Chapter 2
The Language of Geometry and Dynamical Systems: The Linearity Paradigm

La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi agli occhi (io dico l'universo), ma non si può intendere, se prima non s'impara a intender la lingua, e conoscer i caratteri ne'quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi ed altre figure geometriche, senza i quali mezzi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto.
Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth.

Galileo Galilei Il Saggiatore

2.1 Introduction

We can infer from the examples given in Chap. 1 that linear dynamical systems are interesting on their own. Moreover they can be explicitly integrated providing therefore a laboratory to explore new ideas and methods. We will use them systematically to illustrate all new notions and ideas to be introduced in this book.

We begin by elaborating more systematically the elementary, i.e., algebraic theory, for finite-dimensional linear dynamical systems, whose discussion was only initiated in the previous chapter. Later on, we will see how these algebraic ideas can be better expressed using a combination of geometry and analysis, that is, differential geometry.

We will use our experience with linear systems to build the foundations of differential geometry on vector spaces and from there to move to more general carrier spaces. This simple relation between linear algebra and elementary differential geometry is highlighted at the end of the chapter under the name of the 'easy' tensorialization principle, a simple idea that will prove to be very useful throughout the book.

J.F. Cariñena et al., *Geometry from Dynamics, Classical and Quantum*,
DOI 10.1007/978-94-017-9220-2_2

Thus in building a geometrical language for the theory of dynamical evolution we will start by discussing the notion of vector fields versus dynamical systems and the accompanying notions of tangent space, flows, forms, exterior algebra and the more abstract notion of derivations of the algebra of smooth functions on a linear space. This will be the content of Sects. 2.3 and 2.4.

Finally we will address in Sect. 2.5 the integration problem not just for a single vector field, but for a family of them, stating a simple form of the Frobenius theorem that provides necessary and sufficient conditions for the simultaneous integration of a family of vector fields. Then we will use a variation of this idea to solve the integration problem for Lie algebras, offering in this way a proof of Lie's third theorem and introducing in a more systematic way the theory of Lie groups.

2.2 Linear Dynamical Systems: The Algebraic Viewpoint

2.2.1 Linear Systems and Linear Spaces

A mathematical setting that embraces most examples in Chap. 1 is provided by a linear space E (also sometimes called a vector space) and a linear map $A\colon E \to E$ that helps us in defining the dynamics we are interested in. Later on we will discuss the extent to which these assumptions are reasonable so that these considerations will become an argument throughout all the book.

Thus we will consider a real linear space E, i.e. there are two binary operations defined on it: addition, denoted by $u + v$ for any $u, v \in E$, and multiplication by real numbers (scalars), denoted by λu for any $\lambda \in \mathbb{R}$. The linear map A satisfies $A(u + v) = Au + Av$ and $A(\lambda u) = \lambda Au$.

The interpretation of these objects regarding a given dynamics is that the vectors of the linear space E characterize our knowledge of the system at a given time t, thus we may think of vectors of E as describing partially the 'state' of the system we are studying (if the knowledge we can obtain from the system under scrutiny is maximal, i.e., no further information on the system can be obtained besides the one provided by the vectors u, then we would say that description provided by the vectors u is complete and they actually characterize the 'states' of the system). This physical interpretation is the reason that leads in many occasions to consider that there is a given linear structure in E, even though that is at this moment just an assumption we are introducing. In this setting, the trajectories $u(t)$ of our system, the data we can observe, are functions $t \mapsto u(t)$, i.e., curves on E. The parameter t used to describe the evolution of the system is associated to the 'observer' of the system, that is, to the people who are actually performing the experiments that allow description of the change on the vectors u, thus the parameter t has the meaning of a 'time' provided by an observer's clock.

We will assume first that the linear space is finite-dimensional with finite dimension n, thus the choice of any linear basis $\mathcal{B} = \{e_i \mid i = 1, \ldots, n\}$ allows us to identify it with $\mathbb{R}^n$ by means of $u \mapsto (u^i)$, $u = u^i e_i$, thus all notions of differential calculus of several variables can be transported to E by means of the previous identification. Hence if the trajectories are differentiable functions, they are characterized by the tangent vectors du/dt.

The expression for du/dt has to be determined by the experimental data, i.e., the observation of the actual trajectories of the system. If exhaustive experiments could be performed that will give us the value $F(u, t)$ of du/dt for all values of u at all possible times t, then we will have an expression of the form

$$\frac{du}{dt} = F(u, t),$$

that will be the mathematical law describing the dynamics of our system. However that is not the case, because performing such exhaustive measurements is impossible. Thus the determination of du/dt must be done combining experimental data and the ingenuity of the theoretician (at this point we should recall again Einstein's quote mentioned in the introduction (p. 12) of this book). On how to construct a vector field out of experimental data see [MS85].

The simplest situation we may envisage happens when the system has the form:

$$\frac{du}{dt} = A \cdot u. \tag{2.1}$$

(Let us recall that in Sect. 1.2.3 it was shown how it is possible by a simple manipulation to transform a system possessing an inhomogeneous term into a linear one, thus we will omit in what follows possible inhomogeneous terms in Eq. (2.1)). In other words, we assume that the tangent vector to the curve $u(t)$ depends just on the vector $u(t)$ at any time and does it linearly. Any description of a dynamics in the form given by Eq. (2.1) will be called a linear dynamical system or a linear system for short. Notice that the full description of a linear system involves the prescription of a space E with its linear structure $+, \cdot$ and a linear map A on it.

It is interesting to observe that linearity appears at two different levels here. On one side, we are assuming that the vectors describing states of the system can be composed and that the composition law we use for them satisfies the axioms of a linear space. On the other hand, we are considering that the infinitesimal variations to the trajectories of the states of the system, which are vectors by themselves, are tied to the linear structure of the states themselves by means of a linear map. Thus we are identifying a natural mathematical linear structure, that possessed by tangent vectors to curves, with a linear structure on the space of states which depends on the experimental setting we are preparing to describe the system under study. The nature of the exact relation between both structures will be the substance of much discussion in what follows. It will be enough to point out here that it cannot be assumed that the linear structure on the space of states will be uniquely determined in general (what can be thought to be a trivial statement) and that among all possible linear structures on such space, there could be some of them compatible with the dynamics we are trying to describe, in other words, a single dynamics can have various different descriptions as a linear system.

We must also notice here that most of the finite-dimensional examples discussed in the previous chapter were not linear systems. Actually none of the systems for which we have had a direct experience, like the free falling system (Sect. 1.2.4), the motion of celestial bodies, the Kepler system (Sect. 7.4.3), or other systems like the motion of charged particles in constant magnetic fields like the ones exhibited in picture 0.1 (Sect. 1.2.5) are linear systems (notice that we describe them using very specific parameters). Among these simple systems only the harmonic oscillator (Sect. 1.2.7) is a linear system. However all of them can be related to linear systems in different ways, either by direct manipulation or by the general theory of reduction as it will be done exhaustively in Chap. 7.

Contrary to systems in finite dimensions, all infinite-dimensional systems we discussed before, the Klein-Gordon field (Sect. 1.3.1), Maxwell equations (Sect. 1.3.2), and on top of all them, the Schrödinger equation (Sect. 1.3.3), were linear systems. The paramount role played by linearity in the analysis of infinite-dimensional systems has overcast its dynamical character and has taken over its physical interpretation. Without trying to elaborate a full physical analysis of them from a different perspective where linearity will not play a primitive role, we will emphasize later on some of the geometrical structures attached to them not depending on linearity.

2.2.2 Integrating Linear Systems: Linear Flows

Linear systems can be easily integrated. A linear system (2.1) defines a linear homogeneous (and autonomous) differential equation on E. Thus the problem we want to solve is to find differentiable curves $u(t)$ with t, defined on some interval I in $\mathbb{R}$, $t_0 \in I$ with values on E, such that

$$\frac{d}{dt}u(t) = A \cdot u(t), \quad \forall t \in I.$$

Such a curve will be called an integral curve or a (parametrized) solution of the linear system (2.1). We may think of this equation as an initial value problem if we select a vector $u_0 \in E$ and we look for solutions $u(t)$ that at time t_0 satisfies $u(t_0) = u_0$. The pair t_0, u_0 are called the Cauchy data of the initial value problem. The general existence and uniqueness theorem on the theory of differential equations guarantees that if E is a finite-dimensional linear space there exists a unique smooth curve $u(t)$, $t \in \mathbb{R}$ solving such a problem. However we will not need to use such a theorem as we will provide a direct constructive proof of this fact in Sect. 2.2.4. For the moment it will be enough for us to assume that such solutions exist and are unique.

Without loss of generality, see later, we can set $t_0 = 0$ in the present case. A natural way of thinking about solving this equation is to find a family of linear maps

$$\phi_t : E \to E, \quad t \in \mathbb{R}, \tag{2.2}$$

differentiable[1] on the variable t such that for any initial condition $u_0 \in E$ the curve $u(t) = \phi_t(u_0)$ is a solution of Eq. (2.1) with the given Cauchy data, namely,

$$\frac{d}{dt}\phi_t(u_0) = A(\phi_t(u_0)), \quad \phi_0(u_0) = u_0. \tag{2.3}$$

The family of maps $\{\phi_t \mid t \in \mathbb{R}\}$ will be called the flow of the linear system (2.1). Characteristic properties of a flow are:

$$\phi_t \circ \phi_s = \phi_{t+s}, \quad \phi_0 = I. \tag{2.4}$$

(with I the identity map on E). As an example, it is easy to check that the flow of the differential equation associated to the identity map $A = I$ is given by $\phi_t(x) = e^t x$. From the additive properties of the family of maps φ_t it is immediate to get that all of them are invertible and

$$\varphi_t^{-1} = \varphi_{-t}.$$

Both properties in (2.4) are immediate consequences of the uniqueness of solutions for the differential equation (2.1) (see for instance [Ar73] and [HS74] for a general discussion on the subject).

