t) \in T_{\gamma(t)}M$ for all t, we can also define the derivative $\dot Y \equiv \frac{dY}{dt}$ in the direction of $\dot\gamma$ by writing

$$\begin{aligned} Y(t) &= a^i(t)\partial_{x^i}, \\ \dot Y(t) &= \dot a^i(t)\partial_{x^i} + a^i(t)\nabla_{\dot\gamma(t)}\partial_{x^i}. \end{aligned}$$

Here the derivative of the tangent field $\dot\gamma$ is simply the acceleration γ. The field Y is said to be *parallel* iff $\dot Y = 0$. The equation for a field to be parallel is a first order linear ODE, so we see that for any $X \in T_{\gamma(t_0)}M$ there is a unique parallel field $Y(t)$ defined on the entire domain of γ with the property that $Y(t_0) = X$. Given two such parallel fields $Y, Z \in \mathcal{X}^k(M)$, we have that

$$\dot g(Y, Z) = D_{\dot\gamma} g(Y, Z) = g(\dot Y, Z) + g(Y, \dot Z) = 0.$$

Thus X and Y are both of constant length and form constant angles along γ. Hence, 'parallel translation' along a curve defines an orthogonal transformation between the tangent spaces to the manifold along the curve. However, in contrast to Euclidean space, this parallel translation will depend on the choice of curve.

An infinitesimal distance between the two nearby local points m and n on M is defined by an *arc–element*

$$ds^2 = g_{ij}\, dx^i dx^j,$$

and realized by the curves $x^i(s)$ of shortest distance, called *geodesics*, addressed by the *Hilbert 4th problem*. In local coordinates $(x^1(s), ..., x^n(s))$ at a point $m \in M$, the geodesic defining equation is a second order ODE,

$$\ddot x^i + \Gamma^i_{jk}\, \dot x^j\, \dot x^k = 0,$$

where the overdot denotes the derivative with respect to the affine parameter s, $\dot x^i(s) = \frac{d}{ds}x^i(s)$ is the tangent vector to the base geodesic, while the *Christoffel symbols* $\Gamma^i_{jk} = \Gamma^i_{jk}(m)$ (see Appendix) of the *affine connection* (Levi–Civita) ∇ at the point $m \in M$ are defined as

$$\begin{aligned} \Gamma^k_{ij} &= g^{kl}\Gamma_{ijl}, \qquad \text{with} \qquad g^{ij} = (g_{ij})^{-1} \qquad \text{and} \\ \Gamma_{ijk} &= \frac{1}{2}(\partial_{x^i} g_{jk} - \partial_{x^j} g_{ki} + \partial_{x^k} g_{ij}). \end{aligned} \tag{2.91}$$

The *torsion* tensor–field T of the connection ∇ is the function $T : \mathcal{X}^k(M) \times \mathcal{X}^k(M) \to \mathcal{X}^k(M)$ given by

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y].$$

From the skew symmetry ($[X,Y] = -[Y,X]$) of the Lie bracket, follows the skew symmetry ($T(X,Y) = -T(Y,X)$) of the torsion tensor. The mapping T is said to be $f-$bilinear since it is linear in both arguments and also satisfies $T(fX,Y) = fT(X,Y)$ for smooth functions f. Since $[\partial_{x^i}, \partial_{x^j}] = 0$ for all $1 \leq i,j \leq n$, it follows that

$$T(\partial_{x^i}, \partial_{x^j}) = (\Gamma_{ij}^k - \Gamma_{ji}^k)\partial_{x^k}. \tag{2.92}$$

Consequently, torsion T is a $(1,2)$ tensor–field, locally given by

$$T = T_{i\,j}^{k}\, dx^i \otimes \partial_{x^k} \otimes dx^j, \tag{2.93}$$

where the torsion components $T_{i\,j}^{k}$ are given by

$$T_{i\,j}^{k} = \Gamma_{ij}^k - \Gamma_{ji}^k. \tag{2.94}$$

Therefore, the torsion tensor provides a measure of the nonsymmetry of the connection coefficients. Hence, $T = 0$ if and only if these coefficients are symmetric in their subscripts. A connection ∇ with $T = 0$ is said to be *torsion free* or *symmetric*.

The connection also enables us to define many other classical concepts from calculus in the setting of Riemannian manifolds. Suppose we have a function $f \in C^k(M,\mathbb{R})$. If the manifold is not equipped with a Riemannian metric, then we have the differential of f defined by $df(X) = \mathcal{L}_X f$, which is a $1-$form. The dual concept, the *gradient* of f, is supposed to be a vector–field. But we need a metric g to define it. Namely, ∇f is defined by the relationship

$$g(\nabla f, X) = df(X).$$

Having defined the gradient of a function on a Riemannian manifold, we can then use the connection to define the *Hessian* as the linear map

$$\nabla^2 f : TM \to TM, \qquad \nabla^2 f(X) = \nabla_X \nabla f.$$

The corresponding bilinear map is then defined as

$$\nabla^2 f(X,Y) = g(\nabla^2 f(X), Y).$$

One easily checks that this is a symmetric bilinear form. The *Laplacian* of f, Δf, is now defined as the trace of the Hessian

$$\Delta f = \mathrm{Tr}(\nabla^2 f(X)) = \mathrm{Tr}(\nabla_X \nabla f),$$

which is a linear map. It is also called the *Laplace–Beltrami operator*, since Beltrami first considered this operator on Riemannian manifolds.

Riemannian metric has the following mechanical interpretation. Let M be a closed Riemannian manifold with the *mechanical metric* $g = g_{ij}v^i v^j \equiv \langle v, v\rangle$, with $v^i = \dot{x}^i$. Consider the *Lagrangian function*

$$L : TM \to \mathbb{R}, \qquad (x, v) \mapsto \frac{1}{2}\langle v, v\rangle - U(x) \tag{2.95}$$

where $U(x)$ is a smooth function on M called the *potential*. On a fixed level of energy E, bigger than the maximum of U, the Lagrangian flow generated by (2.95) is conjugate to the geodesic flow with metric $\bar{g} = 2(e - U(x))\langle v, v\rangle$. Moreover, the reduced action of the Lagrangian is the distance for $g = \langle v, v\rangle$ [Arn89, AMR88]. Both of these statements are known as the *Maupertius action principle* (see subsection 3.3.5 below).

Geodesics on M

For a $C^k, k \geq 2$ curve $\gamma : I \to M$, we define its *length* on I as

$$L(\gamma, I) = \int_I |\dot{\gamma}|\, dt = \int_I \sqrt{g(\dot{\gamma}, \dot{\gamma})} dt.$$

This length is independent of our parametrization of the curve γ. Thus the curve γ can be reparameterized, in such a way that it has unit velocity. The *distance between two points* m_1 and m_2 on M, $d(m_1, m_2)$, can now be defined as the infimum of the lengths of all curves from m_1 to m_2, i.e.,

$$L(\gamma, I) \to \min.$$

This means that the distance measures the shortest way one can travel from m_1 to m_2.

If we take a variation $V(s, t) : (-\varepsilon, \varepsilon) \times [0, \ell] \to M$ of a smooth curve $\gamma(t) = V(0, t)$ parameterized by arc–length L and of length ℓ, then the first derivative of the arc–length function

$$L(s) = \int_0^\ell |\dot{V}|\, dt, \qquad \text{is given by}$$

$$\frac{dL(0)}{ds} \equiv \dot{L}(0) = g(\dot{\gamma}, X)|_0^\ell - \int_0^\ell g(\gamma, X)\, dt, \tag{2.96}$$

where $X(t) = \frac{\partial V}{\partial s}(0, t)$ is the so–called *variation vector–field*. Equation (2.96) is called the *first variation formula*. Given any vector–field X along γ, one can produce a variation whose variational field is X. If the variation fixes the endpoints, $X(a) = X(b) = 0$, then the second term in the formula drops out, and we note that the length of γ can always be decreased as long as the acceleration of γ is not everywhere zero. Thus the *Euler–Lagrange equation* for the arc–length functional is simply the equation for a curve to be a *geodesic*.

In local coordinates $x^i \in U$, where U is an open subset in the Riemannian manifold M, the geodesics are defined by the *geodesic equation* (see Appendix)

$$\ddot{x}^i + \Gamma^i_{jk}\dot{x}^j\dot{x}^k = 0, \tag{2.97}$$

where overdot means derivative upon the line parameter s, while Γ^i_{jk} are Christoffel symbols of the affine Levi–Civita connection ∇ on M. From (6.18) it follows that the linear *connection homotopy*,

$$\bar{\Gamma}^i_{jk} = s\Gamma^i_{jk} + (1-s)\Gamma^i_{jk}, \qquad (0 \leq s \leq 1),$$

determines the same geodesics as the original Γ^i_{jk}.

Riemannian Curvature on M

The *Riemann curvature tensor* is a rather ominous tensor of type $(1,3)$; i.e., it has three vector variables and its value is a vector as well. It is defined through the Lie bracket (2.4.1) as

$$R\,(X,Y)\,Z = \left(\nabla_{[X,Y]} - [\nabla_X, \nabla_Y]\right) Z = \nabla_{[X,Y]}Z - \nabla_X\nabla_Y Z + \nabla_Y\nabla_X Z.$$

This turns out to be a vector valued $(1,3)-$tensor–field in the three variables $X, Y, Z \in \mathcal{X}^k(M)$. We can then create a $(0,4)-$tensor,

$$R\,(X,Y,Z,W) = g\left(\nabla_{[X,Y]}Z - \nabla_X\nabla_Y Z + \nabla_Y\nabla_X Z, W\right).$$

Clearly this tensor is skew–symmetric in X and Y, and also in Z and $W \in \mathcal{X}^k(M)$. This was already known to Riemann, but there are some further, more subtle properties that were discovered a little later by Bianchi. The *Bianchi symmetry condition* reads

$$R(X,Y,Z,W) = R(Z,W,X,Y).$$

Thus the Riemannian curvature tensor is a symmetric *curvature operator*

$$\mathfrak{R} : \Lambda^2 TM \to \Lambda^2 TM.$$

The *Ricci tensor* is the $(1,1)-$ or $(0,2)-$tensor defined by

$$\mathrm{Ric}(X) = R(\partial_{x^i}, X)\partial_{x^i}, \qquad \mathrm{Ric}(X,Y) = g(R(\partial_{x^i}, X)\partial_{x^i}, Y),$$

for any orthonormal basis (∂_{x^i}). In other words, the Ricci curvature is simply a trace of the curvature tensor. Similarly one can define the *scalar curvature* as the trace

$$\mathrm{scal}(m) = \mathrm{Tr}\,(\mathrm{Ric}) = \mathrm{Ric}(\partial_{x^i}, \partial_{x^i}).$$

When the Riemannian manifold has dimension 2, all of these curvatures are essentially the same. Since $\dim \Lambda^2 TM = 1$ and is spanned by $X \wedge Y$ where

$X, Y \in \mathcal{X}^k(M)$ form an orthonormal basis for T_mM, we see that the curvature tensor depends only on the scalar value

$$K(m) = R(X, Y, X, Y),$$

which also turns out to be the *Gaussian curvature*. The Ricci tensor is a *homothety*

$$\mathrm{Ric}(X) = K(m)X, \qquad \mathrm{Ric}(Y) = K(m)Y,$$

and the scalar curvature is twice the Gauss curvature. In dimension 3 there are also some redundancies as $\dim TM = \dim \Lambda^2 TM = 3$. In particular, the Ricci tensor and the curvature tensor contain the same amount of information.

