Cambridge University Press & Assessment
978-1-107-11388-6 — Optimal Control and Geometry: Integrable Systems
Velimir Jurdjevic
Excerpt
More Information

1

The Orbit Theorem and Lie determined systems

Let us begin with the basic concepts and notations required to set the text in motion starting from differentiable manifolds as the point of departure. Throughout the text, manifolds will be generally designated by the capital letters $M, N, O, \ldots$ and their points by the lower case letters $x, y, z, \ldots$. Unless otherwise stated, all manifolds will be finite dimensional, smooth, and second countable, that is, can be covered by countably many coordinate charts.

Local charts on M will be denoted by (U, ϕ) with U a coordinate neighborhood and ϕ a coordinate map on U. For each point $x \in U$ the coordinates $\phi(x)$ in $\mathbb{R}^n$ will be denoted by $(x_1, x_2, \ldots, x_n)$. For notational simplicity, we will often write $x = (x_1, \ldots, x_n)$, meaning that $(x_1, \ldots, x_n)$ is the coordinate representation of a point x in some coordinate chart (ϕ, U).

The set of smooth functions f defined on open subsets of M will be denoted by $C^\infty(M)$. It is a ring with respect to pointwise addition $(f+g)(x) = f(x) + g(x)$ and pointwise multiplication $(fg)(x) = f(x)g(x)$, both defined on the intersection of their domains.

On smooth manifolds tangent vectors v can be regarded both as the equivalence classes of curves and as the derivations. The first case corresponds to the notion of an "arrow": a tangent vector v at x is defined as the equivalence class of parametrized curves $\sigma(t)$ defined in some open interval I containing 0 such that $\sigma(0) = x$ with $\sigma_1 \sim \sigma_2$ if and only if in each coordinate chart (U, ϕ) $\frac{d}{dt}\phi \circ \sigma_1|_{t=0} = \frac{d}{dt}\phi \circ \sigma_2|_{t=0}$. We shall follow the usual custom and write

$$v = \frac{d}{dt}\phi \circ x(t)|_{t=0} = \left(\frac{dx_1}{dt}, \ldots, \frac{dx_n}{dt}\right),$$

where $x(t)$ is any representative in the equivalence class of curves that defines v.

In the second case, tangent vector are defined by their action on functions resulting in directional derivatives. As such, tangent vectors v at a point x are linear mappings from $C^\infty(M)$ into $\mathbb{R}$ that satisfy the Leibnitz formula,

$v(fg) = f(x)v(g) + g(x)v(f)$, for any functions f and g. In this context, tangent vectors in local coordinates will be written as $v = \sum_{i=1}^n v_i \frac{\partial}{\partial x_i}$, where $\frac{\partial}{\partial x_1}, \dots \frac{\partial}{\partial x_n}$ denotes the usual basis of tangent vectors defined by $\frac{\partial}{\partial x_i} f = \frac{\partial f}{\partial x_i}$, $f \in C^\infty(M)$.

These two notions of tangent vectors are reconciled through the pairing

$$v(f) = \frac{d}{dt} f \circ x(t)|_{t=0} \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}.$$

We will use T_xM to denote the tangent space at x, and use TM to denote the tangent bundle of M; TM is a smooth manifold whose dimension is equal to $2dim(M)$. Each coordinate chart (U, ϕ) in M induces a coordinate chart $(TU, \tilde{\phi})$ in TM with $\tilde{\phi}(v) = (x_1, \dots, x_n, v_1, \dots, v_n)$ for every $v \in T_x(U)$ and $x \in U$.

Recall that cotangent vectors at x are the equivalence classes of functions in $C^\infty(M)$ that vanish at x, with f and g in the same equivalence class if and only if f and g coincide in some open neighborhood of x and $\frac{d}{dt} f \circ \sigma(t)|_{t=0} = \frac{d}{dt} g \circ \sigma(t)|_{t=0}$ for every curve σ on M such that $\sigma(0) = x$. The pairing $\langle [f], [\sigma] \rangle = \frac{d}{dt} f \circ \sigma(t)|_{t=0}$ reflects the duality between tangent and cotangent vectors and identifies cotangent vectors as linear functions on $T_x(M)$.