Exercise 2.1 Prove properties (a), (b) of Eq. (2.4).

Definition 2.1 Given a vector space E, a one-parameter flow (or just a flow) is a family of linear maps φ_t, $t \in \mathbb{R}$ depending smoothly on t and such that they satisfy properties (2.4) above.

In the case of E being finite-dimensional, it is easy to show from (2.4) that the smooth dependence on t of the family φ_t of linear maps defining a flow is equivalent to the much weaker property of the curve $t \mapsto \varphi_t(u)$ from $\mathbb{R}$ into E is continuous for all $u \in E$. In the later case we will say that the family φ_t is a strongly continuous one-parameter family of linear maps. In infinite dimensions the situation is much more subtle leading eventually to Stone's theorem that will be discussed in Chap. 6.

It is also evident from the linearity of Eq. (2.1) that the space of its solutions is a linear space. Thus if $u(t) = \phi_t \cdot u_0$ is the solution with initial condition u_0 and $v(t) = \phi_t \cdot v_0$ is the solution with initial data v_0, then, for any real number λ, $\lambda\, u(t)$ is the solution with initial data $\lambda\, u_0$ and $u(t) + v(t)$ is the solution with initial data

[1] 'Differentiable' here could be understood simply as the statement that the maps $t \mapsto \phi_t e_i$, with $\{e_i\}$ a basis in E, are differentiable.

$u_0 + v_0$. It follows from Eq. (2.4) that $\phi_{-t} = \phi_t^{-1}$, and it is also easy to show that the transformation sending $t \mapsto -t$ sends solutions of the linear equation given by A into solutions of the linear equation given by $-A$, because

$$\frac{d}{dt}(\phi_{-t} \cdot u_0) = -\frac{d}{ds}(\phi_s \cdot x_0)|_{s=-t} = -A \cdot (\phi_s u_0)|_{s=-t} = -A \cdot (\phi_{-t} \cdot u_0).$$

We have just proved that

Proposition 2.2 *Let E be a finite-dimensional space and A a linear map on it. Then the space of solutions $\mathcal{S}$ of the linear system $du/dt = A \cdot u$ is a linear space isomorphic to E. Such isomorphism is not canonical and a linear basis of the space of solutions is called a system of fundamental solutions.*

Proof A way to establish an isomorphism between $\mathcal{S}$ and E is by mapping $u(t)$ into $u(0)$. Clearly this map is invertible with inverse $u_0 \mapsto u(t) = \phi_t(u_0)$. □

For any initial condition $u_0 \in E$, Eq. (2.3) can be written as

$$\left(\frac{d\phi_t}{dt}\right)(u_0) = (A \circ \phi_t)(u_0)\,,$$

and therefore we can also set[2]

$$\frac{d\phi_t}{dt} = A \circ \phi_t, \tag{2.5}$$

or, equivalently,

$$\left(\frac{d\phi_t}{dt}\right) \cdot \phi_t^{-1} = \left(\frac{d\phi_t}{dt}\right) \cdot \phi_{-t} = A. \tag{2.6}$$

If we think of $t \mapsto \phi_t$ as a curve on the space of linear maps on E, fancily written as $\mathrm{End}(E)$, then Eq. (2.5) can be thought as an equation on $\mathrm{End}(E)$ whose solutions are the flow of the linear system we are looking for. In order to put such a problem as a genuine initial value problem we should provide it with the initial data ϕ_0, which we can take without losing generality as being the identity matrix I. Notice that because of the existence and uniqueness theorem for solutions of differential equations, the solution of this initial value problem exists and must coincide with the flow of the linear system $du/dt = A(u)$ described in (2.4). This means that the evolution described by Eq. (2.5) takes place actually in the subset $GL(E) \subset \mathrm{End}(E)$ of invertible linear maps on E. It is clear that the set $GL(E)$ has the algebraic structure

[2] The derivative of ϕ_t could be easily understood as thinking of ϕ_t as a curve of matrices, once we have selected any basis on E, then the space of $n \times n$ matrices can be identified with $\mathbb{R}^{n^2}$.

of a group, that is it carries its binary associative composition law given by the standard composition of linear maps, the identity linear map is the unit element and, finally, any element has an inverse (see later on Sect. 2.6.1 for a detailed description of the class of groups that will be of interest in this book). We call the initial value problem

$$d\phi_t/dt = A \circ \phi_t, \quad \phi_0 = I,$$

the group theoretical picture of the linear system (2.1). It is remarkable that the space of linear maps of E denoted before by $\mathrm{End}(E)$ carries a natural Lie algebra structure given by the standard commutator $[\cdot, \cdot]$ of linear maps, that is:

$$[A, B] = A \circ B - B \circ A. \tag{2.7}$$

(Let us recall that a Lie algebra structure on a linear space L is a skew symmetric bilinear composition law $[\cdot, \cdot] \colon L \times L \to L$ such that it satisfies Jacobi identity, i.e., $[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0, \forall a, b, c \in L$, see 10 for more details on the basic algebraic notions concerning Lie algebras and groups). In what follows we will start exploring the relation between the Lie algebra structure in the space $\mathrm{End}(E)$ and the group structure in the group of automorphisms of E, $GL(E)$.

From the form of our equation on $\mathrm{Aut}\,(E)$ it is clear that ϕ_t can be found easily by exponentiation, namely,[3]

$$\phi_t = \exp(tA) = e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k. \tag{2.8}$$

Indeed, since

$$\frac{d}{dt}\left(\sum_{n=0}^{\infty} \frac{t^n}{n!} A^n\right) = \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} A^n = A \sum_{m=0}^{\infty} \frac{t^m}{m!} A^m = A \exp(tA)\,, \tag{2.9}$$

we see that

$$\left(\frac{d}{dt} e^{tA}\right) \cdot e^{-tA} = A. \tag{2.10}$$

The operator A is called the infinitesimal generator of the one-parameter group $\phi_t = \exp\,(tA)$, $t \in \mathbb{R}$. We recall some useful properties of the exponential map exp defined above, Eq. (2.8):

[3] Notice that the operator-valued power series $\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ is convergent because it can be bounded by the numerical series $\sum_{k=0}^{\infty} \frac{t^k}{k!} ||A||^k$ with $|| \cdot ||$ any norm in the finite-dimensional linear space of linear maps on E.

1. If B is an isomorphism of E and A is in End (E), we have $B^{-1}A^nB = (B^{-1}AB)^n$, for any integer number n, and therefore,

$$e^{B^{-1}AB} = B^{-1}e^AB. \tag{2.11}$$

2. If $A, B \in \text{End}\,(E)$, and $A \circ B = B \circ A$, then, $e^{A+B} = e^Ae^B$.
3. Every exponential is a square, for $e^A = (e^{A/2})^2$. If E is a real, finite-dimensional vector space, then $\det e^A > 0$. It follows that the map $\exp\colon \text{End}(E) \to GL(E)$ is not surjective.
4. For a finite-dimensional vector space E,

$$\det e^A = e^{\text{Tr}\,A},$$

where det stays for the determinant and Tr for the trace.

The flow ϕ_t is a symmetry of the system for any value of t according with the notion of symmetry presented in Sect. 1.2.6. If $\phi_t(u_0)$ is a solution of Eq. (2.1) with initial condition u_0, then $(\phi_s \circ \phi_t)(u_0)$ is again a solution starting at $\phi_s(u_0)$. On the other hand is obvious that:

$$[\phi_t, A] = 0\,. \tag{2.12}$$

Having written $u(t) = e^{tA} \cdot u(0)$ as a solution of our equation $\dot{u} = A \cdot u$, we might believe that the problem of analyzing the linear system is solved. However that is not so. Indeed the series Eq. (2.8) defining e^{tA} may be unsuitable for computations or we could have questions on the dynamics that are not easy to answer from the expression before. For instance we may be interested in knowing if the solution $e^{tA}u(0)$ is periodic or, if $e^{tA}u(0)$ is bounded when $t \to \pm\infty$, and so on.

There are however some situations where most of the questions we may raise about the properties of the solutions have an easy answer:

1. A is diagonalizable. In this generic case E has a basis $\{e_j\}$ of eigenvectors of A with corresponding eigenvalues λ_j, i.e., $Ae_j = \lambda_j e_j$. It follows that $A^n e_j = (\lambda_j)^n e_j$ and therefore the set of curves

$$e_k(t) = e^{tA}\,(e_k) = e^{t\lambda_k}\,(e_k) \tag{2.13}$$

forms a fundamental system of solutions. Hence any solution $u(t)$ will have the form $u(t) = c^k e_k(t)$ (sum over k) and c^k determined by the initial conditions. However we must point it out that even if we know in advance that the operator A is diagonalizable, for instance if it is symmetric, solving the eigenvalue problem is often a difficult problem.
2. The endomorphism A is nilpotent. We recall that $A \in \text{End}\,(E)$ is nilpotent with index n if n is the smallest positive integer for which $A^n = 0$. In this case e^{tA} reduces to a polynomial and the general solution will have the form:

$$u(t) = \left(1 + tA + \cdots + \frac{(tA)^{n-1}}{(n-1)!}\right) u(0). \tag{2.14}$$

The general solution $\exp(tA)u(0)$ will be a vector with components which are polynomials in t of degree less than n.