The *sectional curvature* is a kind of generalization of the Gauss curvature whose importance Riemann was already aware of. Given a 2−plane $\pi \subset T_mM$ spanned by an orthonormal basis $X, Y \in \mathcal{X}^k(M)$ it is defined as

$$\sec(\pi) = R(X, Y, X, Y).$$

The remarkable observation by Riemann was that the *curvature operator is a homothety*, i.e., looks like $\mathfrak{R} = kI$ on $\Lambda^2 T_mM$ iff all sectional curvatures of planes in T_mM are equal to k. This result is not completely trivial, as the sectional curvature is not the entire quadratic form associated to the symmetric operator $\mathfrak{R}$. In fact, it is not true that $\sec \geq 0$ implies that the curvature operator is nonnegative in the sense that all its eigenvalues are nonnegative. What Riemann did was to show that our special coordinates $(x^1, \ldots, x^n)$ at m can be chosen to be *normal* at m, i.e., satisfy the condition

$$x^i = \delta^i_j x^j, \qquad (\delta^i_j x^j = g_{ij})$$

on a neighborhood of m. One can easily show that such coordinates are actually exponential coordinates together with a choice of an orthonormal basis for T_mM so as to identify T_mM with $\mathbb{R}^n$. In these coordinates one can then expand the metric as follows:

$$g_{ij} = \delta_{ij} - \frac{1}{3} R_{ikjl} x^k x^l + O\left(r^3\right).$$

Now the equations $x^i = g_{ij} x^j$ evidently give conditions on the curvatures R_{ijkl} at m.

If $\Gamma^i_{jk}(m) = 0$, the manifold M is flat at the point m. This means that the $(1, 3)$ curvature tensor, defined locally at $m \in M$ as

$$R^l_{ijk} = \partial_{x^j} \Gamma^l_{ik} - \partial_{x^k} \Gamma^l_{ij} + \Gamma^l_{rj} \Gamma^r_{ik} - \Gamma^l_{rk} \Gamma^r_{ij},$$

also vanishes at that point, i.e., $R^l_{ijk}(m) = 0$.

Now, the rate of change of a vector–field A^k on the manifold M along the curve $x^i(s)$ is properly defined by the *absolute covariant derivative*

$$\frac{D}{ds}A^k = \dot{x}^i \nabla_i A^k = \dot{x}^i \left(\partial_{x^i} A^k + \Gamma^k_{ij} A^j\right) = \dot{A}^k + \Gamma^k_{ij} \dot{x}^i A^j.$$

By applying this result to itself, we can get an expression for the second covariant derivative of the vector–field A^k along the curve $x^i(s)$:

$$\frac{D^2}{ds^2}A^k = \frac{d}{ds}\left(\dot{A}^k + \Gamma^k_{ij} \dot{x}^i A^j\right) + \Gamma^k_{ij} \dot{x}^i(\dot{A}^j + \Gamma^j_{mn} \dot{x}^m A^n).$$

In the local coordinates $(x^1(s), ..., x^n(s))$ at a point $m \in M$, if $\delta x^i = \delta x^i(s)$ denotes the *geodesic deviation*, i.e., the infinitesimal vector describing perpendicular separation between the two neighboring geodesics, passing through two neighboring points $m, n \in M$, then the *Jacobi equation of geodesic deviation* on the manifold M holds:

$$\frac{D^2 \delta x^i}{ds^2} + R^i_{jkl} \dot{x}^j \delta x^k \dot{x}^l = 0. \qquad (2.98)$$

This equation describes the *relative acceleration* between two infinitesimally close facial geodesics, which is proportional to the facial curvature (measured by the Riemann tensor R^i_{jkl} at a point $m \in M$), and to the geodesic deviation δx^i. Solutions of equation (6.19) are called *Jacobi fields*.

In particular, if the manifold M is a 2D–surface in $\mathbb{R}^3$, the Riemann curvature tensor simplifies into

$$R^i_{jmn} = \frac{1}{2} R\, g^{ik}(g_{km}\, g_{jn} - g_{kn}\, g_{jm}),$$

where R denotes the *scalar curvature*. Consequently the equation of geodesic deviation (6.19) also simplifies into

$$\frac{D^2}{ds^2}\delta x^i + \frac{R}{2}\delta x^i - \frac{R}{2}\dot{x}^i(g_{jk}\, \dot{x}^j\, \delta x^k) = 0. \qquad (2.99)$$

This simplifies even more if we work in a locally Cartesian coordinate system; in this case the covariant derivative $\frac{D^2}{Ds^2}$ reduces to an ordinary derivative $\frac{d^2}{ds^2}$ and the metric tensor g_{ij} reduces to identity matrix I_{ij}, so our 2D equation of geodesic deviation (6.20) reduces into a simple second order ODE in just two coordinates x^i $(i = 1, 2)$

$$\ddot{x}^i + \frac{R}{2}\delta x^i - \frac{R}{2}\dot{x}^i(I_{jk}\, \dot{x}^j\, \delta x^k) = 0.$$

2.5.2 Global Riemannian Geometry on M

The Second Variation Formula

Cartan also establishes another important property of manifolds with nonpositive curvature. First he observes that all spaces of constant zero curvature have torsion–free fundamental groups. This is because any isometry of finite order on Euclidean space must have a fixed point (the center of mass of any orbit is necessarily a fixed point). Then he notices that one can geometrically describe the L^∞ center of mass of finitely many points $\{m_1, \dots, m_k\}$ in Euclidean space as the unique minimum for the strictly convex function

$$x \to \max_{i=1,\cdots,k} \frac{1}{2}\left\{(d\,(m_i, x))^2\right\}.$$

In other words, the center of mass is the center of the ball of smallest radius containing $\{m_1, \dots, m_k\}$. Now Cartan's observation from above was that the exponential map is expanding and globally distance nondecreasing as a map:

$$(T_m M,\ \text{Euclidean metric}) \to (T_m M,\ \text{with pull–back metric})\,.$$

Thus distance functions are convex in nonpositive curvature as well as in Euclidean space. Hence the above argument can in fact be used to conclude that any Riemannian manifold of nonpositive curvature must also have torsion free fundamental group.

Now, let us set up the *second variation formula* and explain how it is used. We have already seen the first variation formula and how it can be used to characterize geodesics. Now suppose that we have a unit speed geodesic $\gamma\,(t)$ parameterized on $[0, \ell]$ and consider a variation $V\,(s,t)$, where $V\,(0,t) = \gamma\,(t)$. Synge then shows that ($\ddot{L} \equiv \frac{d^2 L}{ds^2}$)

$$\ddot{L}(0) = \int_0^\ell \{g(\dot{X}, \dot{X}) - (g(\dot{X}, \dot{\gamma}))^2 - g(R(X, \dot{\gamma})X, \dot{\gamma})\}dt + g(\dot{\gamma}, A)|_0^\ell\,,$$

where $X\,(t) = \frac{\partial V}{\partial s}\,(0,t)$ is the variational vector–field, $\dot{X} = \nabla_{\dot{\gamma}} X$, and $A\,(t) = \nabla_{\frac{\partial V}{\partial s}} X$. In the special case where the variation fixes the endpoints, i.e., $s \to V\,(s,a)$ and $s \to V\,(s,b)$ are constant, the term with A in it falls out. We can also assume that the variation is perpendicular to the geodesic and then drop the term $g\left(\dot{X}, \dot{\gamma}\right)$. Thus, we arrive at the following simple form:

$$\ddot{L}(0) = \int_0^\ell \{g(\dot{X}, \dot{X}) - g\,(R\,(X, \dot{\gamma})\,X, \dot{\gamma})\}dt = \int_0^\ell \{|\dot{X}|^2 - \sec(\dot{\gamma}, X)\,|X|^2\}dt.$$

Therefore, if the sectional curvature is nonpositive, we immediately observe that any geodesic locally minimizes length (that is, among close–by curves), even if it does not minimize globally (for instance γ could be a closed geodesic). On the other hand, in positive curvature we can see that if a geodesic is too long, then it cannot minimize even locally. The motivation for this result comes from the unit sphere, where we can consider geodesics of length $> \pi$. Globally, we of course know that it would be shorter to go in the opposite direction. However, if we consider a variation of γ where the variational field looks like $X = \sin\left(t \cdot \frac{\pi}{\ell}\right) E$ and E is a unit length parallel field along γ which is also perpendicular to γ, then we get

$$\begin{aligned}
\ddot{L}(0) &= \int_0^\ell \left\{ \left|\dot{X}\right|^2 - \sec(\dot{\gamma}, X) \left|X\right|^2 \right\} dt \\
&= \int_0^\ell \left\{ \left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \sec(\dot{\gamma}, X) \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt \\
&= \int_0^\ell \left(\left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right) dt = -\frac{1}{2\ell}\left(\ell^2 - \pi^2\right),
\end{aligned}$$

which is negative if the length ℓ of the geodesic is greater than π. Therefore, the variation gives a family of curves that are both close to and shorter than γ. In the general case, we can then observe that if $\sec \geq 1$, then for the same type of variation we get

$$\ddot{L}(0) \leq -\frac{1}{2\ell}\left(\ell^2 - \pi^2\right).$$

Thus we can conclude that, if the space is complete, then the diameter must be $\leq \pi$ because in this case any two points are joined by a segment, which cannot minimize if it has length $> \pi$. With some minor modifications one can now conclude that any complete Riemannian manifold (M, g) with $\sec \geq k^2 > 0$ must satisfy $\mathrm{diam}(M, g) \leq \pi \cdot k^{-1}$. In particular, M must be compact. Since the universal covering of M satisfies the same curvature hypothesis, the conclusion must also hold for this space; hence M must have compact universal covering space and finite fundamental group.

In odd dimensions all spaces of constant positive curvature must be orientable, as orientation reversing orthogonal transformation on odd–dimensional spheres have fixed points. This can now be generalized to manifolds of varying positive curvature. Synge did it in the following way: Suppose M is not simply connected (or not orientable), and use this to find a shortest closed geodesic in a free homotopy class of curves (that reverses orientation). Now consider parallel translation around this geodesic. As the tangent field to the geodesic is itself a parallel field, we see that parallel translation preserves the orthogonal complement to the geodesic. This complement is now odd dimensional (even dimensional), and by assumption parallel translation preserves (reverses) the orientation; thus it must have a fixed point. In other words, there must exist a closed parallel field X perpendicular to the closed geodesic γ. We can now use the above second variation formula

$$\ddot{L}(0) = \int_0^\ell \{|\dot{X}|^2 - |X|^2 \sec(\dot{\gamma}, X)\} dt + g(\dot{\gamma}, A)|_0^\ell = -\int_0^\ell |X|^2 \sec(\dot{\gamma}, X)\, dt.$$

Here the boundary term drops out because the variation closes up at the endpoints, and $\dot{X} = 0$ since we used a parallel field. In case the sectional curvature is always positive we then see that the above quantity is negative.