In local coordinates, cotangent vectors will be written as the sums $\sum_{i=1}^n \frac{\partial f}{\partial x_i} dx^i$, where $dx_, \dots, dx_n$ denotes the dual basis relative to the basis $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}$. The cotangent space at x will be denoted by T_x^*M and the cotangent bundle $\cup\{T^*(M) : x \in M\}$ will be denoted by T^*M. The cotangent bundle is also a smooth manifold whose dimension is twice that of the underlying manifold M.

1.1 Vector fields and differential forms

Since these objects are fundamental for this study, it is essential to be precise about their meanings. Recall that a vector field X on M is a smooth mapping from M into TM such that $\pi \circ X = I$, where π denotes the natural projection from TM onto M. Thus $X(x)$ belongs to $T_x(M)$ for each point x in M. In the language of vector bundles, vector fields are sections of the tangent bundle. The space of smooth vector fields will be denoted by $V^\infty(M)$. In local coordinates vector fields X will be represented either by the arrow vector $(X_1(x_1, \dots, x_n), \dots, X_n(x_1, \dots, x_n))$, or by the expression

$$X(x_1, \dots, x_n) = \sum_{i=1}^n X_i(x_1, \dots, x_n) \frac{\partial}{\partial x_i},$$

depending on the context. The function $\sum_{i=1}^{n} X_i(x_1,\dots,x_n)\frac{\partial f}{\partial x_i}(x_1,\dots,x_n)$ will be denoted by Xf. This action of vector fields on functions identifies vector fields with derivations on M, that is, it identifies $V^\infty(M)$ with the linear mappings D on $C^\infty(M)$ that satisfy

$$D(fg)(x) = f(x)(D(g)(x)) + g(x)(D(f)(x)) \tag{1.1}$$

for all functions f and g in $C^\infty(M)$.

The space $V^\infty(M)$ has a rich mathematical structure. To begin with, it is a module over the ring of smooth functions under the operations:

(i) $((fX)g)(x) = f(x)(Xg)(x)$ for any function g and all x in M.
(ii) $(X+Y)f = Xf + Yf$ for all functions f.

Secondly, $V^\infty(M)$ is a Lie algebra under the addition defined by (ii) above and the Lie bracket $[X,Y] = Y\circ X - X\circ Y$, where $[X,Y]$ means that $[X,Y]f = Y(Xf) - X(Yf)$ for every $f \in C^\infty(M)$. The reader can easily show that in local coordinates $[X,Y]$ is given by $Z = \sum_{i=1}^{n} Z_i \frac{\partial}{\partial x_i}$ with

$$Z_i = \sum_{j=1}^{n} \frac{\partial X_i}{\partial x_j} Y_j - \frac{\partial Y_i}{\partial x_j} X_j. \tag{1.2}$$

There seems to be no established convention about the sign of the Lie bracket. In some books the Lie bracket is taken as the negative of the one defined above (for instance, [AM] or [Hl]).

Differential forms are geometric objects dual to vector fields. They are defined analogously, as the smooth mappings ω from M into T^*M such that $\pi\circ\omega = I$, where now π is the natural projection from T^*M onto M. In local coordinates, ω will be written as $\omega(x_1,\dots,x_n) = \sum_{i=1}^{n}\omega_i(x_1,\dots,x_n)dx_i$ for some some smooth functions $\omega_1,\dots,\omega_n$. Differential forms act on vector fields to produce functions $\omega(X)$ given by $\omega(X) = \sum_{i=1}^{n}\omega_i X_i$ in each chart (U,ϕ).

Differential forms are contained in the complex of exterior differential forms in which functions in $C^\infty(M)$ are considered as the forms of degree 0, and the differential forms defined above as the forms of degree 1. Differential forms of degree k can be defined in several ways [BT; Ar]. For our purposes, it will be convenient to define them through the action on vector fields. A differential form ω of degree k is any mapping $\omega : \underbrace{V^\infty(M)\times\cdots\times V^\infty(M)}_{k} \to C^\infty(M)$ that satisifies

$$\begin{aligned}&\omega_x(X_1,\dots X_{i-1}, fX_i + gW_i, X_{i+1},\dots,X_k)\\ &\quad = f\omega(X_1,\dots,X_n) + g(\omega(X_1,\dots,W_i,\dots X_k),\end{aligned}$$

for each $i \in \{1,\dots,k\}$ and each function f and g, and

$$\omega(X_1,\dots X_i,\dots,X_j,\dots,X_k)) = -\omega((X_1,\dots X_j,\dots,X_i,\dots,X_k)),$$

for each index i and j.