As a particular instance, if we consider E to be the set of real polynomials in x of degree less than n, then E is a vector space of dimension n on which the operator d/dx is nilpotent of index n. On the basis $\{1, x, x^2, \ldots, x^{n-1}\}$, the linear operator d/dx is represented by the matrix:

$$\left(\frac{d}{dx}\right) = \begin{pmatrix} 0 & 1 & 0 & & 0 & 0 \\ 0 & 0 & 2 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & & n-2 & 0 \\ 0 & 0 & 0 & \cdots & 0 & n-1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

and

$$\exp(td/dx) = 1 + t\frac{d}{dx} + \cdots + \frac{t^{n-1}}{(n-1)!}\frac{d^{(n-1)}}{dx^{(n-1)}}.$$

If $p \in E$ we find that

$$\left[\exp(td/dx)p\right](x) = p(x) + tp'(x) + \cdots + \frac{t^{n-1}}{(n-1)!}p^{(n-1)}(x) = p(x+t).$$

Indeed the operator $\exp(td/dx)$ is the translation operator by $-t$. It is clear that the series expansion above would still make sense for any real analytic function on the real line, while the translation operator makes sense for any function, even if it is not differentiable.

3. The case of a general endomorphism A. A general endomorphism A can be decomposed in an essentially unique way as $A = D + N$ with D diagonalizable and N nilpotent (see later on the discussion of Jordan canonical forms) such that $[D, N] = 0$. Therefore the general situation will reduce to the preceding cases, but again a similar warning as in the diagonalizable case should be raised: finding D and N is often unpracticable.

2.2.2.1 Linear Changes of Coordinates

If we perform a linear change of coordinates $x \mapsto x' = Px$ in the vector space E where the linear system (2.1) is defined, clearly we obtain a new equation of motion with: $A \mapsto A' = PAP^{-1}$. That is clear because: $\dot{x}' = P\dot{x} = PA \cdot x =$

$PAP^{-1}Px = PAP^{-1} \cdot x'$. Of course all the elements of the family of equations of motion obtained in this way are equivalent under all respects and the corresponding flows are then related by $\phi'_t = P\phi_t P^{-1}$. In fact: $d\left(P\phi_t P^{-1}\right)/dt = P\,(d\phi_t/dt)\,P^{-1} = PA\phi_t P^{-1} = A'P\phi_t P^{-1}$.

From the active viewpoint, if $\Psi : E \to E$ is a linear isomorphism and the curve $x(t)$ is a solution of (2.1), then using the chain rule it is easy to see that the curve $(\Psi \circ x)(t)$ is a solution of,

$$\frac{d}{dt}(\Psi \cdot x) = (\Psi \circ A \circ \Psi^{-1})(\Psi \circ x) \tag{2.15}$$

A linear isomorphism Ψ will be a linear symmetry for our linear dynamical system if it maps solutions of the linear system into solutions of the same system. That means that if $x(t) = \exp(tA) \cdot x(0)$ is a solution, then also $(\Psi \circ x)(t) = (\Psi \circ \exp(tA)) \cdot x(0)$ is a solution of $\dot{x} = A(x)$. This clearly implies that $\Psi \circ A \circ \Psi^{-1} = A$. Conversely, if $\Psi \circ A \circ \Psi^{-1} = A$, then the linear map Ψ sends solutions into solutions. The set of symmetries obtained in this way is thus a group characterized as the subgroup G_A of $GL(E)$ of those isomorphisms Ψ such that $[A, \Psi] = 0$.

2.2.2.2 Symmetries

Given the evolution equation: $dx/dt = A \cdot x$, it is clear that we can construct new evolution equations as

$$\frac{dx}{ds_{(k)}} = A^k \cdot x, \tag{2.16}$$

for any positive integer k (we emphasize here that the evolution parameters are all taken to be independent. Of course: $s_{(1)} \equiv t$). It follows then quite easily that the flows $\Psi_k\left(s_{(k)}\right) = \exp\left\{s_{(k)}A^k\right\}$ are symmetries for the original equation and that, moreover, all these one-parameter groups pairwise commute.

Whenever the characteristic polynomial of A coincides with the minimal polynomial the powers of A yield all the infinitesimal symmetries of our equation, i.e., they generate the algebra of symmetries of A, which turns out then to be Abelian. That is the case when the eigenvalues of A are non-degenerate. If instead A has degenerate eigenvalues, then additional symmetries will be provided by all the elements of the general linear group acting on the corresponding eigenspace on which A acts as a multiple of the identity, and the symmetry group will be in general larger and no longer Abelian (see, however, below Sect. 3.6 for more details).

2.2.3 *Linear Systems and Complex Vector Spaces*

2.2.3.1 The Canonical Form of the Flow of a Linear Complex System

When the linear map A defining a linear system on the real vector space V is not diagonalizable (i.e., the representative matrix associated to it is not diagonalizable over the reals) it is convenient to use complex numbers so that the fundamental theorem of algebra is available. Let then E be a complex n-dimensional vector space (we will discuss its structure immediately after these remarks) and A a complex linear map on E. Then the characteristic polynomial $p(\lambda) = \det(A - \lambda I)$ of $A \in \mathrm{End}\,(E)$ factorizes as

$$p(\lambda) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_m)^{r_m}, \tag{2.17}$$

where the complex numbers $\lambda_1, \ldots, \lambda_m$, $m \leq n$ are the different eigenvalues of A, r_i is the algebraic multiplicity of the eigenvalue λ_i, $i = 1, \ldots, m$, and $r_1 + \cdots + r_m = n$. We consider the operators $(A - \lambda_k I)^{r_k}$ along with their kernels E_k. Each E_k has dimension r_k and $E = \bigoplus_k E_k$. This splitting leads to a decomposition of the identity by means of a family of projections P_i on each subspace E_i in such a way that $E_k = P_k E$, $\sum_{i=1}^m P_i = \mathrm{id}_E$ and $P_i P_j = \delta_{ij} P_i$. Then:

1. A leaves each E_k invariant, that is $A(E_k) \subset E_k$, therefore it makes sense to define $A_k = A P_k$. Note that A-invariance of E_k implies that $P_k A P_k = A P_k$. As the supplementary subspace $E_k' = \bigoplus_{l \neq k} E_l$ is also A-invariant, $P_k A = A P_k$.
2. $A = \sum_k A_k$, with $[A_k, A_j] = 0$, $\forall k, j$. that is a consequence of $A_k A_j = A P_k A P_j = A^2 P_k P_j = \delta_{kj} A^2 P_k$. Also $(A_k - \lambda_k I)^{r_k} = 0$.

We now find the flow of A decomposes as:

$$e^{tA} = e^{tA_1} \cdots e^{tA_m}. \tag{2.18}$$

If a vector $x \in E$ is decomposed as $x = \sum_k P_k x = \sum_k x_k$, we get $\exp(tA)x = \sum_k \exp(tA_k)x_k$. If we write e^{tA_k} as $e^{tA_k} = e^{t(A_k - \lambda_k I) + t\lambda_k I}$, and since the identity I commutes with everything, we have also

$$e^{tA_k} = e^{t(A_k - \lambda_k I)} e^{t\lambda_k I} ;$$

but now, $(A_k - \lambda_k I)$ is nilpotent of index not higher than r_k and we get

$$e^{tA} x = \sum_k e^{t\lambda_k} \mathcal{P}_k(t) x_k, \tag{2.19}$$

where $\mathcal{P}_k(t)$ is a polynomial in t with coefficients in $\mathrm{End}\,(E)$ and of degree less than r_k.

We have shown that the general solution of $\dot{x} = A \cdot x$ is the sum of vectors of the form $e^{t\lambda_k} x_k(t)$, with $x_k(t)$ a vector whose components are polynomials in t of degree less than the multiplicity r_k of the eigenvalues λ_k.

Remark 2.1 This splitting of A into the sum of commuting operators is the prototype of separability of differential equations. We will take back this theme when we discuss the notion of integrability of dynamical systems and we will relate this notion to the existence of normal forms (see Chap. 8).

Remark 2.2 We will come back to this discussion again when analyzing systems with compatible generic Hermitean structures (see Sect. 6.2.5).[4]

2.2.3.2 Complexification of Linear Dynamical Systems

It is time now to discuss with some care a structure that is going to play a relevant role in what follows: complex linear structures.

We recall that a complex linear space E is a vector space over the field of complex numbers $\mathbb{C}$. Any complex vector space has a basis. We will assume in what follows that E is finite-dimensional with complex dimension $\dim_{\mathbb{C}} E = n$. Let $\mathcal{B} = \{u_1, \ldots, u_n\}$ be a basis for E. Thus, any vector $v \in E$ can be written as $v = \sum_k z^k u_k$, with $z^k \in \mathbb{C}$. The map $\varphi_{\mathcal{B}} \colon E \to \mathbb{C}^n$ given by $\varphi_{\mathcal{B}}(v) = (z^1, \ldots, z^n)$ is an isomorphism of complex vector spaces where $\mathbb{C}^n$ is a complex vector space with the natural action of $\mathbb{C}$.