But this means that the closed geodesic has nearby closed curves which are shorter. This is, however, in contradiction with the fact that the geodesic was constructed as a length minimizing curve in a free homotopy class.

In 1941 Myers generalized the diameter bound to the situation where one only has a lower bound for the Ricci curvature. The idea is simply that $\mathrm{Ric}(\dot{\gamma},\dot{\gamma}) = \sum_{i=1}^{n-1} \sec(E_i,\dot{\gamma})$ for any set of vector–fields E_i along γ such that $\dot{\gamma}, E_1, \ldots, E_{n-1}$ forms an orthonormal frame. Now assume that the fields are parallel and consider the $n-1$ variations coming from the variational vector–fields $\sin\left(t \cdot \frac{\pi}{\ell}\right) E_i$. Adding up the contributions from the variational formula applied to these fields then yields

$$\begin{aligned} \sum_{i=1}^{n-1} \ddot{L}(0) &= \sum_{i=1}^{n-1} \int_0^{\ell} \left\{ \left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \sec(\dot{\gamma}, E_i) \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt \\ &= \int_0^{\ell} \left\{ (n-1) \left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - \mathrm{Ric}(\dot{\gamma},\dot{\gamma}) \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt. \end{aligned}$$

Therefore, if $\mathrm{Ric}(\dot{\gamma},\dot{\gamma}) \geq (n-1)\,k^2$ (this is the Ricci curvature of S_k^n), then

$$\begin{aligned} \sum_{i=1}^{n-1} \ddot{L}(0) &\leq (n-1) \int_0^{\ell} \left\{ \left(\frac{\pi}{\ell}\right)^2 \cdot \cos^2\left(t \cdot \frac{\pi}{\ell}\right) - k^2 \sin^2\left(t \cdot \frac{\pi}{\ell}\right) \right\} dt \\ &= -\,(n-1)\, \frac{1}{2\ell} \left(\ell^2 k^2 - \pi^2\right), \end{aligned}$$

which is negative when $\ell > \pi \cdot k^{-1}$ (the diameter of S_k^n). Thus at least one of the contributions $\frac{d^2 L_i}{ds^2}(0)$ must be negative as well, implying that the geodesic cannot be a segment in this situation.

Gauss–Bonnet Formula

In 1926 Hopf proved that in fact there is a Gauss–Bonnet formula for all even–dimensional hypersurfaces $H^{2n} \subset \mathbb{R}^{2n+1}$. The idea is simply that the determinant of the differential of the Gauss map $G : H^{2n} \to S^{2n}$ is the Gaussian curvature of the hypersurface. Moreover, this is an intrinsically computable quantity. If we integrate this over the hypersurface, we get,

$$\frac{1}{\mathrm{vol}\, S^{2n}} \int_H \det(DG) = \deg(G),$$

where $\deg(G)$ is the *Brouwer degree* of the *Gauss map*. Note that this can also be done for odd–dimensional surfaces, in particular curves, but in this case the degree of the Gauss map will depend on the embedding or immersion of the hypersurface. Instead one gets the so–called winding number. Hopf then showed, as Dyck had earlier done for surfaces, that $\deg(G)$ is always half the *Euler characteristic* of H, thus yielding

$$\frac{2}{\operatorname{vol} S^{2n}} \int_H \det(DG) = \chi(H). \tag{2.100}$$

Since the l.h.s of this formula is in fact intrinsic, it is natural to conjecture that such a formula should hold for all manifolds.

Ricci Flow on M

Ricci flow, or the *parabolic Einstein equation*, was introduced by R. Hamilton in 1982 [Ham82] in the form

$$\partial_t g_{ij} = -2R_{ij}. \tag{2.101}$$

Now, because of the minus sign in the front of the Ricci tensor R_{ij} in this equation, the solution metric g_{ij} to the Ricci flow shrinks in positive Ricci curvature direction while it expands in the negative Ricci curvature direction. For example, on the 2−sphere S^2, any metric of positive Gaussian curvature will shrink to a point in finite time. Since the Ricci flow (2.101) does not preserve volume in general, one often considers the *normalized* Ricci flow defined by

$$\partial_t g_{ij} = -2R_{ij} + \frac{2}{n} r g_{ij}, \tag{2.102}$$

where $r = \int R dV \big/ \int dV$ is the average scalar curvature. Under this normalized flow, which is equivalent to the (unnormalized) Ricci flow (2.101) by reparameterizing in time t and scaling the metric in space by a function of t, the volume of the solution metric is constant in time. Also that Einstein metrics (i.e., $R_{ij} = cg_{ij}$) are fixed points of (2.102).

Hamilton [Ham82] showed that on a closed Riemannian 3−manifold M^3 with initial metric of positive Ricci curvature, the solution $g(t)$ to the normalized Ricci flow (2.102) exists for all time and the metrics $g(t)$ converge exponentially fast, as time t tends to the infinity, to a constant positive sectional curvature metric g_∞ on M^3.

Since the Ricci flow lies in the realm of parabolic partial differential equations, where the prototype is the heat equation, here is a brief review of the *heat equation* [CC99].

Let (M^n, g) be a Riemannian manifold. Given a C^2 function $u : M \to \mathbb{R}$, its Laplacian is defined in local coordinates $\{x^i\}$ to be

$$\Delta u = \operatorname{Tr}\left(\nabla^2 u\right) = g^{ij} \nabla_i \nabla_j u,$$

where $\nabla_i = \nabla_{\partial_{x^i}}$ is its associated covariant derivative (Levi–Civita connection). We say that a C^2 function $u : M^n \times [0, T) \to \mathbb{R}$, where $T \in (0, \infty]$, is a solution to the heat equation if

$$\partial_t u = \Delta u.$$

One of the most important properties satisfied by the heat equation is the maximum principle, which says that for any smooth solution to the heat equation, whatever pointwise bounds hold at $t = 0$ also hold for $t > 0$. Let $u : M^n \times [0, T) \to \mathbb{R}$ be a C^2 solution to the heat equation on a complete Riemannian manifold. If $C_1 \leq u\,(x, 0) \leq C_2$ for all $x \in M$, for some constants $C_1, C_2 \in \mathbb{R}$, then $C_1 \leq u\,(x, t) \leq C_2$ for all $x \in M$ and $t \in [0, T)$ [CC99].

Now, given a differentiable manifold M, a one–parameter family of metrics $g\,(t)$, where $t \in [0, T)$ for some $T > 0$, is a solution to the Ricci flow if (2.101) is valid at all $x \in M$ and $t \in [0, T)$. The minus sign in the equation (2.101) makes the Ricci flow a *forward* heat equation [CC99] (with the normalization factor 2).

In local geodesic coordinates $\{x^i\}$, we have [CC99]

$$g_{ij}(x) = \delta_{ij} - \frac{1}{3} R_{ipjq} x^p x^q + O\left(|x|^3\right), \qquad \text{therefore,} \qquad \Delta g_{ij}\,(0) = -\frac{1}{3} R_{ij},$$

where Δ is the standard Euclidean Laplacian. Hence the Ricci flow is like the heat equation for a Riemannian metric

$$\partial_t g_{ij} = 6 \Delta g_{ij}.$$

The practical study of the Ricci flow is made possible by the following short–time existence result: Given any smooth compact Riemannian manifold (M, g_o), there exists a unique smooth solution $g(t)$ to the Ricci flow defined on some time interval $t \in [0, \epsilon)$ such that $g(0) = g_o$ [CC99].

Now, given that short–time existence holds for any smooth initial metric, one of the main problems concerning the Ricci flow is to determine under what conditions the solution to the normalized equation exists for all time and converges to a constant curvature metric. Results in this direction have been established under various curvature assumptions, most of them being some sort of positive curvature. Since the Ricci flow (2.101) does not preserve volume in general, one often considers, as we mentioned in the Introduction, the normalized Ricci flow (2.102). Under this flow, the volume of the solution $g(t)$ is independent of time.

To study the long–time existence of the normalized Ricci flow, it is important to know what kind of curvature conditions are preserved under the equation. In general, the Ricci flow tends to preserve some kind of positivity of curvatures. For example, positive scalar curvature is preserved in all dimensions. This follows from applying the maximum principle to the evolution equation for scalar curvature R, which is

$$\partial_t R = \Delta R + 2 \left| R_{ij} \right|^2.$$

In dimension 3, positive Ricci curvature is preserved under the Ricci flow. This is a special feature of dimension 3 and is related to the fact that the Riemann curvature tensor may be recovered algebraically from the Ricci tensor and the metric in dimension 3. Positivity of sectional curvature is not preserved in general. However, the stronger condition of positive curvature operator is preserved under the Ricci flow. Recall that the Riemann curvature tensor may be considered as a self–adjoint map $Rm : \wedge^2 M \to \wedge^2 M$. We say that a metric g has positive (non–negative) curvature operator if the eigenvalues of Rm are positive (non–negative). We remark that positivity of curvature operator implies the positivity of the sectional curvature (and in dimension 3, the two conditions are equivalent).

Although the condition of positive scalar curvature is preserved in all dimensions, no convergence results are known for metrics satisfying this condition except in dimension 2.

Structure Equations on M

Let $\{X_a\}_{a=1}^m$, $\{Y_i\}_{i=1}^n$ be local orthonormal framings on M, N respectively and $\{e_i\}_{i=1}^n$ be the induced framing on E defined by $e_i = Y_i \circ \phi$, then there exist smooth local coframings $\{\omega_a\}_{a=1}^m$, $\{\eta_i\}_{i=1}^n$ and $\{\phi^*\eta_i\}_{i=1}^n$ on TM, TN and E respectively such that (locally)

$$g = \sum_{a=1}^m \omega_a^2 \qquad \text{and} \qquad h = \sum_{i=1}^n \eta_i^2.$$

The corresponding first *structure equations* are [Mus99]:

$$\begin{aligned} d\omega_a &= \omega_b \wedge \omega_{ba}, & \omega_{ab} &= -\omega_{ba}, \\ d\eta_i &= \eta_j \wedge \eta_{ji}, & \eta_{ij} &= -\eta_{ji}, \\ d(\phi^*\eta_i) &= \phi^*\eta_j \wedge \phi^*\eta_{ji}, & \phi^*\eta_{ij} &= -\phi^*\eta_{ji}, \end{aligned}$$

where the unique 1–forms ω_{ab}, η_{ij}, $\phi^*\eta_{ij}$ are the respective connection forms. The second structure equations are

$$\begin{aligned} d\omega_{ab} &= \omega_{ac} \wedge \omega_{cb} + \Omega_{ab}^M, \qquad d\eta_{ij} = \eta_{ik} \wedge \eta_{kj} + \Omega_{ij}^N, \\ d(\phi^*\eta_{ij}) &= \phi^*\eta_{ik} \wedge \phi^*\eta_{kj} + \phi^*\Omega_{ij}^N, \end{aligned}$$

where the curvature 2–forms are given by

$$\Omega_{ab}^M = -\frac{1}{2} R_{abcd}^M \omega_c \wedge \omega_d \qquad \text{and} \qquad \Omega_{ij}^N = -\frac{1}{2} R_{ijkl}^N \eta_k \wedge \eta_l.$$

The pull back map ϕ^* and the push forward map ϕ_* can be written as [Mus99]

$$\phi^*\eta_i = f_{ia}\omega_a$$

for unique functions f_{ia} on $U \subset M$, so that

$$\phi_* = e_i \otimes \phi^*\eta_i = f_{ia} e_i \otimes \omega_a.$$

Note that ϕ_* is a section of the vector bundle $\phi^{-1}TN \otimes T^*M$.

