Vector fields, flows, and Lie theory

Daniel W. Crews

March 16, 2023

1 Introduction

Give the world the Lie [transform].

- *A.N. Kaufman* after *W. Raleigh*

The natural setting for mechanics is through the theory of vector fields on manifolds.[1] Manifolds are locally Euclidean geometric entities, so one can study vector fields on Euclidean spaces and be well-prepared for global theory. These notes summarizes background for the techniques of differential geometry used in kinetic turbulence theory without going into excessive rigor, as many excellent books exist for that.[2]

Essentially a vector field is a function which attaches a vector to each point of a manifold. Vector fields generate flows as the collection of integral curves induced by the field, e.g. streamlines in fluid mechanics. Flows form one-parameter groups of transformations (with e.g. time as parameter). There is a correspondance between vector fields and flows, and ODEs and their solutions.

The vector field acts as the infinitesimal generator of the system evolution. This was discovered by Sophus Lie and forms the basis of the theory of Lie algebras.[3] Associated with a vector field are two important concepts: its direct action on objects (called a Lie derivative), and the action of its flow on objects (called a Lie transform).

2 Theory of vector spaces

Before understanding vector *fields* it's worthwhile to recall some properties of a vector *space*. The reader is assumed to be familiar with the basic properties of a vector space. This section reviews the idea of the *dual space* of linear functionals on a vector space, sheds light on the vector operations of inner and outer (tensor) products, and discusses the generalization to multilinear functions on a vector space and their own dual pairs (the k-vectors and k-forms).

2.1 Linear functionals and the dual space

Recall that a space of functions may itself be a vector space. Each vector space V is associated with a number of other spaces consisting of certain functionals on that space. The simplest consists

[1] For example, phase space is best described as the cotangent bundle of configuration space.

[2] An introductory book on manifolds is Loring Tu's "An Introduction to Manifolds", a detailed and mathematically rigorous text is that of Shahshahani's "Differentiable Manifolds", an intuitive and concise introduction to Lie algebras and flows is P. Olver's "Application of Lie Groups to Differential Equations", and a good explanation of Lie transform is in Fasso (1990). A good summary of manifolds in variational mechanics is also in the thesis of M. Kraus (2013).

[3] This monumental achievement explains why nearly everything in the theory has the name Lie.

of *linear* functionals and is called the *dual space* V^*. If V is a vector space of dimension N with orthonormal basis $\{e_i\}|_{i=1}^N$, then vectors are represented by

$$\boldsymbol{v} = \sum_{i=1}^{N} v^i e_i \tag{1}$$

with components $v^i \in \mathbb{R}$. Now consider a linear functional $f \in \mathcal{L}(V, \mathbb{R})$, i.e. $f\colon V \to \mathbb{R}$. As a vector space's dimension is essentially its number of independent directions, the dimension of $\mathcal{L}(V, \mathbb{R})$ is also equal to N. In particular consider the linear functional f^i which returns the i'th component of the vector $\boldsymbol{v}$, i.e. $f^i(\boldsymbol{v}) = v^i$ (the i'th projector). Suppose $\mathcal{L}(V, \mathbb{R})$ has some basis $\{b^j\}|_{j=1}^N$ so that $f^i = \sum_{j=1}^N c_j b^j$. Then by its definition

$$f^i(\boldsymbol{v}) = v^i \qquad \Longrightarrow \qquad f^i(\boldsymbol{v}) = \sum_{j=1}^{N}\sum_{i=1}^{N} c_j v^i b^j(e_i) = \sum_{j=1}^{N} c_j v^j = v^i, \tag{2}$$

so that necessary consequences of consistency are that $b^j(e_i) = \delta_i^j$ and $c_j = \delta_j^i$, and therefore $f^i = e^i$. This identifies the *basis* of $\mathcal{L}(V, \mathbb{R})$ as the set $\{e^i\}_{i=1}^N$, those functionals which return the vector components[4]. Since each basis vector e_i is naturally associated with its partner projector functional e^i, the space $\mathcal{L}(V, \mathbb{R})$ is called the *dual space* V^*.

2.2 Inner and outer products

With the dual space concept, associated with any vector $\boldsymbol{v} \in V$ is its *covector*[5] $\boldsymbol{v}^* \in V^*$

$$\boldsymbol{v} = \sum_{i=1}^{N} v^i e_i, \qquad \boldsymbol{v}^* = \sum_{i=1}^{N} v_i e^i \tag{3}$$

whose components $v_i \in \mathbb{R}$ are numerically equal to the components v^i, but v^* is a functional rather than a vector[6]. The distinction is easily understood as that of row and column vectors. The original vector $\boldsymbol{v}$ is a column vector, while its dual is a row vector

$$\boldsymbol{v} = \begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^N \end{bmatrix}, \qquad \boldsymbol{v}^* = \begin{bmatrix} v_1 & v_2 & \dots & v_N \end{bmatrix}. \tag{4}$$

Now the concepts of inner and outer products can be elucidated with the notion of covectors. Considering the action of $\boldsymbol{v}^*$ on another vector $\boldsymbol{u} \in V$,

$$\boldsymbol{v}^*(\boldsymbol{u}) = \sum_{i=1}^{N}\sum_{j=1}^{N} v_i u^j e^i(e_j) = \sum_{i=1}^{N} v_i u^i = \langle \boldsymbol{v}, \boldsymbol{u} \rangle \tag{5}$$

[4] This is called the *standard basis* of the dual space.

[5] Also called its dual vector or one-form. The original vector is a $(1,0)$-tensor, its covector a $(0,1)$-tensor.

[6] Yet as the dual space is also a vector space, the dual vector v^* is a vector as well as a functional. In fact, it should be clear here that the space V and its dual space V^* are isomorphic under the inner product. This correspondence is extended to even infinite-dimensional spaces by the Riesz representation theorem. A caveat is that a vector space V over $\mathbb{C}$ is anti-isomorphic to V^*, in the sense of complex-conjugation.

reveals it as the inner product of $\boldsymbol{u}$ and $\boldsymbol{v}$. Here the dual basis e^i acts on the original basis e_j, giving $e^i(e_j) = \delta^i_j$. On the other hand, the order of operation can be reversed like

$$\boldsymbol{u}\boldsymbol{v}^* = \sum_{i=1}^{N}\sum_{j=1}^{N} u^i v_j e_i e^j = \boldsymbol{u} \otimes \boldsymbol{v} \tag{6}$$

resulting in the outer (tensor) product of u and v (i.e. a $(1,1)$-tensor). Here the dual basis component e^j has not yet operated on a vector. These two possible operations correspond to the inner and outer products, $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \boldsymbol{v}^T \boldsymbol{u}$ and $\boldsymbol{u} \otimes \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^T$.

2.3 Geometric interpretations of vectors (1-vectors) and covectors (1-forms)

Vectors $\boldsymbol{v}$ are naturally thought of as arrows originating from the origin of the vector space. Its associated covector $\boldsymbol{v}^*$ was identified as the function which returns the inner product of $\boldsymbol{v}$ with some other vector. Then geometrically, the covector $\boldsymbol{v}^*$ can be pictured as a set of planes normal to the direction of $\boldsymbol{v}$ and with spacing equal to its magnitude, $|\boldsymbol{v}|$, shown in Fig. 1. The number of times another vector $\boldsymbol{u}$ would pierce these planes is equal to their inner product, $\boldsymbol{v}^*(\boldsymbol{u}) = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$. The key to understanding the picture is that $\boldsymbol{v}^*$ is a function.

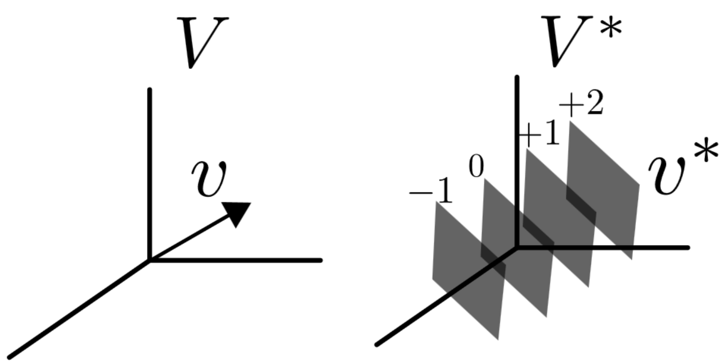

Figure 1: A vector $\boldsymbol{v}$ is an arrow emanating from the origin of its vector space V, while its covector $\boldsymbol{v}^*$ is a series of planes in the dual space V^* defining a function of the number of times another vector would pierce them.

2.4 Multivectors and their duals: the exterior algebra of a vector space

Having noted that the space of linear functionals $\mathcal{L}(V, \mathbb{R})$ is a vector space naturally associated with V, it's natural to also consider the space of bilinear functionals $\mathcal{L}_2(V, \mathbb{R})$ on the space V. This is the space of *two-forms* on V, and their dual objects are the *bivectors* which inhabit the *exterior square* of the vector space. One may continue this construction on up to the maximum dimension N of the vector space, considering k-vectors[7] in the kth exterior power of the space, and their duals the k-linear k-forms.

[7] More precisely, a k-vector is a multivector of grade k.

2.5 The wedge product

To navigate between the exterior spaces, the *wedge product* is defined[8]. As hinted in this section's motivation, it's initially more easily understood in terms of its action on two covectors $\boldsymbol{u}^*$, $\boldsymbol{v}^*$. Let us motivate the definition of this product [Arnold, 1978]. The product of these two linear functions should produce a bilinear function. Also, in order to express geometric relationships, the magnitude of its action on two vectors $\boldsymbol{\xi}$, $\boldsymbol{\zeta} \in V$ should be the oriented area of the parallelogram spanned by the vectors $[\boldsymbol{u}^*(\boldsymbol{\xi}), \boldsymbol{u}^*(\boldsymbol{\zeta})]^T$, $[\boldsymbol{v}^*(\boldsymbol{\xi}), \boldsymbol{v}^*(\boldsymbol{\zeta})]^T$. If the two vectors are equal, that area is zero, so the wedge product should be alternating. The determinant fits these properties perfectly, giving the definition

$$(\boldsymbol{u}^* \wedge \boldsymbol{v}^*)(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \begin{vmatrix} \boldsymbol{u}^*(\boldsymbol{\xi}) & \boldsymbol{v}^*(\boldsymbol{\xi}) \\ \boldsymbol{u}^*(\boldsymbol{\zeta}) & \boldsymbol{v}^*(\boldsymbol{\zeta}) \end{vmatrix}. \tag{7}$$

of the two-form $\boldsymbol{u}^* \wedge \boldsymbol{v}^*$. The space of two-forms is denoted $\bigwedge^2(V)$ to distinguish it from $\mathcal{L}_2(V, \mathbb{R})$, as $\bigwedge^2(V)$ contains only alternating bilinear functions. This definition as a determinant of linear forms provides its properties of bilinearity and skew-symmetry,

$$(\alpha x + \beta y) \wedge z = \alpha x \wedge z + \beta y \wedge z, \tag{8}$$

$$x \wedge y = -y \wedge x. \tag{9}$$

In particular, the wedge product of two basic one-forms[9] e^i, e^j acts on the vectors $\boldsymbol{u}$, $\boldsymbol{v}$ as

$$(e^i \wedge e^j)(\boldsymbol{u}, \boldsymbol{v}) = \begin{vmatrix} u^i & u^j \\ v^i & v^j \end{vmatrix} = u^i v^j - u^j v^i \tag{10}$$

which is the oriented area of the parallelogram spanned by the projections of $\boldsymbol{u}$, $\boldsymbol{v}$ on the e^i, e^j coordinate plane. Because these various planes are independent, the $\binom{N}{2} = \frac{N(N-1)}{2}$ wedges $e^i \wedge e^j \in \bigwedge^2(V)$ form a basis for the two-forms.