Forms of degree k will be denoted by $\Omega^k(M)$. It follows that forms of degree k are k-multilinear and skew-symmetric mappings over $V^\infty(M)$. The skew-symmetry property implies that $\Omega^k(M) = 0$, for $k > dim(M)$.

Alternatively, differential forms can be defined through the wedge products. The wedge product $\omega_1 \wedge \omega_2$ of 1-forms ω_1 and ω_2 is a 2-form defined by

$$(\omega_1 \wedge \omega_2)(X,Y) = \omega_1(X)\omega_2(Y) - \omega_1(Y)\omega_2(X).$$

Any 2-form ω van be expressed as a wedge product of 1-forms. To demonstrate, let $X = \sum_{i=1}^n X^i \frac{\partial}{\partial x_i}$ and $Y = \sum_{i=1}^n Y^i \frac{\partial}{\partial x_i}$. It follows that

$$\omega(X,Y) = \sum_{i,j}^n X^i Y^j \omega\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right) = \sum_{i>j}^n (X^i Y^j - X^j Y^i)\omega\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right).$$

But then $(dx_i \wedge dx_j)(X,Y) = \omega\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right)(X^i Y^j - X^i Y^j)$. Hence,

$$\omega = \sum_{i,j}^n \omega_{ij}(dx_i \wedge dx_j),$$

where ω_{ij} are the functions $\omega\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right)$.

We now come to another indispensable theoretic ingredient, the exterior derivative.

Definition 1.1 The exterior derivative d is a mapping from $\Omega^k(M)$ into $\Omega^{k+1}(M)$ defined by

$$\begin{aligned} df(X) = X(f),\ \text{when } k = 0, d\omega(X_1,\dots,X_{k+1}) \\ = \sum_{i=1}^{k+1}(-1)^{i+1} X_i \omega(X_1,\dots,\hat{X}_i,\dots,X_{k+1}) \\ - \sum_{i<j}(-1)^{i+j}\omega([X_i,X_j],\dots,\hat{X}_i,\dots,X_{k+1}) \end{aligned}$$

for $k > 0$, where the hat above an entry indicates the absence of that entry from the expression. For instance, $\omega(X_1,\hat{X}_2,X_3) = \omega(X_1,X_3)$, $\omega(X_1,X_2,\hat{X}_3) = \omega(X_1,X_2)$.

In particular, the exterior derivative of a 1-form ω is given by

$$d\omega(X_1,X_2) = X_1\omega(X_2) - X_2\omega(X_1) + \omega([X_1,X_2]). \tag{1.3}$$

To show the exterior derivative in more familiar terms [BT], let $X_1 = \frac{\partial}{\partial x_1}, \ldots, X_n = \frac{\partial}{\partial x_n}$ denote the standard basis relative to a system of coordinates $x_1, \ldots, x_n$. If f is a function, then $df = \sum_{i=1}^n \omega_i dx_i$ for some functions $\omega_1, \ldots, \omega_n$. It follows that $\omega_i = df(X_i) = X_i(f) = \frac{\partial f}{\partial x_i}$, and $df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$. Therefore, the exterior derivative of f coincides with the directional derivative.

Consider now the exterior derivative of a 1-form $\omega = \sum_{i=1}^n \omega_i(x) dx_i$. It follows that

$$d\omega(X_i, X_j) = X_i \omega(X_j) - X_j \omega(X_i) = \frac{\partial \omega_i}{\partial x_j} - \frac{\partial \omega_j}{\partial x_i},$$

since $[X_i, X_j] = 0$. Hence,

$$d\omega = \sum_{i=1, j=1}^{n} \left(\frac{\partial \omega_i}{\partial x_j} - \frac{\partial \omega_j}{\partial x_i} \right) dx_i \wedge dx_j.$$