The covariant differential operators are represented as

$$\nabla^M X_a = \omega_{ab} \otimes X_b, \qquad \nabla^N Y_i = \eta_{ij} \otimes Y_j, \qquad \nabla^* \omega_a = -\omega_{ca} \otimes \omega_c,$$

where ∇^* is the dual connection on the cotangent bundle T^*M.

Furthermore, the induced connection ∇^ϕ on E is

$$\nabla^\phi e_i = \big(\eta_{ij}(Y_k) \circ \phi\big)\, e_j \otimes f_{ka}\omega_a.$$

The components of the Ricci tensor and scalar curvature are defined respectively by

$$R^M_{ab} = R^M_{acbc} \qquad \text{and} \qquad R^M = R^M_{aa}.$$

Given a function $f : M \to$, there exist unique functions $f_{cb} = f_{bc}$ such that

$$df_c - f_b\omega_{cb} = f_{cb}\omega_b\,, \tag{2.103}$$

where $f_c = df(X_c)$ for a local orthonormal frame $\{X_c\}_{c=1}^m$. To prove this we take the exterior derivative of $df = \sum_{c=1}^m f_c\omega_c$ and using structure equations, we have

$$0 = [df_c \wedge \omega_c + f_{bc}\omega_b \wedge \omega_{bc}] = [(df_c - f_b\omega_{cb}) \wedge \omega_c]\,.$$

Hence by *Cartan's lemma* (cf. [Wil93]), there exist unique functions $f_{cb} = f_{bc}$ such that

$$df_c - f_b\omega_{cb} = f_{cb}\omega_b.$$

The Laplacian of a function f on M is given by

$$\Delta f = -\operatorname{Tr}(\nabla df),$$

that is, negative of the usual Laplacian on functions.

Basics of Morse Theory

At the same time the variational formulae were discovered, a related technique, called *Morse theory*, was introduced into Riemannian geometry. This theory was developed by Morse, first for functions on manifolds in 1925, and then in 1934, for the loop space. The latter theory, as we shall see, sets up a very nice connection between the first and second variation formulae from the previous section and the topology of M. It is this relationship that we shall explore at a general level here. In section 5 we shall then see how this theory was applied in various specific settings.

If we have a *proper function* $f : M \to \mathbb{R}$, then its Hessian (as a quadratic form) is in fact well defined at its *critical points* without specifying an underlying Riemannian metric. The *nullity* of f at a *critical point* is defined as the dimension of the *kernel* of $\nabla^2 f$, while the *index* is the number of negative eigenvalues counted with multiplicity. A function is said to be a *Morse function* if the nullity at any of its critical points is zero. Note that this guarantees in particular that all critical points are isolated. The first fundamental theorem of Morse theory is that one can determine the topological structure of a manifold from a Morse function. More specifically, if one can order the critical points $x_1, \dots, x_k$ so that $f(x_1) < \cdots < f(x_k)$ and the index of x_i is denoted λ_i, then M has the structure of a CW complex with a cell of dimension λ_i for each i. Note that in case M is closed then x_1 must be a minimum and so $\lambda_1 = 0$, while x_k is a maximum and $\lambda_k = n$. The classical example of Milnor of this theorem in action is a torus in 3–space and f the height function.

We are now left with the problem of trying to find appropriate Morse functions. While there are always plenty of such functions, there does not seem to be a natural way of finding one. However, there are natural choices for Morse functions on the loop space to a Riemannian manifold. This is, somewhat inconveniently, infinite–dimensional. Still, one can develop Morse theory as above for suitable functions, and moreover *the loop space of a manifold determines the topology of the underlying manifold.*

If $m, p \in M$, then we denote by Ω_{mp} the space of all C^k paths from m to p. The first observation about this space is that

$$\pi_{i+1}(M) = \pi_i(\Omega_{mp}).$$

To see this, just fix a path from m to q and then join this path to every curve in Ω_{mp}. In this way Ω_{mp} is identified with Ω_m, *the space of loops fixed at* m. For this space the above relationship between the homotopy groups is almost self-evident.

On the space Ω_{mp} we have two naturally defined functions, the *arc–length* and *energy functionals*:

$$L(\gamma, I) = \int_I |\dot{\gamma}|\, dt, \qquad \text{and} \qquad E(\gamma, I) = \frac{1}{2}\int_I |\dot{\gamma}|^2\, dt.$$

While the energy functional is easier to work with, it is of course the arc–length functional that we are really interested in. In order to make things work out nicely for the arc–length functional, it is convenient to parameterize all curves on $[0, 1]$ and proportionally to arc–length. We shall think of Ω_{mp} as an *infinite–dimensional manifold.* For each curve $\gamma \in \Omega_{mp}$ the natural choice for the tangent space consists of the vector–fields along γ which vanish at the endpoints of γ. This is because these vector–fields are exactly the variational fields for curves through γ in Ω_{mp}, i.e., fixed endpoint variations of γ. An inner product on the tangent space is then naturally defined by

$$(X, Y) = \int_0^1 g(X, Y)\, dt.$$

Now the first variation formula for arc–length tells us that the gradient for L at γ is $-\nabla_{\dot{\gamma}}\dot{\gamma}$. Actually this cannot be quite right, as $-\nabla_{\dot{\gamma}}\dot{\gamma}$ does not vanish at the endpoints. The real gradient is gotten in the same way we find the gradient for a function on a surface in space, namely, by projecting it down into the correct tangent space. In any case we note that the critical points for L are exactly the geodesics from m to p. The second variation formula tells us that the Hessian of L at these critical points is given by

$$\nabla^2 L(X) = \ddot{X} + R(X, \dot{\gamma})\,\dot{\gamma},$$

at least for vector–fields X which are perpendicular to γ. Again we ignore the fact that we have the same trouble with endpoint conditions as above. We now need to impose the Morse condition that this Hessian is not allowed to have any kernel. The vector–fields J for which $\ddot{J} + R(J, \dot{\gamma})\,\dot{\gamma} = 0$ are called *Jacobi fields*. Thus we have to figure out whether there are any Jacobi fields which vanish at the endpoints of γ. The first observation is that Jacobi fields must always come from geodesic variations. The Jacobi fields which vanish at m can therefore be found using the exponential map $\exp_m$. If the Jacobi field also has to vanish at p, then p must be a critical value for $\exp_m$. Now Sard's theorem asserts that the set of critical values has measure zero. For given $m \in M$ it will therefore be true that the arc–length functional on Ω_{mp} is a Morse function for almost all $p \in M$. Note that it may not be possible to choose $p = m$, the simplest example being the standard sphere. We are now left with trying to decide what the *index* should be. This is of course the dimension of the largest subspace on which the Hessian is negative definite. It turns out that this index can also be computed using Jacobi fields and is in fact always finite. Thus one can compute the topology of Ω_{mp}, and hence M, by finding all the geodesics from m to p and then computing their index.

In geometric situations it is often unrealistic to suppose that one can compute the index precisely, but as we shall see it is often possible to given lower bounds for the index. As an example, note that if M is not simply connected, then Ω_{mp} is not connected. Each curve of minimal length in the path components is a geodesic from m to p which is a local minimum for the arc–length functional. Such geodesics evidently have index zero. In particular, if one can show that all geodesics, except for the minimal ones from m to p, have index > 0, then the manifold must be simply connected. We continue the exposition of Morse theory on M in section (4.2.1) below.

Basics of (Co)Bordism Theory

(Co)bordism appeared as a revival of Poincaré's unsuccessful 1895 attempts to define homology using only manifolds. Smooth manifolds (without boundary) are again considered as 'negligible' when they are *boundaries* of smooth manifolds–with–boundary. But there is a big difference, which keeps definition of 'addition' of manifolds from running into the difficulties encountered by Poincaré; it is now the disjoint union. The (unoriented) *(co)bordism relation* between two compact smooth manifolds M_1, M_2 of same dimension n simply means that their disjoint union $\partial W = M_1 \cup M_2$ is the boundary ∂W of an $(n+1)$D smooth manifold–with–boundary W. This is an *equivalence* *relation*, and the classes for that relation of nD manifolds form a *commutative group* $\mathfrak{N}_n$ in which every element has order 2. The direct sum $\mathfrak{N}_\bullet = \oplus_{n\geq 0}\mathfrak{N}_n$ is a ring for the multiplication of classes deduced from the Cartesian product of manifolds.

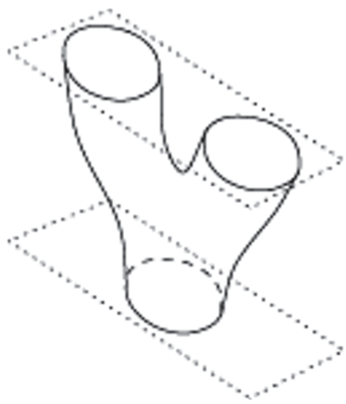
More precisely, a manifold M is said to be a (co)bordism from A to B if exists a diffeomorphism from a disjoint sum, $\varphi \in \mathrm{diff}(A^* \cup B, \partial M)$. Two (co)bordisms $M(\varphi)$ and $M'(\varphi')$ are equivalent if there is a $\Phi \in \mathrm{diff}(M, M')$ such that $\varphi' = \Phi \circ \varphi$. The equivalence class of (co)bordisms is denoted by $M(A,B) \in Cob(A,B)$ [Sto68].