In a similar way, one may define general k-forms as further wedge products defining projected *k-volumes*. For example, the basis three-forms are $e^i \wedge e^j \wedge e^k \in \bigwedge^3(V)$ for the $\binom{N}{3}$-dimensional three-forms, which give the volume spanned by three projected vectors. At the maximal dimension N of the space V, its exterior power $\bigwedge^N(V)$ is one-dimensional, where the N-form $e^i \wedge \cdots \wedge e^N$ gives the oriented volume of the space.

2.6 Multivectors

If we now consider the dual space to the space of two-forms, we will return to vector-like objects associated with a vector space[10]. Using the duality principle, the basis bivectors are the wedge products $e_i \wedge e_j \in \bigwedge_2(V)$ which represent the oriented area of the parallelogram spanned by the standard basis vectors e_i, e_j of V[11]. In the literature, the basis objects $e_i \wedge e_j$ are called 2-blades. Between two vectors $\boldsymbol{u}$, $\boldsymbol{v} \in V$ their product $\boldsymbol{u} \wedge \boldsymbol{v}$ is a bivector whose (i, j) components are given through duality by Eqn. 10,

$$(\boldsymbol{u} \wedge \boldsymbol{v})|_{(i,j)} = \begin{vmatrix} u^i & v^j \\ u^i & v^j \end{vmatrix} e_i \wedge e_j = (u^i v^j - u^j v^i) e_i \wedge e_j. \tag{11}$$

[8] The wedge product is also called the *exterior product*.

[9] Basic one-forms meaning the standard dual basis of the space V.

[10] While the double-dual space $(V^*)^*$ is still a space of functions, it is isomorphic to its original space. This shows that even vectors as "arrows" are isomorphic to functions, which is important in the section on tangent spaces where tangent vectors will be noted as differential operators.

[11] That is, $e_i \wedge e_j$ is the oriented area object itself, as opposed to representation in a functional way as the wedge of dual basis vectors $e^i \wedge e^j$ does.

In a completely analogous way higher k-vectors are defined up the dimension N of V. These represent oriented k-volumes through their basis objects the k-blades, being linear combinations of the k-fold wedges of the standard basis vectors e^i. Combining the exterior powers of the space, $\bigwedge(V) = \bigoplus_{k=0}^{N} \bigwedge_k(V)$, and equipping with the wedge product defines the *exterior algebra* of the space V, and provides a way to measure oriented volumes of any dimension within the space. As $\sum_{k=0}^{N} \binom{N}{k} = 2^n$, the exterior algebra $\bigwedge(V)$ is 2^N-dimensional.

In summary, the kth exterior power $\bigwedge_k(V)$ consists of the antisymmetric type-$(k,0)$ tensors (contravariant) while the kth exterior power of the dual space $\bigwedge^k(V)$ consists of the antisymmetric type-$(0,k)$ tensors (covariant). Combinations form the multivectors of V.

2.7 Example: the exterior algebra of $\mathbb{R}^3$

Intuition for multivectors and covectors can be built in $\mathbb{R}^3$ where objects are visualizable. Written with a standard basis $\{e_1, e_2, e_3\}$ for $\mathbb{R}^3$, general vectors and 1-forms are

$$\boldsymbol{v} = v^1 e_1 + v^2 e_2 + v^3 e_3 = v^i e_i, \tag{12}$$

$$\boldsymbol{v}^* = v_1 e^1 + v_2 e^2 + v_3 e^3 = v_i e^i, \tag{13}$$

and represent an arrow and a series of planes normal to that arrow, as in Fig. 1. The bivectors and 2-forms on $\mathbb{R}^3$ are $\binom{3}{2} = 3$-dimensional, with coordinate representations

$$\omega_2 = \omega^{23} e_2 \wedge e_3 + \omega^{31} e_3 \wedge e_1 + \omega^{12} e_1 \wedge e_2 = \frac{1}{2!}\omega^{ij} e_i \wedge e_j, \tag{14}$$

$$\omega^2 = \omega_{23} e^2 \wedge e^3 + \omega_{31} e^3 \wedge e^1 + \omega_{12} e^1 \wedge e^2 = \frac{1}{2!}\omega_{ij} e^i \wedge e^j, \tag{15}$$

as visualized in Fig. 2. Note that the matrices of the bivectors and 2-forms are antisymmetric, and a factor $\frac{1}{2}$ is included in the summation formula for double-counting. As bivectors are often built out of vectors, it's worthwhile to include the bivector of $\boldsymbol{v} \wedge \boldsymbol{u}$,

$$\boldsymbol{v} \wedge \boldsymbol{u} = (v^2 u^3 - v^3 u^2) e_2 \wedge e_3 + (v^3 u^1 - v^1 u^3) e_3 \wedge e_1 + (v^1 u^2 - v^2 u^1) e_1 \wedge e_2. \tag{16}$$

Note its resemblence to the cross product. Finally, any trivector and 3-form in $\mathbb{R}$ is

$$\omega_3 = \Omega e_1 \wedge e_2 \wedge e_3, \qquad \omega^3 = \Omega e^1 \wedge e^2 \wedge e^3. \tag{17}$$

Because the trivectors are one-dimensional in $\mathbb{R}^3$, they are sometimes called pseudoscalars[12]. For example, the trivector $\boldsymbol{u} \wedge \boldsymbol{v} \wedge \boldsymbol{w}$ is the oriented volume of their spanned parallelepiped,

$$\boldsymbol{u} \wedge \boldsymbol{v} \wedge \boldsymbol{w} = \begin{vmatrix} u^1 & u^2 & u^3 \\ v^1 & v^2 & v^3 \\ w^1 & w^2 & w^3 \end{vmatrix} e^1 \wedge e^2 \wedge e^3 \tag{18}$$

with the associated 3-form giving, of course, the numerical value of this determinant. A general multivector in the exterior algebra $v \in \bigwedge(\mathbb{R}^3)$ is of the form

$$v = \alpha^0 + \alpha^i e_i + \frac{1}{2!}\omega^{ij} e_i \wedge e_j + \Omega e_1 \wedge e_2 \wedge e_3 \tag{19}$$

for some scalars $\alpha^i, \omega^{ij}, \Omega \in \mathbb{R}$.

[12] Note that the dimensionalities of the scalars agrees with the trivectors, and that of the vectors with the bivectors. This is an example of *Hodge duality* on the exterior algebra.

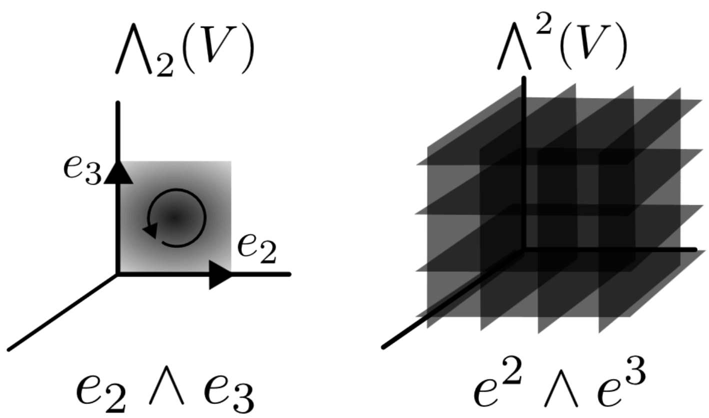

Figure 2: A basis bivector $e_2 \wedge e_3$ is an oriented area (i.e. "directed measure"), while its corresponding basis 2-form $e^2 \wedge e^3$ can be thought of as a series of intersecting planes defining a function of the area spanned by two vectors which it acts on.

2.8 The musical isomorphisms: raising and lowering indices with the metric

Previously it was noted that there is an isomorphism between a vector space and its dual, but the function was not identified. This is because it depends on the *metric* of the space, which defines its inner products. Associate with the vector space V its metric $g(\cdot,\cdot)$, a bilinear functional $g\colon V \times V \to \mathbb{R}$. The map $\tau\colon V \to V^*$ from the base space to its dual is called the *correlation*, and is defined as

$$g(\boldsymbol{v}, \boldsymbol{u}) = \tau(\boldsymbol{v})(\boldsymbol{u}). \tag{20}$$

Provided the metric is non-degenerate (that is, there is always some vector such that $g(\boldsymbol{v}, \cdot) \neq 0$), this establishes an isomorphism between V and V^*. Define the matrix of g such that $g_{ij} = g(e_i, e_j)$, i.e. as its action on the basis vectors of the space. If the matrix of the metric is the identity, then the metric gives the inner product as in Eqn. 5. An important special case is Minkowski spacetime, where the metric tensor g_{ij} contains a -1 on the diagonal. The metric introduces operations which raise and lower indices, called sharp ($\sharp\colon V^* \to V$) and flat ($\flat\colon V \to V^*$) respectively, in analogy with musical notation. The metric has vector arguments, so $\sharp$ and $\flat$ are defined through the metric as

$$\boldsymbol{v}_\flat(\boldsymbol{u}) = g(\boldsymbol{v}, \boldsymbol{u}), \qquad \boldsymbol{v}^*(\boldsymbol{u}) = g(\boldsymbol{v}^{*\sharp}, \boldsymbol{u}). \tag{21}$$

Because $g(\boldsymbol{v}, \boldsymbol{u}) = g_{ij} v^i u^j$, it follows that the components of $\boldsymbol{v}_\flat$ and $\boldsymbol{v}^{*\sharp}$ are given by

$$v_{\flat i} = g_{ij} v^j, \qquad v^{\sharp i} = g^{ij} v_j \tag{22}$$

where metric adjoint is defined so that $g^{ik} g_{kj} = \delta^i_j$. In the physics literature the symbols $\sharp$ and flat $\flat$ are often dropped, and the metric is said to raise and lower the indices. However in the mathematical literature they are usually kept, so it's important to be aware of them.