We can also think of E as a real vector space by considering the action of $\mathbb{R}$ on E given by $\lambda \cdot v = (\lambda + i\, 0)v$, $\lambda \in \mathbb{R}$. We shall denote this real vector space as $E_{\mathbb{R}}$ and we will call it the realification of E. It is clear that the vectors u and $i\, u$ are linearly independent on $E_{\mathbb{R}}$. Hence the set $\mathcal{B}_{\mathbb{R}} = \{u_1, \ldots, u_n, i\, u_1, \ldots, i\, u_n\}$ is a linear basis for $E_{\mathbb{R}}$. We will call this a real basis adapted to the complex structure on E. It is clear that $\dim E_{\mathbb{R}} = 2 \dim_{\mathbb{C}} E$.

The realification $E_{\mathbb{R}}$ of a complex vector space E carries a natural endomorphism J verifying $J^2 = -I$ defined by $J(u) = i\, u$. Conversely if the real vector space V is equipped with a real linear map J such that $J^2 = -I$ it becomes a complex linear space with the action of $\mathbb{C}$ on V defined by

$$z \cdot v = xv + yJ(v)$$

for all $z = x + iy \in \mathbb{C}$ and $v \in V$. We shall denote by $V(J)$ the complex vector space defined in this way. The realification of the complex space $V(J)$ is the original real vector space V and the endomorphism induced on it coincides with J. We see

[4] In Quantum Mechanics a similar decomposition of the total space can be achieved by using a compact group of symmetries for A. The irreducible invariant subspaces of our group will be finite-dimensional and the restriction of the Hamiltonian operator A to each invariant subspace gives raise to a finite-dimensional problem. Motions in central potentials are often studied in this way by using the rotation group and spherical harmonics. The radial part is then studied as a one-dimensional problem.

in this way that a complex structure on a real vector space is defined by a linear map whose minimal polynomial is the simplest possible $p(\lambda) = \lambda^2 + 1$.

Definition 2.3 A linear complex structure on the real vector space E is a real linear map $J : E \to E$ such that $J^2 = -I$.

Exercise 2.2 Prove that E must be even-dimensional as a real space.

Remark 2.3 The canonical model for a complex structure J is provided by the real space $\mathbb{R}^{2n}$ and the endomorphism J_0 given by

$$J_0 = \left(\begin{array}{c|c} 0 & -I_n \\ \hline I_n & 0 \end{array} \right). \tag{2.20}$$

Thus, if we denote by $(x^1, \ldots, x^n, y^1, \ldots, y^n)$ a generic point in $\mathbb{R}^{2n}$, multiplication by J_0 gives, $(-y^1, \ldots, -y^n, x^1, \ldots, x^n)$, hence that is equivalent to multiplication by i if we identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ by $(x^1, \ldots, x^n, y^1, \ldots, y^n) \mapsto (x^1 + iy^1, \ldots, x^n + iy^n)$.

We consider now V to be an n-dimensional real vector space. It is possible however to exploit the previous discussions leading to the structure of the flow of a complex linear system by constructing a complex vector space out of V. Such construction is called the complexification of V and it proceeds as follows:

Consider the set $E = V \times V$. We can endow E with the structure of a complex space by defining [Ar73]:

1. $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$,
2. $(a + ib)(v, w) = (av - bw, bv + aw)$.

Exercise 2.3 Check that $V \times V$ with the binary composition law + defined by (1) above and the action of $\mathbb{C}$ defined by (2) satisfies the axioms of a complex linear space.

The set $V \times V$ when endowed with the new structure of complex space will be denoted $V^{\mathbb{C}}$ and it is said to be the complexification of V. The space V is embedded as a real linear subspace in $V^{\mathbb{C}}$ by means of the map $j : V \to V \times V$, $j(v) = (v, 0)$. Moreover, every element $(v, w) \in V^{\mathbb{C}}$ can be written as $(v, w) = (v, 0) + i(w, 0) = j(v) + i\, j(w)$, because $i(w, 0) = (0, w)$.

A vector of $V^{\mathbb{C}}$ is said to be real if it is of the form $j(v) = (v, 0)$, and then it will be denoted as v instead of $j(v)$. An arbitrary element $(v, w) \in V^{\mathbb{C}}$ can be written as a sum $v + iw$, with v and w being real vectors.

Notice that if $\mathcal{B} = \{e_i \mid i \in I\}$ is a basis of the real space V, then, $\overline{\mathcal{B}} = \{(e_i, 0) \mid i \in I\}$ will be a basis of the complex space $V^{\mathbb{C}}$. Therefore,

$$\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V^{\mathbb{C}} \tag{2.21}$$

Another important remark is that if the real linear space V was endowed with a real inner product $(\cdot,\cdot)$, that is a positive definite symmetric bilinear form on E, then $V^{\mathbb{C}}$ becomes a pre-Hilbert space (actually a Hilbert space in the finite-dimensional case) by means of the Hermitean product:

$$\langle v_1 + i v_2, w_1 + i w_2 \rangle = (v_1, w_1) + (v_2, w_2) + i(v_1, w_2) - i(v_2, w_1) \tag{2.22}$$

which satisfies $\langle v, w\rangle = (v, w)$. We will discuss Hermitean products in depth in Sect. 6.2.

We introduce now the important notion of complex linear map.

Definition 2.4 A complex linear map φ between two complex linear spaces E_1 and E_2, is a map $\varphi\colon E_1 \to E_2$ such that

$$\varphi(z_1 \cdot u_1 + z_2 \cdot u_2) = z_1 \cdot \varphi(u_1) + z_2 \cdot \varphi(u_2), \quad \forall z_1, z_2 \in \mathbb{C}, \quad u_1, u_2 \in E_1.$$

Equivalently, we have:

Proposition 2.5 *A complex linear map φ between two complex linear spaces (E_1, J_1) and (E_2, J_2), is a real linear map $\varphi\colon E_1 \to E_2$ such that*

$$\varphi \circ J_1 = J_2 \circ \varphi.$$

The first example of complex linear maps is provided by the complexification of a real linear map. Given a linear map $A\colon V \to V$ of the real linear space V, It is possible to complexify A as

$$A^{\mathbb{C}}(v + i w) = (Av) + i(Aw) \tag{2.23}$$

and this correspondence satisfies:

1. $(\lambda A)^{\mathbb{C}} = \lambda A^{\mathbb{C}}$.
2. $(A + B)^{\mathbb{C}} = A^{\mathbb{C}} + B^{\mathbb{C}}$.
3. $(AB)^{\mathbb{C}} = A^{\mathbb{C}} B^{\mathbb{C}}$.
4. $(A^T)^{\mathbb{C}} = (A^{\mathbb{C}})^{\dagger}$.

Moreover, if $A^{\mathbb{C}}$ has a real eigenvalue λ, then λ is also an eigenvalue of A. More specifically, if $v + i w$ is an eigenvector of $A^{\mathbb{C}}$ corresponding to the eigenvalue $\lambda + i\mu$, then,

$$Av + i\,Aw = A^{\mathbb{C}}(v + i w) = (\lambda + i\mu)(v + i w) = (\lambda v - \mu w) + i(\lambda w + \mu v)\,,$$

and as a consequence of the uniqueness of the splitting

$$Av = \lambda v - \mu w, \quad Aw = \lambda w + \mu v$$

In particular, when $\mu = 0$, we get:

$$Av = \lambda v, \quad Aw = \lambda w \tag{2.24}$$

But $v + iw$ was an eigenvector of $A^{\mathbb{C}}$ and therefore v and w cannot vanish simultaneously.

Notice that if A is a real linear operator, the complexification $A^{\mathbb{C}}$ will define a complex liner dynamical system:

$$\frac{dz}{dt} = A^{\mathbb{C}} \cdot z; \quad z \in V^{\mathbb{C}} \tag{2.25}$$

We notice that a solution of our complexified equation stays real if the initial condition is real. In fact the operator that associates to any vector $z = x + iy$ in $V^{\mathbb{C}}$ its complex conjugate $\bar{z} = x - iy$, commutes with $A^{\mathbb{C}}$. That means that if $z(t)$ is a solution of Eq. (2.25), then $\bar{z}(t)$ will be a solution too. Then by uniqueness of the existence of a solution for given Cauchy data, if the Cauchy data are real the solution must be real for all values of t. This solution is also a solution of the real differential equation. It is also clear that a curve $t \mapsto z(t) = x(t) + iy(t)$ is a solution of the complexified equation iff the real and imaginary parts are solutions of the real equation. All this follows from the fact that $A^{\mathbb{C}}$ commutes with the multiplication by i.

Now we can look for solutions of $\dot{x} = A \cdot x$ with initial conditions $x(0)$ by solving $\dot{z} = A^{\mathbb{C}} \cdot z$ with initial condition $z(0) = x(0)$. Now a diagonalization is possible, all our previous considerations apply and we get the desired solution. It remains to express it into real form for it is the sum of vectors like $e^{t\lambda_k} x_k(t)$ where $x_k(t)$ is a polynomial of degree at most r_k. The eigenvalue λ_k is not necessarily real. What we need is to use combinations of the real and imaginary part of the solution. If $\lambda = a + ib$, $a, b \in \mathbb{R}$ we shall consider terms of the form $e^{at}\cos(bt)x(t)$ and $e^{at}\sin(bt)x(t)$ where $x(t)$ is a polynomial in t with real coefficients.