Composition c_{Cob} of (co)bordisms comes from gluing of manifolds [BD95]. Let $\varphi' \in \mathrm{diff}(C^* \cup D, \partial N)$. One can glue (co)bordism M with N by identifying B with C^*, $(\varphi')^{-1} \circ \varphi \in \mathrm{diff}(B, C^*)$. We obtain the glued (co)bordism $(M \circ N)(A, D)$ and a semigroup operation,

$$c(A,B,D) : Cob(A,B) \times Cob(B,D) \longrightarrow Cob(A,D).$$

A *surgery* is an operation of cutting a manifold M and gluing to cylinders. A surgery gives new (co)bordism: from $M(A,B)$ into $N(A,B)$. The disjoint sum of $M(A,B)$ with $N(C,D)$ is a (co)bordism $(M \cup N)(A \cup C, B \cup D)$. We got a 2–graph of (co)bordism Cob with $Cob_0 = Man_d$, $Cob_1 = Man_{d+1}$, whose 2–cells from Cob_2 are surgery operations.

There is an $n-$category of (co)bordisms $\mathcal{BO}$ [Lei03] with:

- $0-$cells: $0-$manifolds, where 'manifold' means 'compact, smooth, oriented manifold'. A typical $0-$cell is $\bullet\ \bullet\ \bullet\ \bullet$.
- $1-$cells: $1-$manifolds with corners, i.e., (co)bordisms between $0-$manifolds, such as  (this being a $1-$cell from the 4–point manifold to the 2–point $0-$manifold).
- $2-$cells: $2-$manifolds with corners, such as 
- $3-$cells, $4-$cells,... are defined similarly;

- Composition is gluing of manifolds.

The (co)bordisms theme was taken a step further by by Baez and Dolan in [BD95], when when they started a programme to understand the subtle relations between certain TMFT models for manifolds of different dimensions, frequently referred to as the dimensional ladder. This programme is based on higher–dimensional algebra, a generalization of the theory of categories and functors to $n-$categories and $n-$functors. In this framework a topological quantum field theory (TMFT) becomes an $n-$functor from the $n-$category $\mathcal{BO}$ of $n-$cobordisms to the $n-$category of $n-$Hilbert spaces.

2.5.3 Complex and Kähler Manifolds

Just as a smooth manifold has enough structure to define the notion of differentiable functions, a *complex manifold* is one with enough structure to define the notion of holomorphic (or, analytic) functions $f : X \to \mathbb{C}$. Namely, if we demand that the transition functions $\phi_j \circ \phi_i^{-1}$ in the charts U_i on M (see Figure 2.4) satisfy the *Cauchy–Riemann equations*

$$\partial_x u = \partial_y v, \qquad \partial_y u = -\partial_x v,$$

then the analytic properties of f can be studied using its coordinate representative $f \circ \phi_i^{-1}$ with assurance that the conclusions drawn are patch independent. Introducing local complex coordinates in the charts U_i on M, the ϕ_i can be expressed as maps from U_i to an open set in $\mathbb{C}^{\frac{n}{2}}$, with $\phi_j \circ \phi_i^{-1}$ being a holomorphic map from $\mathbb{C}^{\frac{n}{2}}$ to $\mathbb{C}^{\frac{n}{2}}$. Clearly, n must be even for this to make sense. In local complex coordinates, we recall that a function $h : \mathbb{C}^{\frac{n}{2}} \to \mathbb{C}^{\frac{n}{2}}$ is *holomorphic* if $h(z^1, \bar{z}^1, ..., z^{\frac{n}{2}}, \bar{z}^{\frac{n}{2}})$ is actually independent of all the $\bar{z}^j$.

In a given patch on any even–dimensional manifold, we can always introduce local complex coordinates by, for instance, forming the combinations $z^j = x^j + ix^{\frac{n}{2}+j}$, where the x^j are local real coordinates on M. The real test is whether the transition functions from one patch to another — when expressed in terms of the local complex coordinates — are holomorphic maps. If they are, we say that M is a complex manifold of complex dimension $d = n/2$. The local complex coordinates with holomorphic transition functions provide M with a *complex structure* [Gre96].

Given a smooth manifold with even real dimension n, it can be a difficult question to determine whether or not a complex structure exists. On the other hand, if some differentiable manifold M does admit a complex structure, we are not able to decide whether it is unique, i.e., there may be numerous inequivalent ways of defining complex coordinates on M [Gre96].

Now, in the same way as a *homeomorphism* defines an equivalence between topological manifolds, and a *diffeomorphism* defines an equivalence between smooth manifolds, a *biholomorphism* defines an equivalence between complex manifolds. If M and N are complex manifolds, we consider them to be equivalent if there is a map $\phi : M \to N$ which in addition to being a diffeomorphism,

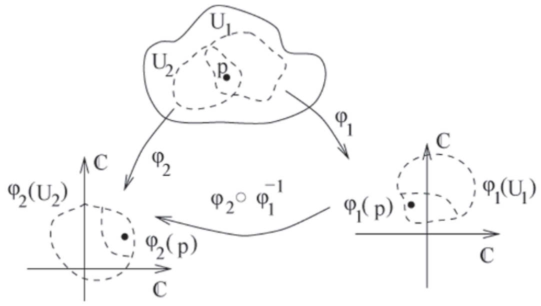


Fig. 2.4. The charts for a complex manifold M have the complex coordinates.

is also a holomorphic map. That is, when expressed in terms of the complex structures on M and N respectively, ϕ is holomorphic. It is not hard to show that this necessarily implies that ϕ^{-1} is holomorphic as well and hence ϕ is known as a biholomorphism. Such a map allows us to identify the complex structures on M and N and hence they are isomorphic as complex manifolds.

These definitions are important because there are pairs of smooth manifolds M and N which are homeomorphic but not diffeomorphic, as well as, there are complex manifolds M and N which are diffeomorphic but not biholomorphic. This means that if one simply ignored the fact that M and N admit local complex coordinates (with holomorphic transition functions), and one only worked in real coordinates, there would be no distinction between M and N. The difference between them only arises from the way in which complex coordinates have been laid down upon them.

Again, recall that a *tangent space* to a manifold M *at a point* p is the closest flat approximation to M at that point. A convenient basis for the tangent space of M at p consists of the n linearly independent partial derivatives,

$$T_pM : \{\partial_{x^1}|_p, ..., \partial_{x^n}|_p\}. \tag{2.104}$$

A vector $v \in T_pM$ can then be expressed as $v = v^\alpha \partial_{x^\alpha}|_p$.

Also, a convenient basis for the dual, *cotangent space* T_p^*M, is the basis of one–forms, which is dual to (2.104) and usually denoted by

$$T_p^*M : \{dx^1|_p, ..., dx^n|_p\}, \tag{2.105}$$

where, by definition, $dx^i : T_pM \to \mathbb{R}$ is a linear map with $dx_p^i(\partial_{x^j}|_p) = \delta_j^i$.

Now, if M is a complex manifold of complex dimension $d = n/2$, there is a notion of the *complexified tangent space* of M, denoted by $T_pM^{\mathbb{C}}$, which is the same as the real tangent space T_pM except that we allow complex coefficients to be used in the vector space manipulations. This is often denoted by writing $T_pM^{\mathbb{C}} = T_pM \otimes \mathbb{C}$. We can still take our basis to be as in (2.104) with an arbitrary vector $v \in T_pM^{\mathbb{C}}$ being expressed as $v = v^\alpha \frac{\partial}{\partial x^\alpha}|_p$, where the v^α can now be complex numbers. In fact, it is convenient to rearrange the basis vectors in

(2.104) to more directly reflect the underlying complex structure. Specifically, we take the following linear combinations of basis vectors in (2.104) to be our new basis vectors:

$$
\begin{aligned}
T_pM^{\mathbb{C}} \quad : \quad & \{(\partial_{x^1} + i\partial_{x^{d+1}})|_p, ..., \\
& (\partial_{x^d} + i\partial_{x^{2d}})|_p, (\partial_{x^1} - i\partial_{x^{d+1}})|_p, ..., (\partial_{x^d} - i\partial_{x^{2d}})|_p\}.
\end{aligned} \tag{2.106}
$$

In terms of complex coordinates we can write the basis (2.106) as

$$T_pM^{\mathbb{C}} : \{\partial_{z^1}|_p, ..., \partial_{z^d}|_p, \partial_{\bar{z}^1}|_p, ..., \partial_{\bar{z}^d}|_p\}.$$

From the point of view of real vector spaces, $\partial_{x^j}|_p$ and $i\partial_{x^j}|_p$ would be considered linearly independent and hence $T_pM^{\mathbb{C}}$ has real dimension $4d$.

In exact analogy with the real case, we can define the dual to $T_pM^{\mathbb{C}}$, which we denote by $T_p^*M^{\mathbb{C}} = T_p^*M \otimes \mathbb{C}$, with the one–forms basis

$$T_p^*M^{\mathbb{C}} : \{dz^1|_p, ..., dz^d|_p, d\bar{z}^1|_p, ..., d\bar{z}^d|_p\}.$$

For certain types of complex manifolds M, it is worthwhile to refine the definition of the complexified tangent and cotangent spaces, which pulls apart the holomorphic and anti–holomorphic directions in each of these two vector spaces. That is, we can write

$$T_pM^{\mathbb{C}} = T_pM^{(1,0)} \oplus T_pM^{(0,1)},$$

where $T_pM^{(1,0)}$ is the vector space spanned by $\{\partial_{z^1}|_p, ..., \partial_{z^d}|_p\}$ and $T_pM^{(0,1)}$ is the vector space spanned by $\{\partial_{\bar{z}^1}|_p, ..., \partial_{\bar{z}^d}|_p\}$. Similarly, we can write

$$T_p^*M^{\mathbb{C}} = T_p^*M^{(1,0)} \oplus T_p^*M^{(0,1)},$$

where $T_p^*M^{(1,0)}$ is the vector space spanned by $\{dz^1|_p, ..., dz^d|_p\}$ and $T_p^*M^{(0,1)}$ is the vector space spanned by $\{d\bar{z}^1|_p, ..., d\bar{z}^d|_p\}$. We call $T_pM^{(1,0)}$ the *holomorphic tangent space*; it has complex dimension d and we call $T_p^*M^{1,0}$ the *holomorphic cotangent space*. It also has complex dimension d. Their complements are known as the *anti–holomorphic* tangent and cotangent spaces respectively [Gre96].

Now, a *complex vector bundle* is a vector bundle $\pi : E \to M$ whose fiber bundle $\pi^{-1}(x)$ is a complex vector space. It is not necessarily a complex manifold, even if its base manifold M is a complex manifold. If a complex vector bundle also has the structure of a complex manifold, and is holomorphic, then it is called a holomorphic vector bundle.

A *Hermitian metric* on a complex vector bundle assigns a *Hermitian inner product* to every fiber bundle. The basic example is the trivial bundle $\pi : U \times \mathbb{C}^2 \to U$, where U is an open set in $\mathbb{R}^n$. Then a positive definite Hermitian matrix H defines a Hermitian metric by

$$\langle v, w \rangle = v^T H \bar{w},$$

where $\bar{w}$ is the complex conjugate of w. By a partition of unity, any complex vector bundle has a Hermitian metric.