2.9 The basics of Hodge duality, or "what is the cross product?"

In the example of Section 2.7, it was observed that in $\mathbb{R}^3$ the dimensionalities of scalars and pseudoscalars ($\bigwedge_0(\mathbb{R}^3) = \bigwedge_3(\mathbb{R}^3) = 1$), and that of vectors and bivectors ($\bigwedge_1(\mathbb{R}^3) = \bigwedge_2(\mathbb{R}^3)$ were equal. This is due to the identity $\binom{N}{k} = \binom{N}{N-k}$, and is at the root of a fundamental result called

Hodge isomorphism. To state it in mathematical precision requires several pages of definitions, so an intuitive sketch is provided here.

When the exterior algebra is built on a metric space (V, g), the metric $g\colon V \times V \to \mathbb{R}$ can be *extended* to a functional $G\colon \bigwedge(V) \times \bigwedge(V) \to \mathbb{R}$ defined such that $G(\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_N; \boldsymbol{u}_1 \wedge \cdots \boldsymbol{u}_N) = \det(g(\boldsymbol{v}_i, \boldsymbol{u}_j))$ and $G(\omega_i; \omega_j) = 0$ if ω_i, ω_j are k-vectors of different sizes. The exterior algebra with an extended metric is called a *Grassmann algebra.* This construction allows the entire exterior space to be oriented. Once oriented, an isomorphism $\star\colon \bigwedge_k(V) \to \bigwedge_{N-k}(V)$ called the *Hodge star* can be defined, which identifies k-vectors with $(N-k)$-vectors. This result holds only for spaces with a *symmetric* metric[13].

In $\mathbb{R}^3$, the Hodge star operator $\star$ identifies bivectors (oriented areas) with their normal vector, and trivectors (pseudoscalars) with their associated volume (i.e. their scalar value). This explains the similarity of Eqn. 16 for the bivector $\boldsymbol{v} \wedge \boldsymbol{u}$ to the cross product $\boldsymbol{v} \times \boldsymbol{u}$; the two are Hodge duals in $\mathbb{R}^3$. The correspondence is shown in Fig. 3. This explains why the cross product does not generalize to higher dimensional metric spaces: the Hodge duality no longer pairs vectors to areas. Hodge duality in $\mathbb{R}^3$ also pairs scalars with pseudoscalars.

To illustrate a lack of cross product in $\mathbb{R}^4$, the dual pairs on the Grassmann algebra are: scalars with pseudoscalars (4-vectors), bivectors with bivectors, and vectors with trivectors. There is no isomorphism between bivectors and vectors, so a cross product (a vector) doesn't make sense to represent oriented areas, while a bivector does so naturally. Note that scalars and N-vectors (pseudoscalars) are always Hodge dual, that is $\star 1 = \Omega_V$, and both represent oriented volumes in the exterior space of $\mathbb{R}^N$ (or any other vector space).

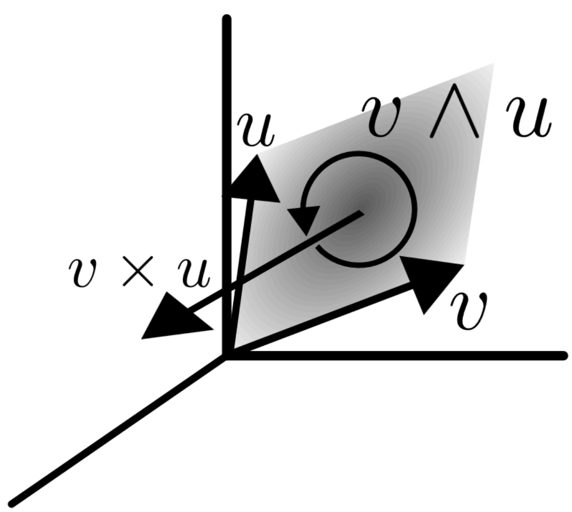

Figure 3: Bivectors and cross products in $\mathbb{R}^3$ are Hodge duals, i.e. $\star(\boldsymbol{v} \wedge \boldsymbol{u}) = \boldsymbol{v} \times \boldsymbol{u}$.

[13] Generally speaking, spaces with a symmetric metric g measure lengths (hence the moniker metric spaces). Spaces with an antisymmetric metric fundamentally measure ares (called symplectic spaces).

3 Theory of vector fields

Let $M \subset \mathbb{R}^n$ be an open subset of the Euclidean space $\mathbb{R}^n$, i.e. a smooth flat manifold. The *tangent space* T_xM to $x \in M$ is a local copy of the vector space $\mathbb{R}^n$ with origin x. The collection of all tangent spaces $TM = \cup_{x\in M} T_xM = M \times \mathbb{R}^n$ is called the *tangent bundle* and is $2n$-dimensional. Consider a smooth function $f\colon m \to \mathbb{R}$, and its directional derivative D_vf at $x \in M$ towards $v \in T_xM$,

$$D_vf = \lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} = \frac{d}{dt}\Big|_{t=0} f(x+tv) \tag{23}$$

$$= \sum_{i=1}^{n} \frac{d(x^i + tv^i)}{dt}\Big|_{t=0} \frac{\partial f}{\partial x^i}(x) = \sum_{i=1}^{n} v^i(x) \frac{\partial f}{\partial x^i}(x). \tag{24}$$

In other words, the directional derivative acts as a functional $D_v\colon C_x^\infty \to \mathbb{R}$ given by

$$D_v = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}\Big|_x. \tag{25}$$

The directional derivative is a type of derivative, which in general are called *derivations*. Just like functions, derivations may be described as members of a vector space, and it may be shown that the tangent vector space T_xM is isomorphic to the space of derivations[T]. This means that a generic vector $v|_x \in T_xM$ can be represented precisely as in Eqn. 25,

$$v|_x = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i} \tag{26}$$

and the appropriate basis for vectors in a tangent space are the partial derivatives $\partial_i \equiv \frac{\partial}{\partial x^i}$. Then *vectors are themselves differential operators.* For $v|_x \in T_xM$,

$$v|_x f = v^i \partial_i f(x) = v^i \frac{\partial f}{\partial x^i}(x). \tag{27}$$

and the action of the vector components at $x \in M$ is to take a derivative of a function. This is the simplest example of a Lie derivative. Therefore *vector fields* are defined as, with $x \in M$ and $f(x) \in T_xM$, the pairs $X(x) = (x, f(x)) \in M \times T_xM$ with natural representation

$$X = \sum_{i=1}^{n} f^i \partial_i. \tag{28}$$

3.1 Cotangent spaces and differentials

Recall that the standard basis e_i of a vector space V and its duals $e^i \in V^*$ satisfied

$$e^i(e_j) = \delta^i_j \tag{29}$$

with members of the dual space called covectors. Now as the tangent space T_xM at $x \in M$ is a vector space it's natural to consider its dual space. This is the *cotangent space* T_x^*M, consisting of those covectors or linear functionals $f\colon T_xM \to \mathbb{R}$. The *cotangent bundle* $T^*M = \cup_{x\in M} T_x^*M$ collects all cotangent spaces over M[14].

[14]Strictly speaking the cotangent bundle is the dual bundle of the tangent bundle, defined through the duals of the fibers T_xM just as described here.

Just as vector fields X assign to each point $x \in M$ a vector $v \in T_xM$, a *differential form*[15] ω_1 assigns a covector $\omega_{1,x} \in T_x^*M$ to each $x \in M$. The natural representation for differential forms ω_1 is in terms of the differentials dx. Recalling that vectors act as differential operators, define the *differential* of $f\colon M \to \mathbb{R}$ as a covector $df \in T^*M$ whose members $df|_x \in T_x^*M$ act on vectors $v|_x \in T_xM$ as the action of the vector on the function,

$$df|_x(v|_x) \equiv v|_x f. \tag{30}$$

To understand this, suppose the function $f = x^i$ picks out the i'th coordinate. Then its action on the basis vectors of the tangent space $\partial_j|_x \in T_xM$ is

$$dx^i|_x(\partial_j|_x) = \partial_j x^i = \delta^i_j. \tag{31}$$

This reveals that *the natural basis of the cotangent space are the coordinate differentials* dx^i. For a general function $f\colon M \to \mathbb{R}$, its differential in this basis has the familiar form,

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i. \tag{32}$$

The simplest differential forms are such differentials, e.g. $\omega^1 = \sum_{i=1}^n \omega_i dx^i$, the 1-forms.

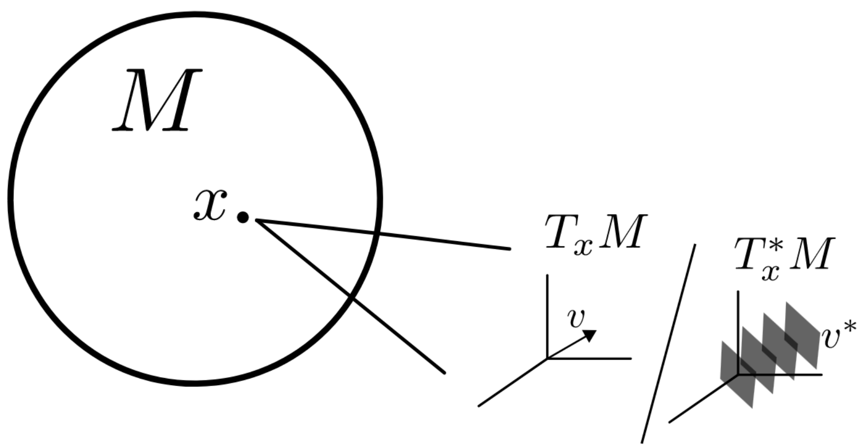

Figure 4: Associated to each point x in a smooth manifold M is the tangent space T_xM and the cotangent space T_x^*M, containing vectors $v \in T_xM$ and their duals $v^* \in T_x^*M$.

3.2 The connection of vector fields and differential equations

Imagine a streamline in a steady flow, parameterized as the smooth curve $\gamma\colon I \to M$ with parameter $t \in I$ in the interval $I \subset \mathbb{R}$. The streamline is parallel to the velocity field of the flow everywhere along the curve $\gamma(t)$, with the vector emanating from that point on the curve. This naturally lends itself to the tangent space description of Section 3.