The exterior derivative of a 2-form $\omega = \omega_1 dx_2 \wedge dx_3 + \omega_2 dx_3 \wedge dx_1 + \omega_3 dx_1 \wedge dx_2$ in $\mathbb{R}^3$ is given by

$$d\omega(X, Y, Z) = \left(\frac{\partial \omega_1}{\partial x_1} + \frac{\partial \omega_2}{\partial x_2} + \frac{\partial \omega_3}{\partial x_3} \right) (X \cdot (Y \wedge Z),$$

where $X \cdot (Y \wedge Z)$ denotes the signed volume defined by the vectors $X = (X_1, X_2, X_3)$, $Y = (Y_1, Y_2, Y_3)$ and $Z = (Z_1, Z_2, Z_3)$.

Differential forms in $\mathbb{R}^3$ are intercheangeably identified with vector fields via the following identification:

$$(w \in \Omega^1(\mathbb{R}^3) \Longleftrightarrow W \in V^\infty(\mathbb{R}^3)) \Longleftrightarrow w(X) = (W \cdot X), X \in V^\infty(\mathbb{R}^3).$$

In particular, if $w = df$ then the corresponding vector field is called the gradient of f and is usually denoted by $grad(f)$.

The expressions $\left(\frac{\partial \omega_2}{\partial x_3} - \frac{\partial \omega_3}{\partial x_2}, \frac{\partial \omega_3}{\partial x_1} - \frac{\partial \omega_1}{\partial x_3}, \frac{\partial \omega_1}{\partial x_2} - \frac{\partial \omega_2}{\partial x_1}\right)$ and $\left(\frac{\partial \omega_1}{\partial x_1} + \frac{\partial \omega_2}{\partial x_2} + \frac{\partial \omega_3}{\partial x_3}\right)$ are known as the curl and the divergence of a vector field $\omega_1 \frac{\partial}{\partial x_1} + \omega_2 \frac{\partial}{\partial x_2} + \omega_3 \frac{\partial}{\partial x_3}$. In this terminology, $d^2 = 0$ coincides with the well-known formulas of vector calculus

$$curl(grad) = 0 \quad \text{and} \quad div(curl) = 0.$$

Differential forms ω for which the exterior derivative is equal to 0 are called closed. The forms ω for which $\omega = d\gamma$ for some form γ are called exact. It can be shown that $d^2 = 0$, therefore exact forms are automatically closed. The quotient of k-closed forms over the exact k-forms is called the kth de Rham cohomology of M.

1.2 Flows and diffeomorphisms

Let us now consider differential equations

$$\frac{d\sigma}{dt}(t) = X(\sigma(t)), \tag{1.4}$$

defined by a vector field X in a manifold M.

Solution curves $\sigma(t)$ are called integral curves of X. In local coordinates, integral curves are the solutions of a system of ordinary differential equations

$$\frac{d\sigma_i}{dt}(t) = X_i(\sigma_1(t), \ldots, \sigma_n(t)), i = 1, \ldots, n.$$

It then follows from the basic theory of differential equations that for each initial point x there exists integral curves $\sigma(t)$ of X defined on an open interval $I = (-\epsilon, \epsilon)$ such that $\sigma(0) = x$. Any such curve can be extended to a maximal open interval $I_x = (e^-(x), e^+(x))$ whose end points are called the negative and the positive escape time. All integral curves of X that pass through a common point x have the same negative and positive escape time. The solution curve $\sigma(t), t \in (e^-(x), e^+(x))$ is called the integral curve of X through x.

Definition 1.2 Let X be a vector field and let $\Delta = \{(x,t) : x \in M, t \in (e^-(x), e^+(x))\}$. The mapping $\phi : \Delta \to M$ defined by $\phi(x,t) = \sigma(t)$ will be called the flow, or a dynamical system induced by X.

The theory of ordinary differential equations concerning the existence and uniqueness of solutions and their smooth dependence on the initial conditions can be summarized by the following essential properties:

1. $\phi(x, 0) = x$ for each $x \in M$.
2. $\phi(x, s+t) = \phi(\phi(x,s), t) = \phi(\phi(x,t), s)$ for all (x,s), (x,t) and $(x, s+t)$ in Δ.
3. ϕ is smooth.
4. $\frac{\partial}{\partial t}\phi(x,t) = X \circ \phi(x,t)$.