In the special case of a complex manifold, the complexified tangent bundle $TM \otimes \mathbb{C}$ may have a Hermitian metric, in which case its real part is a Riemannian metric and its imaginary part is a nondegenerate alternating multilinear form ω. When ω is closed, i.e., in this case a symplectic form, then ω is a *Kähler form*.

On a holomorphic vector bundle with a Hermitian metric h, there is a unique connection compatible with h and the complex structure. Namely, it must be $\nabla = \partial + \bar{\partial}$.

A *Kähler structure* on a complex manifold M combines a Riemannian metric on the underlying real manifold with the complex structure. Such a structure brings together geometry and complex analysis, and the main examples come from algebraic geometry. When M has n complex dimensions, then it has $2n$ real dimensions. A Kähler structure is related to the unitary group $U(n)$, which embeds in $SO(2n)$ as the orthogonal matrices that preserve the *almost complex structure* (multiplication by i). In a coordinate chart, the complex structure of M defines a multiplication by i and the metric defines orthogonality for tangent vectors. On a Kähler manifold, these two notions (and their derivatives) are related.

A Kähler manifold is a complex manifold for which the exterior derivative of the fundamental form ω associated with the given *Hermitian metric* vanishes, so $d\omega = 0$. In other words, it is a complex manifold with a *Kähler structure*. It has a *Kähler form*, so it is also a symplectic manifold. It has a *Kähler metric*, so it is also a Riemannian manifold.

The simplest example of a Kähler manifold is a Riemann surface, which is a complex manifold of dimension 1. In this case, the imaginary part of any Hermitian metric must be a closed form since all 2−forms are closed on a real two–dimensional manifold.

A *Kähler form* is a closed two–form ω on a complex manifold M which is also the negative imaginary part of a Hermitian metric $h = g - iw$ is called a Kähler form. In this case, M is called a Kähler manifold and g, the real part of the Hermitian metric, is called a Kähler metric. The Kähler form combines the metric and the complex structure, $g(X,Y) = \omega(X,JY)$,where J is the *almost complex structure* induced by multiplication by i. Since the Kähler form comes from a Hermitian metric, it is preserved by J, since $h(X,Y) = h(JX,JY)$. The equation $d\omega = 0$ implies that the metric and the complex structure are related. It gives M a Kähler structure, and has many implications.

On $\mathbb{C}^2$, the Kähler form can be written as

$$\omega = -\frac{1}{2}i\left(dz_1 \wedge \overline{dz_1} + dz_2 \wedge \overline{dz_2}\right) = dx_1 \wedge dy_1 + dx_2 \wedge dy_2,$$

where $z_n = x_n + y_n$. In general, the Kähler form can be written in coordinates

$$\omega = g_{ij}\, dz_i \wedge \overline{dz_j},$$

where g_{ij} is a Hermitian metric, the real part of which is the Kähler metric. Locally, a Kähler form can be written as $\partial\bar{\partial}f$, where f is a function called a *Kähler potential*. The Kähler form is a real $(1,1)-$complex form. The Kähler potential is a real–valued function f on a Kähler manifold for which the Kähler form ω can be written as $\omega = i\partial\bar{\partial}f$, where,

$$\partial = \partial_{z_k} dz_k \qquad \text{and} \qquad \bar{\partial} = \partial_{\bar{z}_k} d\bar{z}_k.$$

Since the Kähler form ω is closed, it represents a cohomology class in the De Rham cohomology. On a compact manifold, it cannot be exact because $\omega^n/n! \neq 0$ is the volume form determined by the metric. In the special case of a projective variety, the Kähler form represents an integral cohomology class. That is, it integrates to an integer on any one–dimensional submanifold, i.e., an algebraic curve. The Kodaira embedding theorem says that if the Kähler form represents an integral cohomology class on a compact manifold, then it must be a projective variety. There exist Kähler forms which are not projective algebraic, but it is an open question whether or not any Kähler manifold can be deformed to a projective variety (in the compact case).

A Kähler form satisfies *Wirtinger's inequality*,

$$|\omega(X,Y)| \leq |X \wedge Y|\,,$$

where the r.h.s is the volume of the parallelogram formed by the tangent vectors X and Y. Corresponding inequalities hold for the *exterior powers* of ω. Equality holds iff X and Y form a complex subspace. Therefore, there is a *calibration form*, and the complex submanifolds of a Kähler manifold are *calibrated submanifolds*. In particular, the complex submanifolds are locally volume minimizing in a Kähler manifold. For example, the graph of a holomorphic function is a locally area–minimizing surface in $\mathbb{C}^2 = \mathbb{R}^4$.

Kähler identities is a collection of identities which hold on a Kähler manifold, also called the *Hodge identities*. Let ω be a Kähler form, $d = \partial + \bar{\partial}$ be the exterior derivative, $[A,B] = AB - BA$ be the commutator of two differential operators, and A^* denote the formal adjoint of A. The following operators also act on differential forms α on a Kähler manifold:

$$L(\alpha) = \alpha \wedge \omega, \qquad \Lambda(\alpha) = L^*(\alpha) = \alpha \lrcorner \omega, \qquad d_c = -JdJ,$$

where J is the almost complex structure, $J = -I$, and $\lrcorner$ denotes the interior product. Then

$$[L,\bar{\partial}] = [L,\partial] = 0, \qquad [\Lambda,\bar{\partial}^*] = [\Lambda,\partial^*] = 0,$$
$$[L,\bar{\partial}^*] = -i\partial, \qquad [L,\partial^*] = i\bar{\partial}, \qquad [\Lambda,\bar{\partial}] = -i\partial^*, \qquad [\Lambda,\partial] = -i\bar{\partial}.$$

These identities have many implications. For instance, the two operators

$$\Delta_d = dd^* + d^*d \qquad \text{and} \qquad \Delta_{\bar{\partial}} = \bar{\partial}\bar{\partial}^* + \bar{\partial}^*\bar{\partial}$$

(called Laplacians because they are elliptic operators) satisfy

$$\Delta_d = 2\Delta_{\bar{\partial}}.$$

At this point, assume that M is also a compact manifold. Along with Hodge's theorem, this equality of Laplacians proves the *Hodge decomposition*. The operators L and Λ commute with these Laplacians. By *Hodge's theorem* (see Chapter 4 below), they act on cohomology, which is represented by *harmonic forms*. Moreover, defining

$$H = [L, \Lambda] = \sum (p + q - n)\, \Pi^{p,q},$$

where $\Pi^{p,q}$ is projection onto the $(p, q)-$Dolbeault cohomology, they satisfy

$$[L, \Lambda] = H, \qquad [H, L] = -2L, \qquad [H, \Lambda] = 2L.$$

In other words, these operators provide a group representation of the special linear Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ on the complex cohomology of a compact Kähler manifold (*Lefschetz theorem*).

2.5.4 Conformal Killing–Riemannian Geometry

In this subsection we present some basic facts from *conformal Killing–Riemannian geometry*. In mechanics (see Chapter 3) it is well–known that *symmetries of Lagrangian* or *Hamiltonian* result in *conservation laws*, that are used to deduce constants of motion for the trajectories (geodesics) on the configuration manifold M. The same constants of motion are obtained using geometric language, where a *Killing vector–field* is the standard tool for the description of symmetry [MTW73]. A Killing vector–field ξ^i is a vector–field on a Riemannian manifold M with metrics g, which in coordinates $x^j \in M$ satisfies the *Killing equation*

$$\xi^{i;j} + \xi^{j;i} = \xi^{(i;j)} = 0, \qquad \text{or} \qquad \mathcal{L}_{\xi^i} g_{ij} = 0, \tag{2.107}$$

where semicolon denotes the covariant derivative on M (as in (2.90) above), the indexed bracket denotes the tensor symmetry, and $\mathcal{L}$ is the Lie derivative.

The *conformal Killing vector–fields* are, by definition, infinitesimal conformal symmetries i.e., the flow of such vector–fields preserves the conformal class of the metric. The number of linearly–independent conformal Killing fields measures the degree of *conformal symmetry of the manifold*. This number is bounded by $\frac{1}{2}(n+1)(n+2)$, where n is the dimension of the manifold. It is the maximal one if the manifold is conformally flat [Bau00].

Now, to properly initialize our conformal geometry, recall that *conformal twistor spinor–fields* φ were introduced by R. Penrose into physics (see [Pen67, PR86]) as solutions of the *conformally covariant twistor equation*

$$\nabla^S_X \varphi + \frac{1}{n} X \cdot D\varphi = 0,$$

for each vector–fields X on a Riemannian manifold (M,g), where D is the *Dirac operator*. Each twistor spinor–field φ on (M,g) defines a *conformal vector–field* V_φ on M by

$$g(V_\varphi, X) = i^{k+1}\,\langle X\cdot\varphi,\varphi\rangle.$$

Also, each twistor spinor–field φ that satisfies the *Dirac equation* on (M,g),

$$D\varphi = \mu\varphi,$$

is called a *Killing spinor–field*. Each twistor spinor–field without zeros on (M,g) can be transformed by a conformal change of the metric g into a Killing spinor–field [Bau00].

Conformal Killing Vector–Fields and Forms on M

The space of all conformal Killing vector–fields forms the Lie algebra of the *isometry group* of a Riemannian manifold (M,g) and the number of linearly independent Killing vector–fields measures the *degree of symmetry* of M. It is known that this number is bounded from above by the dimension of the isometry group of the standard sphere and, on compact manifolds, equality is attained if and only if the manifold M is isometric to the standard sphere or the real projective space. Slightly more generally one can consider *conformal vector–fields*, i.e., vector–fields with a flow preserving a given conformal class of metrics. There are several geometric conditions which force a conformal vector–field to be Killing [Sem02].

A natural generalization of conformal vector–fields are the *conformal Killing forms* [Yan52], also called *twistor forms* [MS03]. These are $p-$forms α satisfying for any vector–field X on the manifold M the *Killing–Yano equation*

$$\nabla_X\,\alpha\ -\ \tfrac{1}{p+1}\,X\,\lrcorner\,d\alpha\ +\ \tfrac{1}{n-p+1}\,X^*\,\wedge\,d^*\alpha\ =\ 0, \tag{2.108}$$

where n is the dimension of the manifold (M,g), ∇ denotes the covariant derivative of the Levi–Civita connection on M, X^* is $1-$form dual to X and $\lrcorner$ is the operation dual to the wedge product on M. It is easy to see that a conformal Killing $1-$form is dual to a conformal vector–field. *Coclosed* conformal Killing $p-$forms are called *Killing forms*. For $p=1$ they are dual to Killing vector–fields.

Let α be a Killing $p-$form and let γ be a geodesic on (M,g), i.e., $\nabla_{\dot\gamma}\,\dot\gamma = 0$. Then

$$\nabla_{\dot\gamma}\,(\dot\gamma\lrcorner\,\alpha)\ =\ (\nabla_{\dot\gamma}\,\dot\gamma)\lrcorner\,\alpha\ +\ \dot\gamma\lrcorner\,\nabla_{\dot\gamma}\,\alpha\ =\ 0,$$

i.e., $\dot\gamma\lrcorner\,\alpha$ is a $(p-1)$–form parallel along the geodesic γ and in particular its length is constant along γ.