[15]Occasionally, although rarely, called a covector field.

If such a curve has components $\gamma(t) = (\gamma^1(t), \ldots, \gamma^n(t))$, then its *velocity vector* $\frac{d\gamma}{dt}\big|_{\gamma(t)}$ lives in the tangent space $T_{\gamma(t)}U$ at $\gamma(t)$ and has components

$$\frac{d\gamma}{dt}\Big|_{\gamma(t)} = \Big(\frac{d\gamma^1}{dt}, \ldots, \frac{d\gamma^n}{dt}\Big) \in T_{\gamma(t)}U. \tag{33}$$

Now if X is a vector field on $M \subset \mathbb{R}^n$, then an *integral curve* $\gamma\colon I \to M$ of X satisfies

$$\frac{d\gamma}{dt}\Big|_{\gamma(t)} = X(\gamma(t)), \quad \forall t \in I. \tag{34}$$

For a vector field $X(x) = (x, f(x)) \in TM$, this gives the system of ODEs

$$\frac{d\gamma}{dt} = (f^1(\gamma), \ldots, f^n(\gamma)), \quad \gamma(0) = \gamma_0 \tag{35}$$

and an integral curve is a solution of these ODEs. Interpreting *integral curves as solutions of vector fields* establishes a correspondence between vector fields and differential equations. From the uniqueness theorem of ODEs, the curve $\gamma(t)$ is unique.[16]

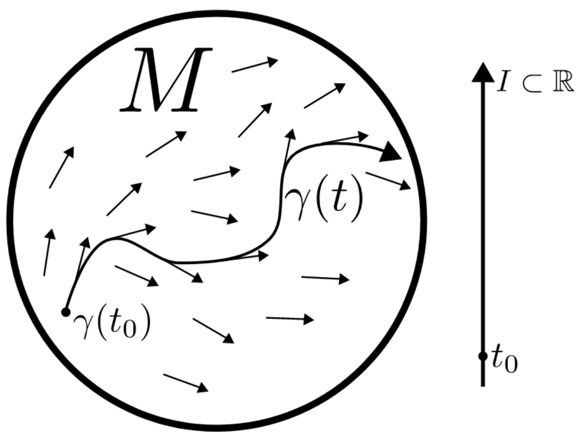

Figure 5: An integral curve $\gamma(t)$ emanates from a point $x_0 = \gamma(t_0) \in M$ and runs along a vector field X on the manifold M, parameterized by an interval $I \subset \mathbb{R}$ of the real line.

3.2.1 Examples of vector fields

Take $M \subset \mathbb{R}$. The following are simple vector fields and their integral curves $\gamma = x(t)$:

- $X = \partial_x$ are the translations, i.e. $\frac{dx}{dt} = 1 \implies x = x_0 + t$;
- $X = x\partial_x$ are the dilations, i.e. $\frac{dx}{dt} = x \implies x = e^t x_0$;
- $X = x^2\partial_x$ are the inversions, i.e. $\frac{dx}{dt} = x^2 \implies x = x_0(1 - x_0 t)^{-1}$.

[16] Of course its parametrization is usually not. For a standardized analysis one singles out the distinguished *maximal integral curves* $\gamma\colon I^x \to M$ corresponding to initial conditions $x \in M$ which use the largest possible interval $I^x \subset \mathbb{R}$. The vector field is called *complete* when the interval $I^x = \mathbb{R}, \forall x \in M$.

3.3 The flow of a vector field

The three examples of vector fields, their corresponding differential equations, and their solutions as transformations reveals the close connection of these concepts. The action of a vector field in generating integral curve solutions is called its flow, and this flow is identical to a local Lie group action. This identification connects the theory of vector fields to that of transformation groups, a useful notion in modern perturbation theory.

As above, the *flow* generated by a vector field $X \in TM$ is defined as the integral curve[17] $\Phi^X(\tau, x)$ with parameter $\tau \in I$ and initial point $x \in M$. As it is the solution of a system of ODEs, it has the properties of time composition $\Phi^X(\delta, \Phi^X(\tau, x)) = \Phi^X(\delta + \tau, x)$ and initial condition $\Phi^X(0, x) = x$. Of course it also satisfies Eqn. 34, which in the Φ notation is

$$\left.\frac{d\Phi^X}{d\tau}\right|_x = X\Phi^X(\tau, x). \tag{36}$$

Given Φ^X, the vector field X may be found in the neighborhood of $\tau = 0$, i.e.

$$\left.\frac{d}{d\tau}\right|_{\tau=0} \Phi^X(x, \tau) = X(x) \tag{37}$$

and for this reason the vector field is called the *infinitesimal generator* of X. The formal solution to Eqn. 36 is given by the operator exponential of the vector field X, recalling that X is in fact a differential operator. Such a solution is called *exponentiation* of the flow

$$\Phi^X(\tau, x) = \exp(\tau X)x. \tag{38}$$

If the system of ODEs is of the form $\frac{d\Phi}{d\tau} = A\Phi$ for A a real matrix, then the flow exponential reduces to the matrix exponential. Taylor expanding the flow exponential yields

$$\exp(\tau X)x^i = (1 + \tau X^i\partial_i + \mathcal{O}(\tau^2))x^i = x^i + \tau X^i + \mathcal{O}(\tau^2) \tag{39}$$

in local coordinates x^i, showing that the flow is locally a translation.

4 Differential forms

The tangent vector and differential one-form concepts are extended to multilinear differential forms through exterior powers much as in Section 2. The introduction of the rules for *exterior differentiation* and *integration of forms* allows for calculus with tensor fields and differential forms on a manifold. This extends the vector field operations such as the gradient and Stokes' theorem to quite general spaces. The concept is important in this work in order to study the *symplectic manifolds* of Hamiltonian mechanics.

As remarked following Eqn. 32, the coordinate differentials dx^i provide a local basis for the differential 1-forms $\phi = \phi_i dx^i$ in the cotangent bundle T^*M at some point $x \in M$, being locally dual to the vector field $X^i\partial_i$ defined on the tangent bundle TM. The higher differential forms are members of the exterior powers of the cotangent bundle $\bigwedge^k(T^*M)$. Just as a vector field is a section of the tangent bundle, i.e. a rule $X\colon TM \to M$ assigning a vector $v \in T_xM$ to each $x \in M$, a differential form of grade k assigns a differential form to each point in M by taking a section of the kth exterior power of the cotangent bundle. The local basis for $\bigwedge^k(T^*M)$ is composed of wedge products of the basic forms on $\bigwedge^1(T^*M)$, i.e. the coordinate differentials.

[17] Technically parametrized maximal integral curve.

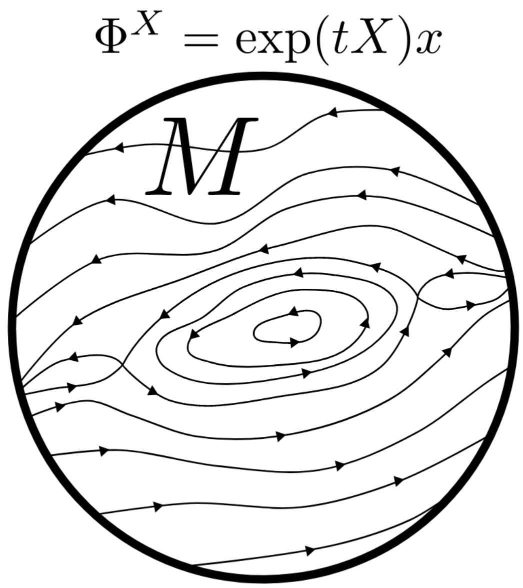

Figure 6: For a time-independent vector field X on a manifold M, the collection of its maximal integral curves is its flow, represented formally by the flow exponential $\exp(tX)$.

4.1 The differential 2-forms and higher forms

In the study of mechanics, the 2-forms are particularly important. Recall that in Eqn. 31 the differentials dx^i were defined through their action on a tangent vector. The basis 2-forms are defined through their action on a *pair* of tangent vectors,

$$(dx^i \wedge dx^j)|_x(A^k\partial_k|_x, B^l\partial_l|_x) = \begin{vmatrix} A^k dx^i(\partial_k) & A^k dx^j(\partial_k) \\ B^l dx^i(\partial_l) & B^l dx^j(\partial_l) \end{vmatrix} \tag{40}$$

$$= \begin{vmatrix} A^k\delta^i_k & A^k\delta^j_k \\ B^l\delta^i_l & B^l\delta^j_l \end{vmatrix} = \begin{vmatrix} A^i & A^j \\ B^i & B^j \end{vmatrix} = A^iB^j - A^jB^i \tag{41}$$

just as in Eqn. 10 for the algebraic 2-forms. The general form of a 2-form on M is

$$\omega^2 = \frac{1}{2}\omega_{ij}dx^i \wedge dx^j \equiv \frac{1}{2}\omega_{ij}dx^i dx^j \tag{42}$$

where the suggestive notation $dx^2dx^3 \equiv dx^2 \wedge dx^3$ is used for economy[18] and the components $\omega_{ij}\colon M \to \mathbb{R}$ are smooth functions on the manifold. Using Eqn. 41, the 2-form acts as

$$\omega^2(A^k\partial_k|_x, B^l\partial_l|_x) = \frac{1}{2}\omega_{ij}(x)(A^iB^j - A^jB^i)(x) \tag{43}$$

where the notation $f(x)$ is used to emphasize how the result depends on the point $x \in M$. The higher forms are defined in an analogous manner, so that a general k-form is

$$\phi^k = \frac{1}{k!}\phi_{i_1\cdots i_k}dx^{i_1}\cdots dx^{i_k} \tag{44}$$

where the $\phi_{i_1\cdots i_k}\colon M \to \mathbb{R}$ are smooth functions on M and $k!$ is for overcounting in the sum.

[18] It's simply to be remembered that $dx^3dx^2 = -dx^2dx^3$ as a differential 2-form and that $dx^3dx^3 = 0$ as the form returns zero area on any two vectors.