Conversely, any smooth mapping $\phi : \Delta \to M$ with Δ an open subset of $M \times \mathbb{R}$, and a neighborhood of $M \times \{0\}$ that satisfies properties (1), (2), (3), necessarily satisfies property (4) with $X = \frac{\partial \phi}{\partial t}(x,t)|_{t=0}$.

Vector field X is called the infinitesimal generator of the flow. The set $\{\phi(x,t) : t \in \mathbb{R}\}$ is called the trajectory through x, or the motion through x.

Definition 1.3 A mapping F from a manifold M onto a manifold N is called a diffeomorphism if F is invertible with both F and its inverse smooth. Manifolds M and N are said to be diffeomorphic if there is a diffeomorphism between them.

Flows of vector fields induce diffeomorphisms in the following sense. Let U be an open set in M whose closure is compact. Then there is an open interval $I = (-a, a)$ such that $U \times I$ is contained in Δ. The mapping $\Phi_t(x) = \phi(x, t)$ is a diffeomorphism from U onto $\Phi_t(U)$. Indeed, $\Phi_t^{-1} = \Phi_{-t}$.

A vector field X is said to be complete if $\Delta = M \times \mathbb{R}$, i.e., if each integral curve of X is defined for all $t \in \mathbb{R}$. Complete vector fields induce global flows $\phi : M \times \mathbb{R} \to M$. If X is a complete vector field then the corresponding family of diffeomorphisms $\{\Phi_t : t \in \mathbb{R}\}$ is called the one-parameter group of diffeomorphisms induced by X. Indeed, $\{\Phi_t : t \in \mathbb{R}\}$ is a group under the composition with $\Phi_t \circ \Phi_s = \Phi_{t+s}$ and $\Phi_t^{-1} = \Phi_{-t}$. If a vector field is not complete then its flow defines a local group of diffeomorphisms in some neighborhood of each point in M. It is known that all vector fields on compact manifolds are complete.

The shift in perspective from dynamical systems to groups of diffeomorphisms suggests another name for the trajectories. The trajectory through x becomes the orbit through x under the one-parameter group of diffeomorphisms. Each name evokes its own orientation, hence both will be used depending on the context.

1.2.1 Duality between points and linear functionals

The two ways of seeing vector fields, as arrows or as derivations, calls for further notational distinctions that elucidate the calculations with their flows. If X is a vector field then it is natural to write $X(q)$ for the induced tangent vector at q, seen as the arrow with its base at q and the direction $X(q)$, We will use a different notation for the same tangent vector seen as a derivation; $\hat{q} \circ X$ will denote the same tangent vector defined by $(\hat{q} \circ X)(f) = X(f)(q)$ for any function f, where now $\hat{q} : C^\infty(M) \to \mathbb{R}$ denotes the evaluation of f at q, i.e., $\hat{q}(f) = f(q)$.

It is easy to verify that for each point $q \in M$, $\hat{q}$ is a homomorphism from $C^\infty(M)$ into $\mathbb{R}$, the latter viewed as the ring under multiplication and addition, that is, $\hat{q}$ is a mapping from $C^\infty(M)$ into $\mathbb{R}$ that satisfies:

1. $\hat{q}(\alpha f + \beta g) = \alpha\hat{q}(f) + \beta\hat{q}(g)$ for all real numbers α and β and all functions f and g, and
2. $\hat{q}(fg) = \hat{q}(f)\hat{q}(g)$ for all functions f and g.

Conversely, for any non-trivial homomorphism ϕ from $C^\infty(M)$ into $\mathbb{R}$ there exists a unique point q in M such that $\hat{q} = \phi$ ([AS]). Therefore, the correspondence $q \to \hat{q}$ identifies points in M with linear functionals on $C^\infty(M)$ that satisfy (1) and (2) above.

The dualism between points and linear functionals carries over to diffeomorphisms. If F is any diffeomorphism on M then $\hat{F}$ will denote the pull back on functions in $C^\infty(M)$ defined by $\hat{F}(f) = f \circ F$. It follows that $\hat{F}$ is a ring automorphism on $C^\infty(M)$. Conversely, any ring automorphism on $C^\infty(M)$ is of the form $\hat{F}$ for some diffeomorphism F, as can be easily shown by the same proof as above.