The l.h.s of equation (2.108) defines a first order elliptic differential operator T, the so–caled *twistor operator*. Equivalently one can describe a conformal

Killing form as a form in the kernel of twistor operator T. From this point of view conformal Killing forms are similar to Penrose's *twistor spinors* in *Lorentzian spin geometry*. One shared property is the conformal invariance of the defining equation. In particular, any form which is parallel for some metric g, and thus a Killing form for trivial reasons, induces non–parallel conformal Killing forms for metrics conformally equivalent to g (by a non–trivial change of the metric) [Sem02].

Conformal Killing Tensors and Laplacian Symmetry on M

In an nD Riemannian manifold (M, g), a *Killing tensor–field* (of order 2) is a symmetric tensor K^{ab} satisfying (generalizing (2.107))

$$K^{(ab;c)} = 0. \tag{2.109}$$

A *conformal Killing tensor–field* (of order 2) is a symmetric tensor Q^{ab} satisfying

$$Q^{(ab;c)} = q^{(a} g^{bc)}, \qquad \text{with} \qquad q^a = (Q^{,a} + 2Q_d^{a;d})/(n+2), \tag{2.110}$$

where comma denotes partial derivative and $Q = Q_d^d$. When the associated *conformal vector* q^a is nonzero, the conformal Killing tensor will be called *proper* and otherwise it is a (ordinary) Killing tensor. If q^a is a Killing vector, Q^{ab} is referred to as a *homothetic Killing tensor*. If the associated conformal vector $q^a = q^{,a}$ is the gradient of some scalar field q, then Q^{ab} is called a *gradient conformal Killing tensor*. For each gradient conformal Killing tensor Q^{ab} there is an associated Killing tensor K^{ab} given by

$$K^{ab} = Q^{ab} - qg^{ab}, \tag{2.111}$$

which is defined only up to the addition of a constant multiple of the inverse metric tensor g^{ab}.

Some authors define a conformal Killing tensor as a *trace–free* tensor P^{ab} satisfying $P^{(ab;c)} = p^{(a} g^{bc)}$. Note that there is no contradiction between the two definitions: if P^{ab} is a trace–free conformal Killing tensor then for any scalar field λ, $P^{ab} + \lambda g^{ab}$ is a conformal Killing tensor and conversely if Q^{ab} is a conformal Killing tensor, its trace–free part $Q^{ab} - \frac{1}{n} Q g^{ab}$ is a trace–free Killing tensor [REB03].

Killing tensor–fields are of importance owing to their connection with quadratic first integrals of the geodesic equations: if p^a is tangent to an affinely parameterized geodesic (i.e., $p^a{}_{;b} p^b = 0$) it is easy to see that $K_{ab} p^a p^b$ is constant along the geodesic. For conformal Killing tensors $Q_{ab} p^a p^b$ is constant along *null* geodesics and here, only the trace–free part of Q_{ab} contributes to the constants of motion. Both Killing tensors and conformal Killing tensors are also of importance in connection with the separability of the Hamilton–Jacobi equations [CH64] (as well as other PDEs).

A Killing tensor is said to be *reducible* if it can be written as a constant linear combination of the metric and symmetrised products of Killing vectors,

$$K_{ab} = a^0 g_{ab} + a^{IJ} \xi_{I(a} \xi_{|J|b)}, \tag{2.112}$$

where ξ_I for $I = 1 \dots N$ are the Killing vectors admitted by the manifold (M, g) and a^0 and a^{IJ} for $1 \leq I \leq J \leq N$ are constants. Generally one is interested only in Killing tensors which are not reducible since the quadratic constant of motion associated with a reducible Killing tensor is simply a constant linear combination of $p^a p_a$ and of pairwise products of the linear constants of motion $\xi_{Ia} p^a$ [REB03].

More generally, any linear differential operator on a Riemannian manifold (M, g) may be written in the form [EG91, Eas02]

$$\mathcal{D} = V^{bc\cdots d} \nabla_b \nabla_c \cdots \nabla_d + \text{ lower order terms},$$

where $V^{bc\cdots d}$ is symmetric in its indices, and $\nabla_a = \partial / \partial x^a$ (differentiation in coordinates). This tensor is called the symbol of $\mathcal{D}$. We shall write $\phi^{(ab\cdots c)}$ for the symmetric part of $\phi^{ab\cdots c}$.

Now, a *conformal Killing tensor* on (M, g) is a symmetric trace–free tensor field, with s indices, satisfying

$$\text{the trace–free part of } \nabla^{(a} V^{bc\cdots d)} = 0, \tag{2.113}$$

or, equivalently,

$$\nabla^{(a} V^{bc\cdots d)} = g^{(ab} T^{c\cdots d)}, \tag{2.114}$$

for some tensor field $T^{c\cdots d}$ or, equivalently,

$$\nabla^{(a} V^{bc\cdots d)} = \tfrac{s}{n+2s-2} g^{(ab} \nabla_e V^{c\cdots d)e}, \tag{2.115}$$

where $\nabla^a = g^{ab} \nabla_b$ (the standard convention of raising and lowering indices with the metric tensor g_{ab}). When $s = 1$, these equations define a *conformal Killing vector*.

M. Eastwood proved the following theorem: *Any symmetry $\mathcal{D}$ of the Laplacian $\Delta = \nabla^a \nabla_a$ on a Riemannian manifold (M, g) is canonically equivalent to one whose symbol is a conformal Killing tensor* [EG91, Eas02].

2.6 Symplectic Geometry in Human–Like Biomechanics

In this section we develop the basic techniques of symplectic geometry on the biomechanical manifold M [Iva04].

2.6.1 Symplectic Algebra

Symplectic algebra works in the category of symplectic vector spaces V_i and linear symplectic mappings $t \in L(V_i, V_j)$ [Put93].

Let V be a nD real vector space and $L^2(V, \mathbb{R})$ the space of all bilinear maps from $V \times V$ to $\mathbb{R}$. We say that a *bilinear map* $\omega \in L^2(V, \mathbb{R})$ is *nondegenerate*, i.e., if $\omega(v_1, v_2) = 0$ for all $v_2 \in V$ implies $v_1 = 0$.

If $\{e_1, ..., e_n\}$ is a basis of V and $\{e^1, ..., e^n\}$ is the dual basis, $\omega_{ij} = \omega(e_i, e_j)$ is the matrix of ω. A bilinear map $\omega \in L^2(V, \mathbb{R})$ is nondegenerate iff its matrix ω_{ij} is nonsingular. The transpose ω^t of ω is defined by $\omega^t(e_i, e_j) = \omega(e_j, e_i)$. ω is *symmetric* if $\omega^t = \omega$, and *skew–symmetric* if $\omega^t = -\omega$.

Let $A^2(V)$ denote the space of skew–symmetric bilinear maps on V. An element $\omega \in A^2(V)$ is called a 2−form on V. If $\omega \in A^2(V)$ is nondegenerate then in the basis $\{e_1, ..., e_n\}$ its matrix $\omega(e_i, e_j)$ has the form $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$.

A *symplectic form* on a real vector space V of dimension $2n$ is a nondegenerate 2−form $\omega \in A^2(V)$. The pair (V, ω) is called a symplectic vector space. If (V_1, ω_1) and (V_2, ω_2) are symplectic vector spaces, a linear map $t \in L(V_1, V_2)$ is a *symplectomorphism* (i.e., a symplectic mapping) iff $t^*\omega_2 = \omega_1$. If (V, ω) is a symplectic vector space, we have an orientation Ω_ω on V given by

$$\Omega_\omega = \frac{(-1)^{\frac{n(n-1)}{2}}}{n!}\omega^n.$$

Let (V, ω) be a $2n$D symplectic vector space and $t \in L(V, V)$ a symplectomorphism. Then t is volume preserving, i.e., $t^*(\Omega_\omega) = \Omega_\omega$, and $\det_{\Omega_\omega}(t) = 1$.

The set of all symplectomorphisms $t : V \to V$ of a $2n$D symplectic vector space (V, ω) forms a group under composition, called the *symplectic group*, denoted by $Sp(V, \omega)$.

In matrix notation, there is a basis of V in which the matrix of ω is $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$, such that $J^{-1} = J^t = -J$, and $J^2 = -I$. For $t \in L(V, V)$ with matrix $T = [T^i_j]$ relative to this basis, the condition $t \in Sp(V, \omega)$, i.e., $t^*\omega = \omega$, becomes

$$T^t J T = J.$$

In general, by definition a matrix $A \in M_{2n\times 2n}(\mathbb{R})$ is symplectic iff $A^t J A = J$.

Let (V, ω) be a symplectic vector space, $t \in Sp(V, \omega)$ and $\lambda \in \mathbb{C}$ an eigenvalue of t. Then λ^{-1}, $\bar{\lambda}$ and $\bar{\lambda}^{-1}$ are eigenvalues of t.

2.6.2 Symplectic Geometry on *M*

Symplectic geometry is a globalization of symplectic algebra [Put93]; it works in the category **Symplec** of symplectic manifolds M and symplectic diffeomorphisms f. The phase–space of a conservative dynamical system is a symplectic manifold, and its time evolution is a one–parameter family of symplectic diffeomorphisms.

A *symplectic form* or a symplectic structure on a smooth (*i.e.*, C^k) manifold M is a nondegenerate closed 2−form ω on M, i.e., for each $x \in M$ $\omega(x)$ is nondegenerate, and $d\omega = 0$. A *symplectic manifold* is a pair (M, ω) where M is a smooth $2n$D manifold and ω is a symplectic form on it. If (M_1, ω_1) and (M_2, ω_2) are symplectic manifolds then a smooth map $f : M_1 \to M_2$ is called *symplectic map* or *canonical transformation* if $f^*\omega_2 = \omega_1$.

For example, any symplectic vector space (V, ω) is also a symplectic manifold; the requirement $d\omega = 0$ is automatically satisfied since ω is a constant map. Also, any orientable, compact surface Σ is a symplectic manifold; any nonvanishing 2−form (volume element) ω on Σ is a symplectic form on Σ.

If (M, ω) is a symplectic manifold then it is orientable with the standard volume form

$$\Omega_\omega = \frac{(-1)^{\frac{n(n-1)}{2}}}{n!}\omega^n,$$

If $f : M \to M$ is a symplectic map, then f is volume preserving, $\det_{\Omega_\omega}(f) = 1$ and f is a local diffeomorphism.

In general, if (M, ω) is a $2n$D compact symplectic manifold then ω^n is a volume element on M, so the De Rham cohomology class $[\omega^n] \in H^{2n}(M, \mathbb{R})$ is nonzero. Since $[\omega^n] = [\omega]^n$, $[\omega] \in H^2(M, \mathbb{R})$ and all of its powers through the nth must be nonzero as well. The existence of such an element of $H^2(M, \mathbb{R})$ is a *necessary condition* for the compact manifold to admit a symplectic structure.