4.2 Exterior differentiation

The functions $f\colon M \to \mathbb{R}$ are 0-forms on M, and in Eqn. 32 it was pointed out that $df = \partial_i f dx^i$ was a 1-form on M. Its action on a tangent vector $df(v|_x) = \partial_i f v^j dx^i(\partial_j) = \partial_i f v^i$ was revealed to be the directional derivative[19] of f along the tangent vector v at $x \in M$. This operation, when extended to higher forms, is called the *exterior derivative* $d\colon \bigwedge^k(T^*M) \to \bigwedge^{k+1}(T^*M)$. The exterior derivative of a general k-form is

$$d\phi^k = \frac{1}{k!} d(\phi_{i_1 \cdots i_k}) dx^{i_1} \cdots dx^{i_k} \tag{45}$$

$$= \frac{1}{k!}(\partial_{i_1}\phi_{i_1\cdots i_k} dx^{i_1} + \cdots + \partial_{i_k}\phi_{i_1\cdots i_k} dx^{i_k}) dx^{i_1} \cdots dx^{i_k} \tag{46}$$

or simply the differential of the component functions times the basis wedges. Because $dx^{i_1}dx^{i_1} = 0$, this prescription generally simplifies. For example, if $\phi^1 = A(x,y)dx + B(x,y)dy$ is a 1-form on $\mathbb{R}^2$, its exterior derivative is the 2-form

$$d\phi^1 = dAdx + dBdy = (A_x dx + A_y dy)dx + (B_x dx + B_y dy)dy \tag{47}$$

$$= A_x dxdx + A_y dydx + B_x dxdy + B_y dydy \tag{48}$$

$$= (B_x - A_y)dxdy \equiv (B_x - A_y)dx \wedge dy. \tag{49}$$

The operation of exterior differentiation is an extension of the directional derivative to higher forms. In the above example, it gives the change of the form ϕ^1 along the oriented area spanned by two vectors[20]. If ϕ^k, ψ^k are k-forms and ϕ^l an l-form, exterior differentiation has the algebraic properties of:

$$d(\alpha\phi^k + \beta\psi^k) = \alpha d\phi^k + \beta d\psi^k, \quad \alpha, \beta \in \mathbb{R} \tag{50}$$

$$d(\phi^k \wedge \phi^l) = d\phi^k \wedge \phi^l + (-1)^k \phi^k \wedge d\phi^l \tag{51}$$

with the alternating Leibniz property following from that of the wedge product.

4.2.1 Closed and exact forms, potentials, and Poincaré's lemma

Notice that in the example of Eqn. 49, if the 1-form had been a total differential, that is if $\phi^1 = A_x dx + A_y dy = dA$, then $d\phi^1 = (A_{xy} - A_{yx})dxdy$. If A has continuous second partials, then the Clairaut-Schwarz theorem states $A_{xy} = A_{yx}$, so $d\phi^1 = d(dA) = 0$.

Poincaré's lemma[21] determines that this holds for any k-form ϕ^k, so that $d(d\phi^k) = 0$, or symbolically the exterior derivative satisfies $d^2 = 0$. Differential forms ϕ^k with $d\phi^k = 0$ are called *closed forms*, and further a form is called *exact*[22] if it has the property that $\phi^k = d\psi^{k-1}$ for some $(k-1)$-form ψ^{k-1}. Such an underlying form ψ such that $\phi = d\psi$ is called a *potential* for the form ϕ. Potentials are not unique, since any closed form Ψ may be added to it, $d(\psi + \Psi) = d\psi = \phi$, to yield the same differential form.

A direct corollary of the fact that $d^2 = 0$ is that every exact form is closed, or that if a form has a potential then it is closed. Now under what conditions does a closed form have a potential? In fact this depends on the topology of the manifold M. The *converse of Poincaré's lemma* states

[19] This operation is written $\boldsymbol{v} \cdot \nabla f$ in classic vector analysis.
[20] It is no coincidence that the result is the "curl" of the vector field corresponding to ϕ^1!
[21] Many sources state the converse as Poincaré's lemma, but apparently Élie Cartan wrote it this way.
[22] Notice the continuity of an exact form with that of an exact differential.

that closed k-forms ($k > 0$) on a manifold M are exact if M is a smoothly-contractible manifold[23]. In other words, closed differential forms are always exact on simply-connected domains (see Fig. 7). A classic example of a closed but not exact form is $d\theta$ in polar coordinates on $\mathbb{R}^2 - \{0\}$ (as its domain is punctured),

$$d\theta = \frac{ydx - xdy}{x^2 + y^2}. \tag{52}$$

Note that the failure of a form to be exact is typically associated with *singularities*.

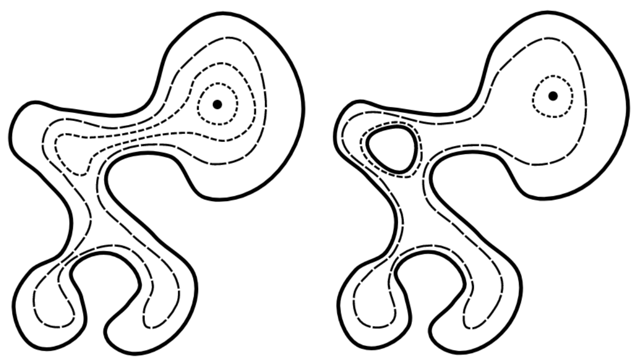

Figure 7: The domain on the left is smoothly contractible while that on the right is not. Closed forms on the left are necessarily exact, while those on the right are not necessarily exact. The contraction of the domain with a hole leaves behind a *residual*.

4.3 Contraction, or the interior product

It was seen in Section 2 that vectors (type $(1,0)$-tensors) and their duals (type $(0,1)$-tensors) naturally pair to produce scalars (type $(0,0)$-tensors) by evaluating the covector onto the vector. This is the simplest example of a *tensor contraction*. Extension to pairing vector fields X and differential k-forms $\phi^k \in \bigwedge^k(T^*M)$ occurs through their *interior product* $\imath_X\colon \bigwedge^k(T^*M) \to \bigwedge^{k-1}(T^*M)$, defined by partially evaluating the k-form on the field X,

$$\imath_X \phi^k(v_1, \cdots, v_{k-1}) = \phi^k(X, v_1, \cdots, v_{k-1}) \tag{53}$$

resulting in a $(k-1)$-form acting on $\{v_i\}_{i=1}^{k-1}$ other fields. Note that the partial evaluation occurs in the first slot. The interior product is an example of an *antiderivation* and so possesses similar properties as the exterior derivative,

$$\imath_X(\alpha\phi^k + \beta\psi^k) = \alpha\imath_X\phi^k + \beta\imath_X\psi^k, \quad \alpha, \beta \in \mathbb{R}, \tag{54}$$

$$\imath_X(\phi^k \wedge \phi^l) = \imath_X\phi^k \wedge \phi^l + (-1)^k\phi^k \wedge \imath_X\phi^l. \tag{55}$$

[23]The proof of this relies on an important concept in topology called *homotopy* (lit. "same space"). Two maps $f_0\colon M \to N$, $f_1\colon M \to N$ between manifolds M and N are *smoothly homotopic* if they can be deformed into one another by a smooth map $F\colon M \times I \to N$ with $I \subset \mathbb{R}$ (a "time") such that $F(x,0) = f_0$ and $F(x,1) = f_1$ for each $x \in M$. The *homotopy formula* for differential forms states that there exists a form $\psi(\phi)$ corresponding to homotopic maps f_1, f_0 such that $f_1^*(\phi) - f_0^*(\phi) = d(\psi(\phi)) + \psi d\phi$ (where $f^*(\phi)$ is the pull-back of ϕ by f, discussed shortly). If the identity map (say $f_1^*(\phi) = \phi$) is homotopic to one bringing all of M to a single point (say $f_0^*(\phi) = 0$), then $\phi = d\psi$ when $d\phi = 0$. In fact this method provides a formula to calculate a potential $\psi(\phi)$, so is not just of theoretical importance!

In the big picture of tensor contraction, the interior product is contraction of a type-$(1,0)$ tensor field with a type-$(0,k)$ antisymmetric cotensor field (differential form). Note that from the alternating property, $\imath_X \imath_X \phi^k = 0$ as a sort-of Poincaré's lemma for contractions.

4.4 Push-forwards and pull-backs

It must be determined how these operations of differentiation, interior product, etc. interact with a map[24] $f\colon M \to N$ between manifolds M and N. For example the *flow* on a manifold M, studied in Section 3, is a map from M to itself. The main idea is that vectors push-forward (covariance) while differential forms (including functions) pull-back (contravariance).

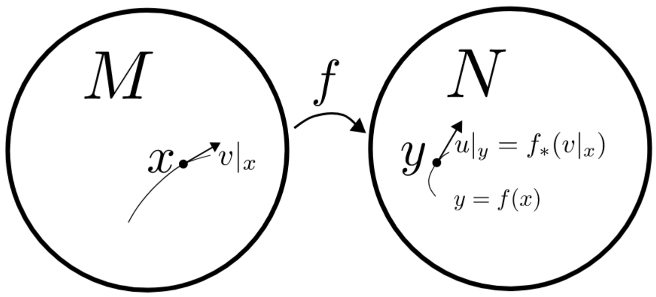

Figure 8: Under the action of a map $f\colon M \to N$, points and tangent vectors $(x, v|_x) \in M \times T_xM$ push-forward to $(f(x), f_*(v|_x)) \in N \times T_{f(x)}N$. In the case that f is a diffeomorphism (*i.e.* a smooth, invertible, topology-preserving transformation such as flows), then vector fields push forward in the very same way.

The more straight-forward[25] of these two notions is that of a *push-forward*. Clearly points $x \in M$ are "pushed forward" to points $f(x) \in N$. Now each tangent vector at $x \in M$ on the source manifold is tangent to some curve $\gamma(t)$, so that $v|_x = \gamma'(t_0)$ at $x = \gamma(t_0)$. The curve transforms to $f(\gamma(t)) \in N$, so the tangent vector at $f(x)$ is given by the chain rule,

$$u|_{f(x)} = (f(\gamma(t)))'|_{t=t_0} = f'(\gamma(t_0))\gamma'(t_0) = f'(x)v|_x. \tag{56}$$

Thus the transformation f on points *induces* one on tangent vectors called the push-forward, denoted by f_*. If the map is given by $y^j = f^j(x^i)$ in local coordinates, then Eqn. 56 implies that the tangent basis pushes-forward as $f_*(\partial_i) = \partial_i f^j \partial_j$, giving the push-forward of a tangent vector by the Jacobian matrix of f at $x \in M$,

$$f_*(v^i \partial_i) = v^i \partial_i f^j \partial_j. \tag{57}$$

To determine the transformation induced on differential forms by f, consider its relationship with the simplest forms: the functions. It is generally not well-defined to push functions forward[26], so instead functions $g\colon N \to \mathbb{R}$ are *pulled-back* to functions $f^*(g)\colon M \to \mathbb{R}$ through pre-composition, $f^*(g) = g(f)$. This agrees with the theme of duality throughout this study.