We will now extend this notation to the flows Φ_t induced by vector fields on M, and let $\exp tX$ denote $\hat{\Phi}_t$ for the flow $\{\Phi_t : t \in R\}$ induced by X. It follows that

$$\exp(t+s)X = \exp tX \circ \exp sX = \exp sX \circ \exp tX, s, t \in R. \tag{1.5}$$

The above implies that $\exp tX|_{t=0} = I$ and $\exp -tX = (\exp tX)^{-1}$. Moreover,

$$\frac{d}{dt}\exp tX = X \circ \exp tX = \exp tX \circ X. \tag{1.6}$$

Let us now note an important fact that will be useful in the calculations below. Suppose that $\{\Phi_t : t \in \mathbb{R}\}$ is the flow induced by a vector field X. Then, $\{F \circ \Phi_t \circ F^{-1} : t \in \mathbb{R}\}$ is a one-parameter group of diffeomorphisms on M for any diffeomorphism F, and hence is generated by some vector field Y. It follows that $Y = F_* X \circ F^{-1}$, where $F_\star$ denotes the tangent map induced by F. Recall that $F_*(v) = w$, where $v = \frac{d\sigma}{dt}|_{t=0}, \sigma(0) = q$, and $w = \frac{d}{dt}(F(\sigma(t))|_{t=0}$. Since $F \circ \Phi_t \circ F^{-1}$ acts on points, $Y = F_* X \circ F^{-1}$ is the arrow representation of the infinitesimal generator of $\{F \circ \Phi_t \circ F^{-1} : t \in \mathbb{R}\}$.

As a derivation, $Y = \hat{F}^{-1} \circ \exp tX \circ \hat{F}$ by the following calculation:

$$Yf = \frac{d}{dt} f \circ (F \circ \Phi_t \circ F^{-1}) = X(\hat{F}(f)) \circ F^{-1} = (\hat{F}^{-1} \circ X \circ \hat{F})(f),$$

and therefore

$$\exp tY = \hat{F}^{-1} \circ \exp tX \circ \hat{F}. \tag{1.7}$$

Equation (1.6) yields the following asymptotic formula:

$$\exp tX \approx I + tX + \frac{t^2}{2}X^2 + \cdots + \frac{t^n}{n!} + \cdots . \tag{1.8}$$

Then,

$$\begin{aligned}
&\exp tY \circ \exp tX \circ \exp -tY \circ \exp -tX \\
&\quad = \left(I + tY + \frac{t^2}{2}Y^2 + \cdots\right) \circ \left(I + tX + \frac{t^2}{2}X^2 + \cdots\right) \\
&\qquad \circ \left(I - Yt + \frac{t^2}{2}Y^2 + \cdots\right) \circ \left(I - tX + \frac{t^2}{2}X^2 + \cdots\right)
\end{aligned}$$

Cambridge University Press & Assessment
978-1-107-11388-6 — Optimal Control and Geometry: Integrable Systems
Velimir Jurdjevic
Excerpt
More Information

$$= \left(I + t(X+Y) + \frac{t^2}{2}(X^2 + 2X\circ Y + Y^2) + \cdots\right)$$
$$\circ \left((I - t(X+Y) + \frac{t^2}{2}(X^2 + 2X\circ Y + Y^2) + \cdots\right)$$
$$= I + t^2(Y\circ X - X\circ Y) + \cdots,$$

which in turn yields an important formula

$$\frac{d}{dt}\exp\sqrt{t}Y \circ \exp\sqrt{t}X \circ \exp -\sqrt{t}Y \circ \exp -\sqrt{t}X|_{t=0} = [X,Y]. \tag{1.9}$$

Similar calculations show that the Lie bracket $[X,Y]$ can alternatively be defined by the formula

$$\frac{\partial^2}{\partial t\partial s}\exp -tX \circ \exp sY \circ \exp tX|_{t=s=0} = [X,Y]. \tag{1.10}$$

The preceeding formula can be seen in slightly more general terms according to the following definitions.