However, if M is a $2n$D compact manifold without boundary, then there does not exist any exact symplectic structure, $\omega = d\theta$ on M, as its total volume is zero (by Stokes' theorem),

$$\int_M \Omega_\omega = \frac{(-1)^{\frac{n(n-1)}{2}}}{n!}\int_M \omega^n = \frac{(-1)^{\frac{n(n-1)}{2}}}{n!}\int_M d(\theta \wedge \omega^{n-1}) = 0.$$

For example, spheres S^{2n} do not admit a symplectic structure for $n \geq 2$, since the second De Rham group vanishes, i.e., $H^2(S^{2n}, \mathbb{R}) = 0$. This argument applies to any compact manifold without boundary and having $H^2(M, \mathbb{R}) = 0$.

In mechanics, the phase–space is the cotangent bundle T^*M of a configuration space M. There is a natural symplectic structure on T^*M that is usually defined as follows. Let M be a smooth nD manifold and pick local coordinates $\{dq^1, ..., dq^n\}$. Then $\{dq^1, ..., dq^n\}$ defines a basis of the tangent space T_q^*M, and by writing $\theta \in T_q^*M$ as $\theta = p_i dq^i$ we get local coordinates $\{q^1, ..., q^n, p_1, ..., p_n\}$ on T^*M. Define the canonical symplectic form ω on T^*M by

$$\omega = dp_i \wedge dq^i.$$

This 2−form ω is obviously independent of the choice of coordinates $\{q^1, ..., q^n\}$ and independent of the base point $\{q^1, ..., q^n, p_1, ..., p_n\} \in T_q^*M$; therefore, it is locally constant, and so $d\omega = 0$.

The canonical 1−form θ on T^*M is the unique 1−form with the property that, for any 1−form β which is a section of T^*M we have $\beta^*\theta = \theta$.

Let $f : M \to M$ be a diffeomorphism. Then T^*f preserves the canonical $1-$form θ on T^*M, i.e., $(T^*f)^*\theta = \theta$. Thus T^*f is symplectic diffeomorphism.

If (M, ω) is a $2n$D symplectic manifold then about each point $x \in M$ there are local coordinates $\{q^1, ..., q^n, p_1, ..., p_n\}$ such that $\omega = dp_i \wedge dq^i$. These coordinates are called canonical or symplectic. By the Darboux theorem, ω is constant in this local chart, i.e., $d\omega = 0$.

2.6.3 Momentum Map and Symplectic Reduction of M

Let (M, ω) be a connected symplectic manifold and $\phi : G \times M \to M$ a symplectic action of the Lie group G on M, that is, for each $g \in G$ the map $\phi_g : M \to M$ is a symplectic diffeomorphism. If for each $\xi \in \mathfrak{g}$ there exists a globally defined function $\hat{J}(\xi) : M \to \mathbb{R}$ such that $\xi_M = X_{\hat{J}(\xi)}$, then the map $J : M \to \mathfrak{g}^*$, given by

$$J : x \in M \mapsto J(x) \in \mathfrak{g}^*, \qquad J(x)(\xi) = \hat{J}(\xi)(x)$$

is called the *momentum map* for ϕ [MR99, Put93].

Since ϕ is symplectic, $\phi_{\exp(t\xi)}$ is a one parameter family of canonical transformations, i.e., $\phi^*_{\exp(t\xi)}\omega = \omega$, hence ξ_M is locally Hamiltonian and not generally Hamiltonian. That is why not every symplectic action has a momentum map. $\phi : G \times M \to M$ is Hamiltonian iff $\hat{J} : \mathfrak{g} \to C^k(M, \mathbb{R})$ is a Lie algebra homomorphism.

Let $H : M \to \mathbb{R}$ be G–invariant, that is $H\left(\phi_g(x)\right) = H(x)$ for all $x \in M$ and $g \in G$. Then $\hat{J}(\xi)$ is a constant of motion for dynamics generated by H.

Let ϕ be a symplectic action of G on (M, ω) with the momentum map J. Suppose $H : M \to \mathbb{R}$ is G–invariant under this action. Then the Noether's theorem states that J is a constant of motion of H, i.e., $J \circ \phi_t = J$, where ϕ_t is the flow of X_H.

A *Hamiltonian action* is a symplectic action with an Ad^*–equivariant momentum map J, i.e.,

$$J\left(\phi_g(x)\right) = Ad^*_{g^{-1}}\left(J(x)\right),$$

for all $x \in M$ and $g \in G$.

Let ϕ be a symplectic action of a Lie group G on (M, ω). Assume that the symplectic form ω on M is exact, i.e., $\omega = d\theta$, and that the action ϕ of G on M leaves the one form $\theta \in M$ invariant. Then $J : M \to \mathfrak{g}^*$ given by $\left(J(x)\right)(\xi) = \left(i_{\xi_M}\theta\right)(x)$ is an Ad^*–equivariant momentum map of the action.

In particular, in the case of the cotangent bundle $(M = T^*M,\, \omega = d\theta)$ of a mechanical configuration manifold M, we can *lift* up an action ϕ of a Lie group G on M to obtain an action of G on T^*M. To perform this lift, let G act on M by transformations $\phi_g : M \to M$ and define the *lifted action* to the cotangent bundle by $(\phi_g)_* : T^*M \to T^*M$ by pushing forward one forms, $(\phi_g)_*(\alpha) \cdot v = \alpha\left(T\phi_g^{-1}v\right)$,where $\alpha \in T^*_qM$ and $v \in T_{\phi_g(q)}M$. The lifted action $(\phi_g)_*$ preserves the canonical one form θ on T^*M and the momentum map for $(\phi_g)_*$ is given by

$$J : T^*M \to \mathfrak{g}^*, \qquad J(\alpha_q)(\xi) = \alpha_q(\xi_M(q)).$$

For example, let $M = \mathbb{R}^n$, $G = \mathbb{R}^n$ and let G act on $\mathbb{R}^n$ by translations:

$$\phi : (t, q) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto t + q \in \mathbb{R}^n.$$

Then $\mathfrak{g} = \mathbb{R}^n$ and for each $\xi \in \mathfrak{g}$ we have $\xi_{\mathbb{R}^n}(q) = \xi$.

In case of the group of rotations in $\mathbb{R}^3$, $M = \mathbb{R}^3$, $G = SO(3)$ and let G act on $\mathbb{R}^3$ by $\phi(A, q) = A \cdot q$. Then $\mathfrak{g} \simeq \mathbb{R}^3$ and for each $\xi \in \mathfrak{g}$ we have $\xi_{\mathbb{R}^3}(q) = \xi \times q$.

Let G act transitively on (M, ω) by a Hamiltonian action. Then $J(M) = \{Ad^*_{g^{-1}}(J(x))\, | g \in G\}$ is a coadjoint orbit.

Now, let (M, ω) be a symplectic manifold, G a Lie group and $\phi : G \times M \to M$ a Hamiltonian action of G on M with Ad^*–equivariant momentum map $J : M \to \mathfrak{g}^*$. Let $\mu \in \mathfrak{g}^*$ be a regular value of J; then $J^{-1}(\mu)$ is a submanifold of M such that $\dim\left(J^{-1}(\mu)\right) = \dim(M) - \dim(G)$. Let $G_\mu = \{g \in G | Ad^*_g \mu = \mu\}$ be the isotropy subgroup of μ for the coadjoint action. By Ad^*–equivariance, if $x \in J^{-1}(\mu)$ then $\phi_g(x) = J^{-1}(\mu)$ for all $g \in G$, i.e., $J^{-1}(\mu)$ is invariant under the induced G_μ–action and we can form the quotient space $M_\mu = J^{-1}(\mu)/G_\mu$, called the *reduced phase–space* at $\mu \in \mathfrak{g}^*$.

Let (M, ω) be a symplectic $2n$D manifold and let $f_1, ..., f_k$ be k functions in involution, i.e., $\{f_i, f_j\}_\omega = 0$, $i = 1, ..., k$. Because the flow of X_{f_i} and X_{f_j} commute, we can use them to define a symplectic action of $G = \mathbb{R}^k$ on M. Here $\mu \in \mathbb{R}^k$ is in the range space of $f_1 \times ... \times f_k$ and $J = f_1 \times ... \times f_k$ is the momentum map of this action. Assume that $\{df_1, ..., df_k\}$ are independent at each point, so μ is a regular value for J. Since G is Abelian, $G_\mu = G$ so we get a symplectic manifold $J^{-1}(\mu)/G$ of dimension $2n - 2k$. If $k = n$ we have *integrable systems*.

For example, let $G = SO(3)$ and $(M, \omega) = \left(\mathbb{R}^6, \sum_{i=1}^3 dp_i \wedge dq^i\right)$, and the action of G on $\mathbb{R}^6$ is given by $\phi : (R, (q, p)) \mapsto (R_q, R_p)$. Then the momentum map is the well known angular momentum and for each $\mu \in \mathfrak{g}^* \simeq \mathbb{R}^3 \mu \neq 0$, $G_\mu \simeq S^1$ and the reduced phase–space (M_μ, ω_μ) is $(T^*\mathbb{R},\, \omega = dp_i \wedge dq^i)$, so that $\dim(M_\mu) = \dim(M) - \dim(G) - \dim(G_\mu)$. This reduction is in celestial mechanics called by Jacobi 'the elimination of the nodes'.

The equations of motion: $\dot{f} = \{f, H\}_\omega$ on M reduce to the equations of motion: $\dot{f}_\mu = \{f_\mu, H_\mu\}_{\omega_\mu}$ on M_μ (see [MR99]).

2.7 The Covariant Force Functor

We summarize this geometrical Chapter by stating that our central construct, the *covariant force law*, $F_i = m g_{ij} a^j$ (see subsection A.1.4 in Appendix), in categorical language represents the *covariant force functor* $\mathcal{F}_*$ defined by the following commutative diagram:

$$\begin{tikzcd}
TT^*M \arrow[rr, "\mathcal{F}_*"] & & TTM \\
T^*M = \{x^i, p_i\} \arrow[u, "F_i = \dot{p}_i"] & & TM = \{x^i, v^i\} \arrow[u, "a^i = \dot{\bar{v}}^i"'] \\
& M = \{x^i\} \arrow[ul, "p_i"] \arrow[ur, "v^i = \dot{x}^i"'] &
\end{tikzcd}$$

saying that the force 1–form–field $F_i = \dot{p}_i$, defined on the mixed tangent–cotangent bundle TT^*M, causes the acceleration vector–field $a^i = \dot{\bar{v}}^i$, defined on the second tangent bundle TTM of the configuration manifold M.

The Lie biomechanical functors (defined in the section 3.5 below) represent special versions of the fundamental force functor $\mathcal{F}_* : TT^*M \longrightarrow TTM$.

The corresponding *contravariant acceleration functor* is defined as its inverse map $\mathcal{F}^* : TTM \longrightarrow TT^*M$.