[24] For a completely proper treatment, a manifold should be described as a separable Hausdorff space with an atlas of compatible coordinate charts from open subsets $O \subset \mathbb{R}^n$ to neighborhoods of points $x \in M$. This is what is meant by a manifold being locally Euclidean. Then smooth maps $f\colon M \to N$ are those which smoothly operate on the coordinate charts. But that is far too much detail for our purposes.

[25] Pun not necessarily intended.

[26] For instance, if two points on M map to the same point on N, it's unclear which value this pushed-forward function should have

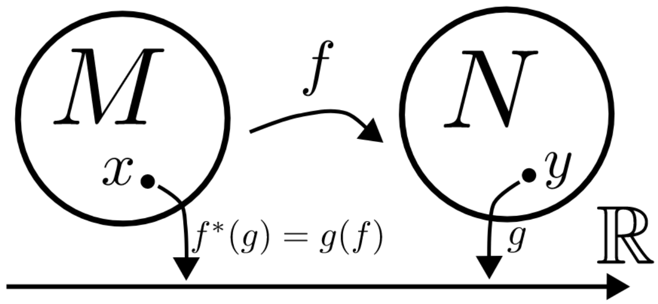

Figure 9: Functions $g\colon N \to \mathbb{R}$ on the codomain N pull-back to functions $f^*(g)\colon M \to \mathbb{R}$ on the domain M under mappings $f\colon M \to N$ through precomposition. In the diagram shown, the value of $f^*(g)$ on a point $x \in M$ is $f^*(g)(x) = g(f(x)) = g(y)$.

Note that pulled-back functions are evaluated on M, but their arguments are still in N, *i.e.* $f^*(g)(x) = g(f_*(x)) = g(f(x))$, so that the point x is pushed-forward to the point $f(x)$. For k-forms ϕ, their arguments are tuples of tangent vectors rather than points. Thus, in the very same way, to evaluate a pulled-back form on tangent vectors in M with precomposition, the argument must be pushed-forward to N to evaluate them. Hence, $f^*(\phi) = \phi(f_*)$, or

$$f^*(\phi(v_1, \cdots, v_k)) = \phi(f_*(v_1), \cdots, f_*(v_k)) \tag{58}$$

is the pull-back of a k-form $\phi \in \bigwedge^k(T^*N)$ to a k-form $f^*(\phi) \in \bigwedge^k(T^*M)$ induced by f.

Let us pull-back the dual basis dy^i on N to M under a map f by considering its action on a basis vector ∂_j on M. Take $f^*(dy^i)(\partial_j) = dy^i(f_*(\partial_j)) = dy^i(\partial_j f^k \partial_k) = \partial_j f^k dy^i(\partial_k) = \partial_j f^i$. Hence $f^*(dy^i) = \partial_j f^i dx^j$ where dx^j is the dual basis on M. Comparing to the result for push-forward, we have the basis transformation rules (that the reader may find familiar),

$$f_*\left(\frac{\partial}{\partial x^i}\right) = \partial_i f^j \frac{\partial}{\partial y^j}, \tag{59}$$

$$f^*(dy^i) = \partial_j f^i dx^j, \tag{60}$$

so that in local coordinates, the pull-back and push-forward are related through the *transpose* of the Jacobian matrix. Now in practice it's often easier to compute a pull-back by "substitution" (see example below) which is equivalent to Eqn. 59.

The operations of push-forward and pull-back are natural, so to speak, and so under a transformation they preserve geometric relationships by transforming everything directly. That is, the pull-back under f of forms ϕ, ψ and functions g, h satisfies the properties:

$$f^*(g\phi + h\psi) = f^*(g)f^*(\phi) + f^*(h)f^*(\psi), \tag{61}$$

$$f^*(\phi\psi) = f^*(\phi)f^*(\psi), \tag{62}$$

$$f^*(d\phi) = df^*(\phi). \tag{63}$$

4.5 Hodge duality with differential forms: circulation, flux, and beyond

As noted in Section 2, there is a dimension-dependent Hodge duality between certain pairs of geometric objects in the exterior algebra of a vector space, *e.g.* vectors and oriented areas in $\mathbb{R}^3$. This correspondence carries over to vector fields too. Let us use $\mathbb{R}^3$ for example. First, there is the clear duality of vector fields X and 1-forms ϕ. The idea of Hodge duality was that vectors (resp. 1-forms) were dual to bivectors (resp. 2-forms). Hence, the correspondence

$$X = A\partial_x + B\partial_y + C\partial_z, \tag{64}$$
$$\phi = Adx + Bdy + Cdz, \tag{65}$$
$$\omega = \star\phi = Adzdy + Bdzdx + Cdxdy \tag{66}$$

where ϕ is called the field's *circulation form* and ω its *flux form* (as integrals of them are these quantities). The reader is familiar with the ideas of gradient, divergence, and curl. In fact, these correspond to exterior derivatives of the corresponding forms. That is,

- If $f\colon \mathbb{R}^3 \to \mathbb{R}$ is a function, then $df \Leftrightarrow \nabla f$ (in duality of 1-form with vector field),
- If X is a vector field with circulation form ϕ, then $d\phi$ has dual $(\star d)\phi \Leftrightarrow \nabla \times X$,
- If X is a vector field with flux form $\omega = \star\phi$, then $d\omega = (d\star)\phi \Leftrightarrow \nabla \cdot X$.

In words: the gradient corresponds to the exterior derivative of a function; the curl corresponds to the area of the exterior derivative of the field's circulation form; the divergence corresponds to the volume of the exterior derivative of a field's flux form. Two more ingredients are needed for products (for $X \Leftrightarrow \phi$, $Y \Leftrightarrow \psi$):

- The dot product $X \cdot Y$ corresponds to the 0-form[27] $\star(\phi \wedge \star\psi) \equiv \star(\phi \star \psi)$,
- The cross product $X \times Y$ corresponds to the 1-form[28] $\star(\phi \wedge \psi) \equiv \star(\phi\psi)$.

From this unified point of view, it is easier to find identities. For example, as the Laplacian is $\nabla^2 f = (\star d)^2 f$, then (using that $\star(f \wedge \phi) = f \wedge (\star\phi)$ for 0-forms f and k-forms ϕ),

$$(\star d)^2(fg) = (\star d \star d)(fg) = (\star d\star)(gdf + fdg) \tag{67}$$
$$= (\star d)(g(\star df) + f(\star dg)) \tag{68}$$
$$= \star(g(d \star df) + dg(\star df) + f(d \star dg) + df(\star dg)) \tag{69}$$
$$= f(\star d)^2 g + g(\star d)^2 f + \star(df \star dg) + \star(dg \star df) \tag{70}$$
$$\Leftrightarrow \tag{71}$$
$$\implies \nabla^2(fg) = f\nabla^2 g + g\nabla^2 f + 2\nabla f \cdot \nabla g. \tag{72}$$

That example holds in $\mathbb{R}^n$ for $(\star d)^2 f \Leftrightarrow \nabla^2 f$ on functions. It should be noted that on 1-forms in $\mathbb{R}^3$, $(\star d)^2\phi \Leftrightarrow \nabla \times \nabla \times X \neq \nabla^2 X$. Now for an $\mathbb{R}^3$ specific example, consider $\nabla \cdot (X \times Y)$. In exterior calculus language, the divergence is the exterior derivative of the flux form of the vector field, and the flux form of $X \times Y$ is just $\phi \wedge \psi$. So differentiating the flux form,

$$\star d(\phi \wedge \psi) = \star(d\phi \wedge \psi) - \star(\phi \wedge d\psi) \tag{73}$$
$$= \star(\star(\star d\phi) \wedge \psi) - \star(\phi \wedge \star(\star d\psi)) \tag{74}$$
$$\Leftrightarrow \tag{75}$$
$$\nabla \cdot (X \times Y) = (\nabla \times X) \cdot Y - X \cdot (\nabla \times Y) \tag{76}$$

where it was used that $\star(\star\phi) = \phi$ for all forms in $\mathbb{R}^3$ (the sign may change in other spaces).

In higher dimensions $\mathbb{R}^n$ with $n \geqslant 3$, the 1-form corresponding to a field is Hodge dual to an $(n-1)$-form. While this allows a generalization of the divergence theorem to higher spaces, a generalized curl is trickier as there is no correspondence of bivectors with vectors.

[27] Note that the dot product is the volume spanned by one vector and another vector's dual area.
[28] As noted previously, the cross product is defined only in $\mathbb{R}^3$.

4.6 Stokes' theorem, or the generalized FTC

The differential forms formalism is put together with integration in mind, as forms measure weighted volumes spanned by tuples of tangent vectors. The great insight afforded by forms is that the various manifestations of Stokes' theorem (*e.g.* Green's theorem, the divergence theorem, *etc.*) are in fact generalizations of the fundamental theorem of calculus (FTC).

Integration on a manifold M of a form ϕ is defined by pulling the form back via a coordinate chart $f\colon R \to M$ to a reference domain $R \subset \mathbb{R}^n$ where it is a definite integral,

$$\int_M \phi \equiv \epsilon_M \int_R f^*(\phi). \tag{77}$$

The manifold M must have an orientation ϵ_M (simply the choice of a right-handed ($\epsilon_M = +1$) or left-handed ($\epsilon_M = -1$) basis on $T_x M$) to make sense of the alternating wedge product. If the manifold is Euclidean, $M \subset \mathbb{R}^n$, then integration of an n-form $\phi^n = \Phi(x)dx^1 \cdots dx^n$ is

$$\int_M \phi^n \equiv \epsilon_M \int \Phi(x)dx^1 \cdots dx^n \tag{78}$$

and may be computed as an iterated integral. Integration of k-forms with $k < n$ is done on k-dimensional submanifolds of M by pulling-back via their parameterization to the reference shape. For example, a line integral on some curve C parameterized as $\gamma(t)$ with $t \in I$ is

$$\int_C \phi^1 \equiv \epsilon_C \int_I \gamma^*(\phi^1) \tag{79}$$

where the orientation ϵ_C of the submanifold C is induced by that of the manifold ϵ_M. In the case of Euclidean manifolds, this is simply a formalization of material covered in a vector calculus course. The real benefit of the formalism is *generalized Stokes' theorem*,

$$\int_M d\phi = \int_{\partial M} \phi \tag{80}$$

that the integration of an exact form $\psi = d\phi$ on a manifold M is the integration of its potential ϕ on the manifold's boundary ∂M. From the discussed correspondence of vector fields and differential forms, this theorem underlies all the identities of vector calculus: Green's theorem, Stokes' theorem, and the divergence theorem. Intrinsically it is a generalization of the FTC, as for $I = [a, b]$ a 1-manifold and F an exact 0-form on it (a function),

$$\int_I dF = \int_{\partial I} F = F(b) - F(a) \tag{81}$$

as the interval's boundary is $\partial I = \{b\} \cup \{-a\}$. Similarly in $\mathbb{R}^3$, as $d\omega$ is the divergence of the flux-form ω of a vector field, applying GST to its integration within a volume V

$$\int_V d\omega = \int_{\partial V} \omega \qquad \Leftrightarrow \qquad \int (\nabla \cdot F)dV = \int F \cdot dS \tag{82}$$

reveals the divergence theorem. Similarly, as the 2-form $d\phi$ is the curl of a field with circulation 1-form ϕ, then applying GST to integration on a surface S shows

$$\int_S d\phi = \int_{\partial S} \phi \qquad \Leftrightarrow \qquad \int (\nabla \times F)dS = \oint F \cdot d\ell \tag{83}$$

relates the circulation of a field to its vorticity by the usual Stokes' theorem. Finally it's worthwhile to state *de Rham's theorem*, that for ϕ a k-form on an n-manifold M, if

$$\int_C \phi = 0 \tag{84}$$

for all k-cycles on M (technically closed k-chains, or intuitively closed k-dimensional submanifolds), then ϕ is an exact form. For example, analytic functions in $\mathbb{C}$ have $\oint_\gamma f dz = 0$ for all closed curves γ. This is because analytic functions are exact differentials.