The following proposition shows that all complex structures in a given linear space are isomorphic, and isomorphic to the canonical model $(\mathbb{R}^{2n}, J_0)$.

Proposition 2.6 *Let (E, J) be a complex linear space. Then, there exists a linear isomorphism $\varphi \colon E \to \mathbb{R}^{2n}$ such that $J_0 \circ \varphi = \varphi \circ J$, i.e., (E, J) is complex isomorphic to $(\mathbb{R}^{2n}, J_0)$.*

Proof We give two proofs. Because E is a complex linear space of dimension n, let $u_1, \ldots, u_n$ denote a basis for it. The real linear space $E_{\mathbb{R}}$ is the realification of the complex linear space (E, J). A linear basis of it is provided by $u_1, \ldots, u_n, Ju_1, \ldots, Ju_n$. Then, we identify E with $\mathbb{R}^{2n}$ by means of the linear isomorphism $\psi \colon \mathbb{R}^{2n} \to E$ given by:

$$\psi(x^1, \ldots, x^n, y^1, \ldots, y^n) = x^k u_k + y^k J(u_k).$$

Then notice that

$$\psi(J_0(x^1, \ldots, x^n, y^1, \ldots, y^n)) = \psi(-y^1, \ldots, -y^n, x^1, \ldots, x^n) = -y^k u_k + x^k J(u_k)$$

and then,

$$\psi(J_0(x^1,\ldots,x^n,y^1,\ldots,y^n)) = J(y^k J(u_k) + x^k u_k) = J(\psi(x^1,\ldots,x^n,y^1,\ldots,y^n)),$$

and the result is established.

An alternative proof that will be useful for us later on, works as follows.

Consider the complexification $E^{\mathbb{C}}$ of the real space E and complexity J to $E^{\mathbb{C}}$. Thus if we denote by $u^{\mathbb{C}} = u_1 + i\,u_2$ a vector on $E^{\mathbb{C}}$, the complexified map $J^{\mathbb{C}}$ acts as $J^{\mathbb{C}}(u^{\mathbb{C}}) = J(u_1) + i\,J(u_2)$. Moreover $(J^{\mathbb{C}})^2 = -I$, hence $E^{\mathbb{C}}$ carries two alternative complex structures. We can diagonalize $J^{\mathbb{C}}$ with respect to the complex structure on $E^{\mathbb{C}}$ defined by multiplication by i, then denote by $K_\pm = \ker(J^{\mathbb{C}} \mp i)$, the eigenspaces of $\pm i$ respectively. Denote by $R_\pm = \mathrm{Im}(J^{\mathbb{C}} \pm i)$. Then, because $(J^{\mathbb{C}} + i)(J^{\mathbb{C}} - i) = 0$ it is clear that $R_\pm = K_\pm$. But $E^{\mathbb{C}}/K_\pm \cong R_\mp$. Hence,

$$\dim R_+ = \dim R_- = \dim K_+ = \dim K_-,$$

and $\dim_{\mathbb{C}} E^{\mathbb{C}}$ is even (hence the real dimension of E is even). Let us call this dimension $2n$. The dimension of K_+ is n and let $w_1^{\mathbb{C}},\ldots,w_n^{\mathbb{C}}$ a complex basis of it. The vectors $w_k^{\mathbb{C}}$ have the form $w_k^{\mathbb{C}} = u_k + i v_k$, $k = 1,\ldots,n$, and they satisfy

$$J(u_k) = v_k, \quad J(v_k) = -u_k.$$

Thus we have found a basis such that the natural identification with $\mathbb{R}^{2n}$ provided by it gives the desired map. □

The set of complex linear isomorphisms $\varphi\colon E \to E$ defines a subgroup of the real general linear group $GL(E)$. We shall denote such group by $GL(E, J)$.

As it was shown in Proposition 2.6 we can identify (E, J) with $(\mathbb{R}^{2n}, J_0)$, hence the group $GL(E, J)$ is isomorphic with a subgroup of the group $GL(2n, \mathbb{R})$. Such subgroup will be denoted by $GL(n, \mathbb{C})$ and is characterized as the set of matrices $A \in GL(2n, \mathbb{R})$ such that

$$AJ_0 = J_0 A$$

or, equivalently that A has the block form

$$A = \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix},$$

with X, Y $n \times n$ real matrices. Notice that if we identify E with $\mathbb{C}^n$, then $GL(E, J)$ becomes simply the group of invertible complex $n \times n$ matrices, i.e., the group $GL(n, \mathbb{C})$. The identification between these two representations of the same group (the fundamental one and the $2n$-dimensional) is given by the map

$$Z = X + i\,Y \mapsto \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}.$$

The (real) dimension of the group $GL(n, \mathbb{C})$ is $2n^2$.

Notice that if we consider the realification of $V^{\mathbb{C}}$, then $V^{\mathbb{C}}_{\mathbb{R}}$ is isomorphic to $V \oplus V$. A basis for $V^{\mathbb{C}}_{\mathbb{R}}$ is obtained from a basis $B = \{e_i\}_{i\in I}$ of V by $\bar{B} = \{(e_i, 0), (0, e_i)\}_{i\in I}$. In such a basis, the $\mathbb{R}$-linear map J of $V^{\mathbb{C}}_{\mathbb{R}}$ corresponding to multiplication for the imaginary unit i is represented by the standard model matrix J_0 Eq. (2.20). Notice that

$$V \cap JV = \{\mathbf{0}\}, \qquad V^{\mathbb{C}}_{\mathbb{R}} = V \oplus JV\,. \tag{2.26}$$

2.2.4 Integrating Time-Dependent Linear Systems: Dyson's Formula

We can consider again the problem of integrating a linear dynamical system. In the previous section we found the flow of the system by integrating the equation on the group and checking the convergence of the series found in that way. A similar idea can be used to integrate and find explicitly the flow of a time-dependent (i.e., non-autonomous) linear system on E like

$$\frac{dx}{dt} = A\,(t) \cdot x. \tag{2.27}$$

To discuss the solution of this equation we reconsider first the time-independent case and put it into a slightly different perspective.

We assume that the Eq. (2.27) has a solution,

$$x(t) = \phi(t, t_0)x(t_0) \tag{2.28}$$

i.e., $\phi(t, t_0)$ is the evolution matrix of the Cauchy datum $x_0 = x(t_0)$. Of course, due to the previous results we know that if A is constant $\phi(t, t_0) = e^{A(t-t_0)}$, however for the time being we are arguing independently of it. Taking the time derivative of $x(t)$ we get,

$$\dot{x}(t) = \dot{\phi}(t, t_0)x(t_0) = \dot{\phi}(t, t_0)\phi^{-1}(t, t_0)x(t) \tag{2.29}$$

therefore we have $A = \dot{\phi}(t, t_0)\phi^{-1}(t, t_0)$. Because by assumption A is independent of time, the right-hand side can be computed for any time t. Thus we find that our initial equation on E can be replaced by the equation on $GL(n, \mathbb{R})$,

$$\frac{d}{dt}\phi = A\phi \tag{2.30}$$

We use now our knowledge that $\phi(t, t_0) = e^{A(t-t_0)}$ is a solution to consider the series expansion,

$$\phi(t, t_0) = I + A\,(t - t_0) + \cdots + A^n \,\frac{(t - t_0)^n}{n!} + \cdots \tag{2.31}$$

We denote the term $A^n \frac{(t-t_0)^n}{n!}$ in the previous expansion by $R_n(t, t_0)$ and we notice that,

$$R_{n+1}(t, t_0) = \int_{t_0}^{t} A R_n(s, t_0)\, ds \tag{2.32}$$

Therefore, we can consider the matrix,

$$S_n(t, t_0) = I + \sum_{k=0}^{n} R_k(t, t_0) \tag{2.33}$$

as providing us with an approximate solution of our initial equation. It is clear that because of the independence of A on t we have,

$$R_{n+1}(t, t_0) = A^{n+1} \int_{t_0}^{t} \frac{(s - t_0)^n}{n!}\, ds \tag{2.34}$$

Therefore, the sequence: $R_0(t, t_0), \ldots, R_n(t, t_0)$ converges uniformly to the limit,

$$S(t, t_0) = I + \int_{t_0}^{t} A\, S(s, t_0)\, ds \tag{2.35}$$

Now, $S(t, t_0)$ is differentiable and satisfies:

$$\frac{d}{dt} S = A\, S \tag{2.36}$$

and we have found our solution in terms of an integral.

This new way of finding $S = e^{A(t-t_0)}$ holds true in the time dependent case as long as $A \colon I \subset \mathbb{R} \to \mathfrak{gl}(n, \mathbb{R})$ is a continuous map with M an upper bound for $||A(t)||$ in I. Then we define again,

$$R_{n+1}(t, t_0) = \int_{t_0}^{t} A(s) R_n(s, t_0)\, ds, \qquad R_0(t, t_0) = I \tag{2.37}$$

and notice that $||R_n(t, t_0)|| \leq |t - t_0|^n \frac{M^n}{n!}$. Therefore the sequence $R_0(t, t_0) + \cdots + R_n(t, t_0)$ converges uniformly on I to the limit,

$$S(t, t_0) = I + \int_{t_0}^{t} A(s) S(s, t_0)\, ds \tag{2.38}$$

with S satisfying the differential equation,

$$\frac{d}{dt} S(t, t_0) = A(t) S(t, t_0) \tag{2.39}$$

with initial condition: $S(t_0, t_0) = I$.