Definition 1.4 If X is a vector field then $adX : V^\infty(M) \to V^\infty(M)$ denotes the mapping $adX(Y) = [X,Y], Y \in V^\infty(M)$. If F is a diffeomorphism on M, then $Ad_F : V^\infty(M) \to V^\infty(M)$ is defined as

$$Ad_F(X) = \hat{F}^{-1} \circ X \circ \hat{F}$$

for all X in $V^\infty(M)$.

It then follows from (1.10) that

$$\frac{d}{dt}Ad_{\exp tX} = Ad_{\exp tX} \circ adX = adX \circ Ad_{\exp tX}. \tag{1.11}$$

1.3 Orbits of families of vector fields: the Orbit theorem

It is well known that each orbit of a one-parameter group of diffeomorphisms $\{\Phi_t\}$ generated by a vector field X is a submanifold of the ambient manifold M; the orbit through a critical point of X is zero dimensional, otherwise an orbit is one dimensional. These orbits are often referred to as the leafs of X, in which case M is said to be foliated by the leaves of X. There are two pertinent observations about these orbits that are relevant for the text below:

1. The orbits are not of the same dimension whenever X has critical points.
2. It may happen that an orbit is an immersed rather than an embedded submanifold of M. Recall that a submanifold is called embedded if its

topology coincides with the relative topology induced by the topology of the ambient manifold. For immersed submanifolds, all relatively open sets are open in the submanifold topology, but there may be other open sets which are not in this class.

For instance, each orbit of the flow $\Phi_t(z,w) = \{z\exp t\theta, w\exp t\phi, z\in\mathbb{C}, w\in\mathbb{C}, |z|^2 = 1, |w|^2 = 1\}$ is dense on the torus $T^2 = \{z\in\mathbb{C} : |z|^2 = 1\}\times\{w\in\mathbb{C} : |w|^2 = 1\}$ whenever the ratio $\frac{\theta}{\phi}$ is irrational. The sets $\{\Phi_t(z,w) : t\in(t_0,t_1)\}$ are open in the orbit topology, but are not equal to the intersections of open sets in T^2 with the orbit.

Remarkably, the manifold structure of orbits generated by one vector field extends to arbitrary families of vector fields, and that is the content of the Orbit theorem. To be more precise, let $\mathcal{F}$ be an arbitrary family of vector fields (finite or infinite) which, for simplicity of exposition only, will be assumed to consist of complete vector fields.

For each $X\in\mathcal{F}$, Φ_t^X will denote the one-parameter group of diffeomorphisms on M generated by X, and $G(\mathcal{F})$ will denote the group of diffeomorphisms generated by $\cup\{\Phi_t^X : X\in\mathcal{F}, t\in\mathbb{R}\}$. A typical element in $G(\mathcal{F})$ is of the form

$$g = \Phi_{t_p}^{X_p}\circ\Phi_{t_{p-1}}^{X_{p-1}}\circ\cdots\Phi_{t_1}^{X_1} \tag{1.12}$$

for a subset $\{X_1,\ldots,X_p\}$ of $\mathcal{F}$ and some numbers $t_1,t_2,\ldots,t_p$, or

$$\hat{g} = \exp t_1X_1\circ\exp t_2X_2\circ\exp t_{p-1}X_{p-1}\circ\exp t_pX_p. \tag{1.13}$$

Definition 1.5 The set $\{g(x) : g\in G(\mathcal{F})\}$ will be called the orbit of $\mathcal{F}$ through x and will be denoted by $\mathcal{O}_{\mathcal{F}}(x)$.

The orbit of $\mathcal{F}$ through a point x can be defined analogously on the space of functions as the set of automorphisms ϕ on $C^\infty(M)$ of the form

$$\phi = \hat{x}\circ\exp t_1X_1\circ\exp t_2X_2\circ\exp t_{p-1}X_{p-1}\circ\exp t_pX_p.$$

with the understanding that the automorphism $\hat{x}\circ F$ is identified with the point $F(x)$.

Proposition 1.6 The Orbit theorem *Each orbit $\mathcal{O}_{\mathcal{F}}(x)$ is a connected (possibly immersed) submanifold of M.*

This theorem, well known in the control community, has not yet found its proper place in the literature on geometry and therefore may not be so familiar to the general reader. Partly for that reason, but mostly because of the importance for the subsequent applications, we will outline the most important