4.7 Easy examples on (the manifold) $\mathbb{R}^3$ to build intuition

This example is intended to demonstrate that exterior differentiation, interior products, and pull-backs are nothing mysterious, and that the reader may in fact already intuit the rules. Consider the 1-forms

$$\phi = ydx - xdy \tag{85}$$
$$\psi = zdz, \tag{86}$$

on the first power of the cotangent bundle $\bigwedge^1(T^*\mathbb{R}^3)$. Their product $\omega = \phi\psi$ is the 2-form

$$\omega = \phi\psi = (ydx - xdy)zdz = -yzdz \wedge dx - xzdy \wedge dz. \tag{87}$$

The form ψ is closed, because $d\psi = d(zdz) = dz \wedge dz = 0$. A potential is $\Psi = \frac{1}{2}z^2$. On the other hand, ϕ is not closed as $d\phi = dy \wedge dx - dx \wedge dy = -2dx \wedge dy$. Now switching gears, consider the point $p = (1, 2, 3)$ and tangent vectors

$$X|_p = 4\partial_x + 3\partial_y + 2\partial_z, \tag{88}$$
$$Y|_p = 5\partial_x \tag{89}$$

at the point p. Then the action of the forms on the tangent vectors at this point are

$$\phi(X|_p) = (2)dx(X|_p) - (1)dy(X|_p) = 2(4) - 1(3) = 5, \tag{90}$$
$$\psi(Y|_p) = (3)dz(Y|_p) = 3(0) = 0, \tag{91}$$
$$\omega(X|_p, Y|_p) = -(2)(3)\begin{vmatrix} X^z & X^x \\ Y^z & Y^x \end{vmatrix} - (1)(3)\begin{vmatrix} X^y & X^z \\ Y^y & Y^z \end{vmatrix} \tag{92}$$
$$= -6\begin{vmatrix} 2 & 4 \\ 0 & 5 \end{vmatrix} - 3\begin{vmatrix} 3 & 2 \\ 0 & 0 \end{vmatrix} = -6(2(5) - 4(0)) - 0 = -60. \tag{93}$$

Consider a vector field $V = x\partial_x + y\partial_y + z\partial_z$ on $\mathbb{R}^3$. The interior product of ω with V is found by considering its action on another vector field $W = W^i\partial_i$,

$$\imath_V\omega(W) = xzdxdy(V, W) = xz\begin{vmatrix} V^x & V^y \\ W^x & W^y \end{vmatrix} = xz(V^xW^y - V^yW^x) \tag{94}$$
$$= xz(xW^y - yW^x) \tag{95}$$
$$\implies \imath_V\omega = x^2zdy - xyzdx. \tag{96}$$

Now consider the transformation $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by $f(u, v) = (u^2 + v^2, u, v)$, *i.e.*

$$x = u^2 + v^2 \quad \implies \quad dx = 2udu + 2vdv, \tag{97}$$
$$y = u, \quad \implies \quad dy = du, \tag{98}$$
$$z = v, \quad \implies \quad dz = dv. \tag{99}$$

Let's compute the pull-back of the forms on $\mathbb{R}^3$ and the push-forward of a tangent vector on $\mathbb{R}^2$. As advertised, pull-backs are done by substitution (equivalent to using the Jacobian),

$$f^*(\phi) = f^*(ydx - xdy) = f^*(y)f^*(dx) - f^*(x)f^*(dy), \tag{100}$$
$$= (u)(2udu + 2vdv) - (u^2 + v^2)(du), \tag{101}$$
$$= (u^2 - v^2)du + 2uvdv. \tag{102}$$

For the pull-back of the 2-form ω under f, using $du \wedge du = 0$,

$$f^*(\omega) = f^*(xzdxdy) = f^*(x)f^*(z)f^*(dx)f^*(dy), \tag{103}$$
$$= (u^2 + v^2)(v)(2udu + 2vdv)(du), \tag{104}$$
$$= -2v^2(u^2 + v^2)dudv. \tag{105}$$

Now consider a tangent vector $U|_q = 2\partial_u + 3\partial_v$ at a point $q = (1, 1)$ of $\mathbb{R}^2$. Using the Jacobian matrix, the tangent vector in $\mathbb{R}^2$ pushes forward under the transformation like

$$f_*(U|_q) = \begin{bmatrix} 2u & 2v \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}_{(1,1)} = \begin{bmatrix} 2 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \\ 3 \end{bmatrix}_{(2,1,1)} \tag{106}$$

to a tangent vector in $\mathbb{R}^3$ at the point $f(1, 1) = (2, 1, 1)$.

5 Lie theory

Recall that a *group* is a set G with: an associative operation; an identity; and an inverse for all elements $g \in G$. A *subgroup* S of G is a group with $S \subset G$ which is closed under the group operation. When the set G is a k-dimensional smooth manifold it is called a k-parameter *Lie group*. For a simple example, $G = \mathbb{R}$ is a one-parameter Lie group under addition.[29] A *Lie subgroup* is (roughly speaking) a submanifold of a Lie group[30].

5.1 Transformation groups

All manifolds are locally Euclidean, so the idea of Lie was to consider the linearization of the group as a manifold about a point. Linearization is formalized by the k-parameter *local Lie group*, one where the group manifold $V \subset \mathbb{R}^k$ is an open subset of Euclidean space $\mathbb{R}^k$, equipped with a group operation $m(u, v)$ for $u, v \in V$ which is associative, i.e. $m(u, m(v, w)) = m(m(u, v), w)$, and whose identity is the origin $0 \in V_0 \subset V$.

When each group element $u \in V$ is associated with a map from a manifold U to itself, the result is a *group of transformations* acting on U. The transformation map $\Psi \colon \mathcal{U} \to U$ is defined on a domain[31] $\mathcal{U} \subset V \times U$ and must have the group properties:

[29] In fact, all one-parameter Lie groups are isomorphic to either $\mathbb{R}$ (translations) or SO(2) (rotations).

[30] A wrinkle arises in that Lie subgroups are not necessarily regular submanifolds. For example, consider the set $H_\omega = \{(\tau, \omega\tau) \bmod 2\pi \colon \tau \in \mathbb{R}\} \subset T^2$ consisting of lines of slope ω through the torus T^2. It is a submanifold of T^2, and a one-parameter Lie subgroup of the toroidal group T^2. If ω is rational, then the lines eventually close on themselves and the subgroup is isomorphic to SO(2). If ω is irrational, then the lines densely fill T^2 and the subgroup is isomorphic to $\mathbb{R}$. The irrational case is not a regular submanifold.

[31] If $\mathcal{U}$ is an open subset of $V \times U$, the transformation group is called *local*. If $\mathcal{U} = V \times U$, it is *global*.

- associativity, for $(u, x) \in \mathcal{U}$ and $(v, \Psi(u, x)) \in \mathcal{U}$

$$\Psi(v, \Psi(u, x)) = \Psi(v \cdot u, x); \tag{107}$$

- identity, for some $i \in V$ and all $x \in U$, then $\Psi(i, x) = x$;
- inverse, each $(u, x) \in \mathcal{U}$ has $(u^{-1}, x) \in \mathcal{U}$ so that $\Psi(u^{-1}, \Psi(u, x)) = x$.

5.1.1 Examples of global transformation groups

- The Lie group $\mathbb{R}$ under addition acts on the manifold $\mathbb{R}^m$ to generate the translations

$$\Psi(t, x) = x + tv, \quad t \in \mathbb{R}, \quad v \in \mathbb{R}^m. \tag{108}$$

- The Lie group $\mathbb{R}^+$ under multiplication acts on $\mathbb{R}^m$ to generate scalings,

$$\Psi(\lambda, x) = (\lambda^{a_1} x^1, \ldots, \lambda^{a_m} x^m), \quad \lambda \in \mathbb{R}^+, \quad x \in \mathbb{R}^m \tag{109}$$

  for a set of scaling exponents a_i not all zero.

The Lie groups $\mathbb{R}$ (under addition) and $\mathbb{R}^+$ (under multiplication) are isomorphic, so the "same" Lie group can act on a manifold in different ways.

5.2 Correspondence of flow with the local group action of $\mathbb{R}$

Reexamining the properties of the flow Φ^X of a vector field $X \in TU$, it is seen to be *identical* to the local group action of $\mathbb{R}$ on the space $U \in \mathbb{R}^n$. Time-composition of the flow map is the associativity property, the initial condition is the identity, and time-reversal is the inverse. In view of the linearization Eqn. 39, the vector field X plays dual roles as generator of the flow and also infinitesimal generator of a one-parameter group of transformations.