The matrix we have found is called the resolvent or the *resolvent kernel* of the equation $\dot{x}(t) = A(t)x(t)$.

In general one can write for $R_n(t, t_0)$ the formula,

$$R_n(t, t_0) = \int_{t_0 \leq s_1 \leq \ldots \leq s_n \leq t} A(s_n) \cdots A(s_1)\, ds_1 \cdots ds_n \tag{2.40}$$

where due care is required for the order of factors because in general $[A(s_i), A(s_j)] \neq 0$.

Defining a time-ordering operation T as,

$$T\left\{A(s)\, A\left(s'\right)\right\} = \begin{cases} A(s)\, A\left(s'\right), & s \geq s' \\ A\left(s'\right) A(s), & s < s' \end{cases} \tag{2.41}$$

and similarly for products with more than two factors (i.e., $T\{\cdot\}$ will order factors in decreasing (or non-increasing) order of the time arguments from left to right), it is not hard to convince oneself that,

$$R_n(t, t_0) = \frac{1}{n!} \int_{t_0}^{t} ds_1 \cdots ds_n T\left\{A(s_1) \cdots A(s_n)\right\} \tag{2.42}$$

and hence that $S(t, t_0)$ can be expressed a 'time-ordered exponential',[5] quoted as Dyson's formula:

$$S(t, t_0) = T\left\{\exp \int_{t_0}^{t} ds\, A(s)\right\} \tag{2.43}$$

[5] Also-called a 'product integral' (see, e.g.: [DF79]).

The matrix $S(t, t_0)$ can be given a simple interpretation. Starting (in the time-independent case, for the time being) from the linear equation, $dx/dt = A \cdot x$, we can consider a fundamental system of solutions, say $\{x_\alpha(t)\}$, $\alpha = 1, \ldots, n$, such that any other solution can be written as: $y(t) = c^\alpha x_\alpha(t)$, $y(0) = c^\alpha x_\alpha(t = 0)$. We can then construct an $n \times n$ matrix $X(t)$ whose α-th column ($\alpha = 1, \ldots, n$) is given by $x_\alpha(t)$, i.e., $X_{j\alpha}(t) = x^j_\alpha(t)$. Then, for this 'matrix of solutions' we have

$$X(t) = e^{tA} X(0)$$

i.e.,

$$e^{tA} = X(t) \circ X^{-1}(0).$$

Therefore, from a fundamental set of solutions we can construct

$$S(t, t_0) = X(t) \circ X^{-1}(t_0). \tag{2.44}$$

This relation holds true also for time-dependent equations and any solution can be written as

$$x(t) = S(t, t_0)\, x(t_0). \tag{2.45}$$

2.2.5 *From a Vector Space to Its Dual: Induced Evolution Equations*

From the equations of motion on the vector space E it is possible to induce equations of motion on any other vector space that we may build canonically out of E. We consider first the induced motion on the dual space E^*, the vector space of linear functions on E.

Starting with the linear system $dx/dt = A \cdot x$, we consider its linear flow φ_t, then we define the linear flow $\widetilde{\varphi}_t$ on E^* defined as:

$$(\widetilde{\varphi}_t \alpha)(x) = \alpha(\varphi_{-t}(x)), \quad \forall \alpha \in E^*, \quad x \in E, \tag{2.46}$$

It is a simple matter to check that:

$$\widetilde{\varphi}_{t+s} = \widetilde{\varphi}_t \circ \widetilde{\varphi}_s, \quad \widetilde{\varphi}_0 = I,$$

and that the dependence on t of $\widetilde{\varphi}_t$ is the same as that of φ_t, i.e., the family of linear maps $\widetilde{\varphi}_t$ defines a flow on E^*. Moreover, a simple computation shows:

$$\left(\frac{d}{dt}\widetilde{\varphi}_t\alpha\right)(x) = \frac{d}{dt}\alpha(\varphi_t^{-1}x) = \alpha\left(\frac{d}{dt}\varphi_t^{-1}x\right) = -\alpha(\varphi_t^{-1}A \cdot x)$$

i.e.,

$$\frac{d}{dt}\left(\widetilde{\varphi}_t\alpha\right) = -A^*\alpha$$

where $A^*\colon E^* \to E^*$ is the linear map defined as $A^*\alpha = \alpha \circ A$ and usually called the dual map to A. A simple computation shows us that the matrix representing the linear map A^* in the dual basis of a given basis in E is the transpose of the matrix representing the linear operator A. Thus we conclude stating that the dynamics on the dual vector space is induced by $-A^*$, the *opposite* of the dual map to A, or in linear coordinates the opposite of the transpose matrix which defines the dynamics on E. If we consider now the curves $x(t) = \varphi_t(x_0)$ and $\alpha(t) = \widetilde{\varphi}_t\alpha_0$ on E and E^* respectively, it is trivial to check that:

$$\frac{d}{dt}[\alpha\,(x)] = 0 \tag{2.47}$$

that is the quantity $\alpha\,(x) = \alpha_0(x_0)$ is constant under evolution.

The particular requirement we have considered here to induce a dynamics on E^* is instrumental to inducing an isospectral dynamics on linear maps or, equivalently, a dynamics that is compatible with the product of linear maps. The same can be stated for more complicated tensorial objects.

Once we have defined the induced flow on E^* it is easy to extend it to any other tensor space. For instance the induced flow on linear maps $B\colon E \to E$, that is on tensors of order $(1, 1)$, we require the following diagram,

$$\begin{tikzcd} E \arrow[r, "B"] \arrow[d, "\widetilde{\varphi}_t"'] & E \arrow[d, "\widetilde{\varphi}_t"] \\ E \arrow[r, "B_t"] & E \end{tikzcd}$$

to be commutative, i.e., $B\,(t) = \widetilde{\varphi}_t \circ B \circ \widetilde{\varphi}_t^{-1}$.

From this we get,

$$\frac{d}{dt}B\,(t) = [A, B\,(t)]\,, \tag{2.48}$$

and the evolution is isospectral. This last equation is the analog of Heisenberg equations in a classical context.

Returning to dual dynamics on the dual vector space E^*, let us consider a linear transformation of the most elementary type, namely: $B = x \otimes \alpha$, with $x \in E$ and $\alpha \in E^*$. By using the derivation property on the tensor product we find: $dB/dt = \dot{x} \otimes \alpha - x \otimes \dot{\alpha} = A\,(x \otimes \alpha) - (x \otimes \alpha)\,A$, i.e. once again Eq. (2.48).

Remark 2.4

1. If we require the evolution on linear maps Φ_t to satisfy

$$\Phi_t (M \circ N) = \Phi_t (M) \circ \Phi_t (N) , \tag{2.49}$$

we find that there exists a one-parameter group $\{\phi_t \mid t \in \mathbb{R}\}$ of automorphisms of the vector space such that

$$\Phi_t (M) = \phi_t \circ M \circ \phi_t^{-1} . \tag{2.50}$$

In one direction the statement is easily verified. In the opposite direction it relies on preservation of the 'row-by-column' product $\alpha (x) = \Phi_t [\alpha (x)] = \big(\Phi_t^* (\alpha)\big) \circ (\Phi_t (x)) = \alpha \circ \phi_t^{-1} \circ \phi_t (x)$. Then, by writing only elementary blocks in M and N, say $y \otimes \beta$, $x \otimes \alpha$, we find

$$(y \otimes \beta) \cdot (x \otimes \alpha) = \beta (x) \, y \otimes \alpha \tag{2.51}$$

and,

$$\begin{aligned} \Phi_t (\beta (x) \, y \otimes \alpha) &= \Phi_t (y) \cdot \Phi_t (\beta (x)) \otimes \Phi_t (\alpha) \\ &= \Phi_t (y) [(\Phi_t (\beta)) (\Phi_t (x))] \otimes \Phi_t (\alpha) \\ &= (\Phi_t (y) \otimes \Phi_t (\beta)) \cdot (\Phi_t (x) \otimes \Phi_t (\alpha)) \\ &= [\Phi_t (y \otimes \beta)] \cdot [\Phi_t (x \otimes \alpha)] \end{aligned}$$

and thus, Eq. (2.49) is satisfied.
2. In describing the evolution of the so-called open quantum systems one has to give up the requirement of the preservation of the product structure on linear maps (in which they represent density states) and we get a more general dynamics which cannot be described in terms of a commutator bracket.

2.3 From Linear Dynamical Systems to Vector Fields

2.3.1 Flows in the Algebra of Smooth Functions

In the previous sections we have been discussing some aspects concerning the structure and properties of linear systems using elementary notions from linear algebra.

Now we are going to present the basic tools from calculus and geometry needed to get a deeper understanding of them. For this purpose we are going to depart from the presentation of these subjects found in most elementary textbooks. We will emphasize the fact that the linear structure of the carrier space is not relevant for construction of a differential calculus in the sense that any other linear structure will define the same differential calculus.

This will be made explicit by using systematically a description in the space of (smooth) functions, where the particular choice of a linear structure will not play any role. This will have also far reaching consequences when we discuss linearization of vector fields and vice versa.

Together with the development of differential calculus we will find immediately the basic notions of differential geometry, vector fields, differential forms, etc. Then we will construct the exterior differential calculus and we will discuss its appealing algebraic structure. Finally we will use all these ideas to present the elementary geometry of dynamical systems, thus setting the basis for later developments.

We assume a basic acquaintance of the reader with the elementary notions of differential calculus and linear algebra in $\mathbb{R}^n$ or in finite-dimensional linear spaces E, such as they are discussed in standard textbooks such as [HS74], etc. We would like to discuss them again briefly here, to orient the reader in building what are going to be the notions on which this book is founded. A systematic use of concepts like those of rings, modules, algebras and the such will done throughout the text, thus the basic definitions are collected for the benefit of the reader in Appendix A. Related concepts, such as that of graded algebras, graded Lie algebras and graded derivations are also discussed in the same Appendix.

We denote by E a real linear space, finite-dimensional for most of the present discussion. In a finite-dimensional linear space all norms are equivalent, thus we will not specify which one we are using in the underlying topological notions we will introduce. Whenever we deal instead with infinite-dimensional spaces we will assume them to be Banach spaces with a given fixed norm $\|\cdot\|$, and often, more specifically Hilbert spaces.

We know from elementary analysis that functions with good regularity properties do not have to be polynomial; let us think of trigonometric or exponential functions. However these functions do share the property that they can be well approximated pointwise by polynomial functions of arbitrary degree. In fact, we may say that a function f is differentiable of class C^r at x if there exists a polynomial function P_r of degree r such that $f(y) - P_r(y)$ is continuous at the point x, and goes to zero faster than $||y - x||^r$ when $y \to x$. The class of smooth or C^∞ functions is defined as the family of functions which are of class C^r for every r at any point x in E. Thus, smooth functions are approximated by polynomial functions of arbitrary degree in the neighborhood of any point and it is the class of functions that extends most naturally the properties of the algebra of polynomials.[6] In this sense we can say that the algebra $\mathcal{F}(E)$ (or $\mathcal{F}$ for short if there is no risk of confusion) of the smooth functions on E extends the algebra $\mathcal{P}$ of the polynomials.[7]

Exercise 2.4 Prove that if f is a function on E homogeneous of degree 1 and differentiable at 0, then it must be linear.

[6] See however below, Remark 2.5.

[7] To be quite honest, the class of functions that extend more naturally the algebra of polynomials is the algebra of real analytic functions. However in this book we will restrict our attention to the algebra of smooth functions.

Exercise 2.5 Find examples of functions which are homogeneous of degree k and which are not k-tic forms, $k \geq 1$.

Because linear functions generate the algebra of polynomials and polynomials approximate arbitrarily well smooth functions in the neighborhood of any point, we will say that linear functions will 'generate' the algebra of smooth functions (we will make precise the notion of 'approximation' when introducing the notion of differentiable algebras in Sect. 3.2.2 and the meaning of 'generating' used here in Sect. 3.2.3).

As we will see in the chapters to follow, the algebra $\mathcal{F}(E)$ will play a central role in our exposition and we will come back to its structure and properties when needed.

Remark 2.5 All that has been said up to now (what we mean by linear functions and so on) depends of course in a crucial way on the linear structure that is assumed to have been assigned on E, and a different linear structure (see below, Sect. 3.5.3 for a discussion of this point) will lead to different notions of 'linear functions', 'polynomials' and so on. On the other hand, being smooth is a property of a function that is independent of the linear structure. We can conclude therefore that $\mathcal{F}(E)$ is obtained anyway as the closure (in the sense specified above) of the polynomial algebras associated with the different linear structures.

2.3.2 Transformations and Flows

Invertible linear maps L from E to E are the natural transformations of E preserving its linear structure. They form a group, the group $GL(E)$ of automorphisms of E. If we fix a linear basis $\{e_i \mid i = 1, \ldots, n\}$ on E, then E is identified with $\mathbb{R}^n$, $n = \dim E$, and linear maps from E to E are represented by square matrices. Invertible linear maps correspond in this representation to regular matrices and the group $GL(E)$ becomes the general linear group $GL(n, \mathbb{R})$. Later on, we will discuss in more detail this and other related groups of matrices (see Sect. 2.6.1). Using the previous identification of E with $\mathbb{R}^n$, any map $\phi\colon E \to E$ can be written as a n-tuple of component functions $\phi = (\phi^1, \ldots, \phi^n)$. Thus the notion of smoothness translates directly to ϕ via the components ϕ^i.

A smooth diffeomorphism of E is a smooth invertible map $\phi\colon E \to E$ whose inverse is also smooth. Clearly, if f is a smooth function, $f \circ \phi$ is also smooth and we obtain in this way a map $\phi^*\colon \mathcal{F} \to \mathcal{F}$, called the pull-back map along ϕ, as

$$\phi^*(f)(x) = (f \circ \phi)(x) \tag{2.52}$$

which preserves the product structure and, as $(\phi_2 \circ \phi_1)^* = \phi_1^* \circ \phi_2^*$, it is invertible if ϕ is a diffeomorphism, $(\phi^*)^{-1} = (\phi^{-1})^*$. Thus the set of diffeomorphisms transforms smooth functions into smooth functions and they are the natural set of maps preserving the differentiability properties of functions, as we shall see shortly.

The composition of two diffeomorphisms is again a diffeomorphism, hence they constitute a group, denoted as Diff (E), because of the associativity property of the composition of maps, and clearly contains the group $GL(E)$. The group of diffeomorphisms of E will also be called the group of transformations of the space E. We must remark that a transformation ϕ will destroy in general the linear structure on E but will leave invariant the class of smooth functions $\mathcal{F}$. This means that the notion of smoothness is not related to any particular linear structure on the space E (see Remark 2.5) and depends only on what is called the differential structure of E.

We can also consider local transformations, i.e., smooth maps defined only on open sets of E and which are invertible on their domains. In what follows we will not pay attention to the distinction between local transformations and transformations in the sense discussed above, because, given a transformation, we can always restrict our attention to an arbitrary open set on E.

2.3.3 The Dual Point of View of Dynamical Evolution

One of the aspects we would like to stress in this chapter is that all notions of differential calculus can be rephrased completely in terms of functions and their higher order analogues, differential forms, i.e., we can take the dual viewpoint and use instead of points in a linear space E, functions on it. In fact, let us consider again the algebra $\mathcal{F}(E)$ of smooth functions on E. It is clear that $\mathcal{F}(E)$ contains the same information as the set of points E itself. In fact, as we discussed earlier, we can reconstruct E from $\mathcal{F}(E)$ by considering the set of homogeneous functions of degree one.

This attitude, one of the main aspects of this book, is very common in the physical construction of theories where, implicitly, states and observables are used interchangeably to describe physical systems. The states are usually identified with points making up the space E and the observables with smooth functions (or more general objects) in E. This means that they define the value taken by the observable on any possible state. If the description of a system in terms of observables is complete, we can reconstruct the states from the observables by taking appropriate measures on them. This is the essence of Gelfand-Naimark theorem.

Evolution can then be described not in terms of how points (states) actually evolve, but in terms of the evolution of the observables themselves. This approach is often taken in physical theories where the actual description of states can be very complicated (if possible at all) but we know instead a generating set of observables. In fact, that is what it is usually done in elementary textbooks, where points are described by their coordinates x^i, thus if we say, for instance, that the position at time t of the point is given by $x^i(t)$, what we mean is that the basic position observables x^i have evolved and at time t they are given by the new functions $x^i(t)$ that turn out to be a new set of 'basic' position observables. Because the position observables x^i generate the full algebra $\mathcal{F}(E)$, describing how they evolve gives us the evolution of all other observables. An interesting observation is that the evolution described by means of observables or functions is always linear.

If we are given two observables f, g and their sum $f+g$ at a given time t_0, then their 'evolved' functions $f(t), g(t)$ and $(f+g)(t)$ satisfy $f(t)+g(t) = (f+g)(t)$. It is also clear that evolution must be invertible, thus if we denote by Φ_t the evolution operator, then $\Phi_t : \mathcal{F}(E) \to \mathcal{F}(E)$ is a linear isomorphism of the (infinite-dimensional) vector spaces $\mathcal{F}(E)$. Using this notation we will write $\Phi_t(f)$ for the observable $f(t)$ and the previous equation would be written as,

$$\Phi_t\,(f) + \Phi_t\,(g) = \Phi_t\,(f+g)\,.$$

It is also clear that, when considering the (pointwise) product $f \cdot g$, we have,

$$\Phi_t\,(f \cdot g) = \Phi_t\,(f) \cdot \Phi_t\,(g) \tag{2.53}$$

i.e., that Φ_t preserves products as well. The family $\{\Phi_t\}_{t\in\mathbb{R}}$ appears therefore as a one-parameter family of automorphisms of the algebra $\mathcal{F}(E)$ of functions on E.

Later on, see Chap. 6, if E is a Hilbert space, we would describe the dynamics on expectation value functions, a subspace of $\mathcal{F}(E)$.

Remark 2.6 Of course not all linear automorphisms of $\mathcal{F}(E)$ are also algebra automorphisms. For example the mapping: $f \mapsto \exp(\lambda k)\, f$ with k a fixed function is a linear automorphism which is not an algebra automorphism.