5.3 Lie transformations and the Lie series

Often one wishes to know how functions and vector fields on a manifold evolve with flow. Having identified flow with a local Lie group acting on the manifold, such evolution is called a Lie transform. For $X \in TU$ a vector field on a manifold U and $f : U \to \mathbb{R}$ a smooth function on the manifold, the *Lie transform* of f generated by X is the transformation $f \mapsto \Phi_\tau^{*X} f$

$$\Phi_\tau^{*X} f(x) = f(\Phi_\tau^X(\tau, x)). \tag{110}$$

It is the pull-back of $f(x)$ by the flow Φ_τ^X generated by the vector field X. In words, it is the value of f at a time τ along the flow. Using the flow exponential notation, the Lie transform of $f(x)$ may be represented as $f(\exp(\tau X)x)$, which leads to a very useful representation for the Lie transform as a power series. Representing the vector field in local coordinates as $X = X^i \partial_i$, consider the first derivative

$$\begin{aligned} \frac{d}{d\tau} f(\exp(\tau X)x) &= \frac{d}{d\tau}[\exp(\tau X)x^i] \partial_i f|_{\exp(\tau X)x} && (111) \\ &= X^i \partial_i f|_{\exp(\tau X)x} && (112) \\ &= X(\exp(\tau X)x) f && (113) \end{aligned}$$

using the chain rule and the differential equation for the flow, Eqn. 36. The derivative is the action of the vector field X along the flow at time τ on the function f. At time $\tau = 0$,

$$\frac{d}{d\tau}\Big|_{\tau=0} f(\exp(\tau X)x) = X^i \partial_i f|_x = X(f)(x). \tag{114}$$

This is the Lie derivative of f at x along the field X. Taking another derivative at $\tau = 0$,

$$\frac{d^2}{d\tau^2} f(\exp(\tau X)x)\Big|_{\tau=0} = X(X(f))(x) \equiv X^2(f)(x). \tag{115}$$

Continuing and inserting into the Taylor expansion for $f(\exp(\tau X)x)$ gives the *Lie series*,

$$f(\exp(\tau X)x) = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} X^k(f)(x) \tag{116}$$

for the Lie transform of the function f. Equation 116 describes how a function transforms under flow, and in that sense provides a formal solution of flow (which can be solved in closed form only in special cases).

5.4 The Lie derivative: evolution under a flow

We have seen that a vector field on a manifold gives rise to a flow upon it. All geometric objects, *i.e.* vector fields, differential forms, multivector fields, etc. will evolve within a flow. The rate of change under this evolution is given by the *Lie derivative.* The simplest example has already been discussed, that functions evolve within a flow in terms of their differential. Notice that Eqn. 114 can be written as the flow pull-back. This defines its Lie derivative

$$\mathcal{L}_X f \equiv \frac{d}{d\tau}\Big|_{\tau=0} \Phi_\tau^{*X} f. \tag{117}$$

The Lie derivative of a differential k-form ϕ is analogously defined as

$$\mathcal{L}_X \phi \equiv \frac{d}{d\tau}\Big|_{\tau=0} \Phi_\tau^{*X} \phi \tag{118}$$

giving the instantaneous rate of change of differential forms (including functions, *i.e.* scalar fields) within a flow generated by X. In practice one uses *Cartan's "magic formula"*[32]

$$\mathcal{L}_X \phi = d(\imath_X \phi) + \imath_X(d\phi). \tag{119}$$

Considering Eqn. 118 as a differential equation for the pull-back under the flow, an exponential map in terms of the Lie derivative is obtained

$$\Phi_\tau^{*X} = \exp(\tau \mathcal{L}_X). \tag{120}$$

This relation is of great use in perturbation theory. For example, to understand the Lie derivative of a form, consider its action on a scalar field f (*i.e.* a function, or 0-form),

$$\mathcal{L}_X f = d(\imath_X f) + \imath_X(df) = 0 + \imath_X(\partial_i f dx^i) = X^i \partial_i f. \tag{121}$$

[32] Also called the homotopy formula, see footnote (ref).

Now consider the Lie derivative of a vector field. As discussed, tangent vectors push-forward rather than pull-back. Then the Lie derivative of a field $Y \in TM$ is given by

$$\mathcal{L}_X Y \equiv \frac{d}{d\tau}\Big|_{\tau=0} \Phi^X_{*,\tau} Y. \tag{122}$$

Suppose that both fields X and Y varied across the manifold, but the tangent vectors of Y remained aligned with the Y field when transported along the flow of the field X. In such a situation with $\mathcal{L}_X Y = 0$, it must be that their relative changes cancel out. In fact, direct calculation shows that the Lie derivative of a vector field along the flow of another is given by their commutator. The field $\mathcal{L}_X Y$'s action on a function f is then

$$\mathcal{L}_X Y(f) = [X, Y](f) = X(Y(f)) - Y(X(f)) \tag{123}$$

or in local coordinates, the resultant field is, for $X = X^i \partial_i$, $Y = Y^i \partial_i$, given by

$$[X, Y] = \sum_{i=1}^{n} \sum_{j=1}^{n} (X^j \partial_j Y^i - Y^j \partial_j X^i) \partial_i. \tag{124}$$

The commutator of two fields $[X, Y]$ is called their *Lie bracket*. The set of vector fields on a manifold which are related by the Lie bracket is called a *Lie algebra*.

6 Hamiltonian theory

Consider as manifold an n-dimensional configuration space Q (corresponding perhaps to the positions of n particles). The corresponding phase space Ω is the $2n$-dimensional cotangent bundle T^*Q of Q equipped with a symplectic form ω (*i.e.* a closed, nondegenerate differential two-form). This is the space for Hamiltonian mechanics. Since the cotangent bundle is even dimensional, this defines phase space as a symplectic manifold, a case of the Poisson manifolds. The coordinates of Ω are given by tuples (q^i, p_i) with $q^i \in Q$ and $p_i \in TQ$.

6.1 Symplectic manifolds and extended phase space

Symplectic manifolds are those even-dimensional manifolds with an associated closed, nondegenerate two-form called the *symplectic form*. Such a space fundamentally measures *areas* rather than *distances* as metric spaces do. The notion of a distance in phase space simply doesn't make sense dimensionally, as one is adding apples and oranges (lengths and momenta). However, adding *areas* makes sense, as one can add the product of apples and oranges to another product of apples and oranges. The symplectic form is given by

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq^i \equiv dp_i dq^i. \tag{125}$$

Its potential, called the *symplectic potential* is the generating form

$$\gamma = p_i(z) dq^i \tag{126}$$

for $z \in Q$ such that $d\gamma = \omega$. Now in the *extended phase space*, where time and energy are conjugate variables, the symplectic form is

$$\widetilde{\omega} = dp_i dq^i - dH dt \tag{127}$$

and its potential, the *fundamental Poincaré-Cartan form* is

$$\tilde{\gamma} = p_i dq^i - H dt, \tag{128}$$

also called the *phase space Lagrangian* in plasma physics, due to its similarity with the Lagrangian function $L = p_i \dot{q}^i - H$.

For two analytic functions f, g on Ω, their *Poisson bracket* is defined as

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}, \tag{129}$$

a skew-symmetric bilinear form which satisfies

$$\{f, gh\} = \{f, g\}h + \{f, h\}g, \tag{130}$$

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \tag{131}$$

The Poisson bracket is a type of Lie bracket, so that Ω with $\{\cdot, \cdot\}$ is in fact a Lie algebra. Given a function S, the *Hamiltonian vector field* X_S associated with S is

$$X_S = \{\cdot, S\} = \sum_{i=1}^{n} \frac{\partial S}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial S}{\partial q^i}\frac{\partial}{\partial p_i} \tag{132}$$

and its action on functions corresponds to the Lie derivative generated by S. The *Hamiltonian flow* of the field X_S is given by the flow exponential on a system point $z \in \Omega$,

$$\Phi_\tau^S(\tau, z) = \exp(\tau X_S)z = \sum_{k=0} \frac{\tau^k}{k!} X_S^k z. \tag{133}$$

As a formal solution, the flow generates a canonical transformation $z = Z(\tau)$ as

$$Z(\tau) = \exp(\tau X_S)z \tag{134}$$

and it follows that the inverse transformation is given by the field $X_{S'} = \{\cdot, S'\}$ with $S' = -S$. This is a special case of a property that for any function f analytic in Ω,

$$\exp(\tau X_S)f(z) = f(Z) \tag{135}$$

which is further explored in another note on Hamiltonian perturbation theory.

6.2 Simple example of the Lie transformation as a formal solution

Consider the harmonic oscillator with Hamiltonian

$$\mathcal{H} = \frac{1}{2}p^2 + \frac{1}{2}q^2. \tag{136}$$

The equations of motion and their corresponding Lie series are

$$\frac{dq}{dt} = X_{\mathcal{H}} q = p \quad \Longrightarrow \quad q = \exp(t\mathcal{H})q_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!} X_{\mathcal{H}}^k q_0, \tag{137}$$

$$\frac{dp}{dt} = X_{\mathcal{H}} p = -q \quad \Longrightarrow \quad p = \exp(t\mathcal{H})p_0 = \sum_{k=0}^{\infty} \frac{t^k}{k!} X_{\mathcal{H}}^k p_0. \tag{138}$$

Recalling that the notation $X_{\mathcal{H}}^2 = X_{\mathcal{H}} X_{\mathcal{H}}$ stands for repeated composition,

$$
\begin{aligned}
X_{\mathcal{H}} q &= \{q, \mathcal{H}\} = p, & X_{\mathcal{H}} p &= \{p, \mathcal{H}\} = -q, \\
X_{\mathcal{H}}^2 q &= \{\{q, \mathcal{H}\}, \mathcal{H}\} = \{p, \mathcal{H}\} = -q, & X_{\mathcal{H}}^2 p &= \{\{p, \mathcal{H}\}, \mathcal{H}\} = \{-q, \mathcal{H}\} = -p, \\
X_{\mathcal{H}}^3 q &= -p & X_{\mathcal{H}}^3 p &= q, \\
X_{\mathcal{H}}^4 q &= q, & X_{\mathcal{H}}^4 p &= p,
\end{aligned}
$$

the Lie series factors to give the expected solution

$$q = q_0 (1 - \frac{1}{2!} t^2 + \frac{1}{4!} t^4 + \cdots) + p_0 (t - \frac{1}{3!} t^3 + \cdots) \tag{139}$$

$$p = -q_0 (t - \frac{1}{3!} t^3 + \cdots) + p_0 (1 - \frac{1}{2!} t^2 + \frac{1}{4!} t^4 + \cdots) \tag{140}$$

$$\Longrightarrow \tag{141}$$

$$q = q_0 \cos t + p_0 \sin t, \tag{142}$$

$$p = -q_0 \cos t + p_0 \sin t, \tag{143}$$

In this case truncation of the Lie series is not a good approximation, yet this example reveals that *the formal solution is given through repeated Poisson bracket composition with the Hamiltonian.* In another sense, the Lie transformation

$$\exp(\phi X_{\mathcal{H}}) = \sum_{k=0}^{\infty} \frac{\phi^k}{k!} X_{\mathcal{H}}^k \tag{144}$$

can be said to generate a simple phase advance by ϕ for $\mathcal{H} = \frac{1}{2}(q^2 + p^2)$.