Another example borrowed from Quantum Mechanics is given by the linear map $\Phi_t(f) = \exp(-i\Delta^2 t)(f)$ (Δ now denotes the Laplace operator). The family Φ_t is a one-parameter group on square integrable functions but it does not define a group of automorphisms for the product because the infinitesimal generator is a second-order differential operator, which is not a derivation.

Equation (2.53) or, otherwise stated, the requirement that time evolution of 'observables' (i.e., functions in $\mathcal{F}(E)$) should preserve both the linear structure and the algebra structure, has some interesting consequences. The most relevant among them is that we can characterize the evolution operator Φ_t in more mundane terms, and precisely as a transformation on E. With reference again to Eq. (2.53) we can go further and think of the effects of the iteration of evolution, i.e., of the result of applying Φ_t and Φ_s successively. In an autonomous world, i.e., the system has no memory of the previous history of the state it is acting upon, then necessarily we must have,

$$\Phi_t(\Phi_s(f)) = \Phi_{t+s}(f). \tag{2.54}$$

Obviously,

$$\Phi_0(f) = f, \tag{2.55}$$

and we find again (cfr. Eq. (2.50)) a one-parameter group. These properties may be satisfied even without the existence of ϕ_t^{-1}; in this ocurrence we would have a semigroup. Thus, evolution will be given by a one-parameter group of isomorphisms of the algebra $\mathcal{F}(E)$. Then, we can conclude this paragraph by postulating the axioms of autonomous evolution as given by a smooth one-parameter group of automorphisms of the algebra $\mathcal{F}(E)$.

Notice that the axioms for autonomous evolution are satisfied for a smooth one-parameter group of diffeomorphisms φ_t on E:

$$\Phi_t(f) = f \circ \varphi_{-t} = \varphi^*_{-t}(f). \tag{2.56}$$

(In Sect. 3.2.2 it will be shown that this is the most general situation).

2.3.4 Differentials and Vector Fields: Locality

From their transformation properties, smooth functions are scalar quantities, i.e., they verify a transformation rule that (cfr. Eq. (2.52)) can be stated as 'the transformed function at a transformed point takes the value of the untransformed function at the given point'.

The usual definition of the differential of a given function f as

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

requires however the explicit introduction of a coordinate system, i.e., an identification of E with $\mathbb{R}^n$. However, partial derivatives and differentials transform under changes of coordinates (local transformations) in a contragradient manner. In other words, under a change of coordinates

$$x^i \mapsto y^i = \phi^i(x), \tag{2.57}$$

we have

$$df = \frac{\partial f}{\partial x^i} dx^i = \frac{\partial f}{\partial y^i} dy^i,$$

or

$$dx^i \frac{\partial}{\partial x^i} = dy^i \frac{\partial}{\partial y^i}.$$

All that is rather elementary, but proves that the association of differentials to functions is an invariant operation. In more intrinsic terms, we can rewrite the above invariance property as the commutation of the operator d and the pull-back along the map ϕ (see later Sect. 2.4.1):

$$\phi^* \circ d = d \circ \phi^*,$$

and will refer to this property by saying that "d" is a "scalar" operator, or that it is "natural" with respect to the group Diff (E) of diffeomorphisms of E.

It is well-known from elementary courses in calculus that the differential of the function $f\colon E \to \mathbb{R}$ at the point x defines a linear map $df(x)\colon E \to \mathbb{R}^n$ via $df(x)(v) = (\partial f(x)/\partial x^i)v^i$, the v^i's being the components of $v \in E$ in the given coordinate system. It is an easy exercise to prove that $df(x)(v)$ is actually a coordinate-independent expression. The differential $df(x)$ belongs therefore to the dual space E^*, i.e., to the space of (real) linear functionals on E, whose elements are called *covectors*. Thus, the differential of f at x is a covector. A basis of covectors at x will be provided by the dx^i's, which denote consistently the differentials of the linear maps $x^i\colon E \to \mathbb{R}$, $x^i(v) = v^i$, and a covector at x will be given, in local coordinates, by an expression of the form $\alpha = \alpha_i\, dx^i$, with the α_i's transforming in the appropriate way.

The operator d does actually something more. Acting on the algebra $\mathcal{F}(E)$ of smooth functions it will produce a smooth field of covectors. Smooth fields of covectors are called 1-forms and their space is denoted by $\Omega^1(E)$. So, d is actually a map $d : \mathcal{F}(E) \to \Omega^1(E)$, but more on this later in this chapter.

Another class of objects to be defined starting from a coordinate representation but that has actually an invariant nature is that of vector fields. Any such an object will be defined, in a given system of coordinates, x^i say, as a first-order differential operator of the form $X = f^i(x)\partial/\partial x^i$. If we require the set f^i to transform under a change of coordinates like (2.57), as

$$f^i(x) \mapsto g^i(y) = f^k(x)\,\frac{\partial y^i}{\partial x^k}, \tag{2.58}$$

(i.e., just as the dx^i's do in df), then X will acquire an intrinsic character as well.

By using the chain rule we see immediately that there is associated, in a natural way, to a vector field a first-order ordinary differential equation, namely,

$$\frac{dx^i}{dt} = f^i\,(x)\,. \tag{2.59}$$

In this way we are 'reading' the components of the vector field as those of a velocity field on E. We obtain therefore a definition of the action of X on functions as,

$$X(h) = \frac{dh}{dt} = \frac{\partial h}{\partial x^i}\frac{dx^i}{dt} = \frac{\partial h}{\partial x^i} f^i(x).$$

It is precisely the requirement that the evolution of a function be expressed in an invariant manner (or equivalently, that the previous system of ordinary differential equations be covariant with respect to changes of coordinates) that fixes the transformation laws for the components of a vector field. As $X(f)$ contains only the basic ingredients defining both X and df, we may read it also as an action of df on X itself,

$$X(f) = df(X). \tag{2.60}$$

Notice that, in this sense, we are identifying the value of X at a given point x with a vector dual to $df(x)$, thus $X(x)$ is a vector on E. Then a vector field is simply what its name indicates, a field of vectors, i.e., a map $X \colon E \to E$. More specifically, $X : E \to E \times E$, $x \mapsto (x, X(x))$; i.e., they are vectors along with their point of application.

2.3.5 *Vector Fields and Derivations on the Algebra of Smooth Functions*

As we have already pointed out, from an algebraic point of view, the set of smooth functions $\mathcal{F}(E)$ on a vector space is an algebra (over the field $\mathbb{R}$ in the present context). Now, a derivation over a ring $\mathcal{A}$ is a map $D \colon \mathcal{A} \to \mathcal{A}$, such that, $D(f+g) = D(f) + D(g)$ and

$$D(fg) = D(f)g + fD(g), \tag{2.61}$$

for every $f, g \in \mathcal{A}$. This equation is known as Leibniz's rule. If, instead of being simply a ring, $\mathcal{A}$ is an algebra over a field $\mathbb{K}$ (we will consider always only $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) we can extend the requirement of linearity in an obvious manner as,

$$D(\alpha f + \beta g) = \alpha D(f) + \beta D(g) \,, \quad \alpha, \beta \in \mathbb{K} \,,\, f, g, \in \mathcal{A}. \tag{2.62}$$

We can turn the set of derivations into an $\mathcal{A}$-module by defining,

$$(D_1 + D_2)(f) = D_1(f) + D_2(f)$$

and

$$(f\, D)(g) = f\, D(g) \,, \quad \forall f, g \in \mathcal{A} \,.$$

Furthermore, we can define a product of derivations as the commutator,

$$[D_1, D_2](f) = D_1(D_2(f)) - D_2(D_1(f)) \,, \qquad \forall f \in \mathcal{A} \,. \tag{2.63}$$

One can check that if the algebra is associative, then $[D_1, D_2]$ is again a derivation. Notice that, however, $D_1 \circ D_2$ and $D_2 \circ D_1$ separately are not derivations. It is also easy to check that the Lie bracket $[\cdot, \cdot]$ above, Eq. (2.63), satisfies the *Jacobi identity*,

$$[D_1, D_2], D_3] + [[D_3, D_1], D_2] + [[D_2, D_3], D_1] = 0 \,, \quad \forall D_1, D_2, D_3 \,. \tag{2.64}$$

In this way, the set of derivations over a ring becomes actually a Lie algebra.

Now we have the following:

Proposition 2.7 *The derivations over $\mathcal{F}(E)$ are the vector fields on E. Explicitly the Lie bracket of two vector fields X, Y is given by*

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

Proof It is clear that vector fields are derivations. We shall consider now an arbitrary derivation D and prove that it defines a vector field. Let $D(x^i) = X^i$ be the images of a coordinate set of functions x^i. Then let us consider the first-order Taylor expansion of f around a given point x_0. Then,

$$f(x) = f(x_0) + \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x^i}\right|_{x'} (x^i - x_0^i), \tag{2.65}$$

with x' lying in the segment joining x_0 and x. Then,

$$(Df)(x) = \sum_{i=1}^{n} \left(D\left(\left.\frac{\partial f}{\partial x^i}\right|_{x'} \right) (x^i - x_0^i) + \left.\frac{\partial f}{\partial x^i}\right|_{x'} X^i(x) \right). \tag{2.66}$$

If we take the limit $x \to x_0$ in the previous equation we get,

$$(Df)(x_0) = \sum_{i=1}^{n} X^i(x_0) \left.\frac{\partial f}{\partial x^i}\right|_{x_0}, \tag{2.67}$$

namely, $D(f) = X(f)$ for the vector field defined by the local components X^i. □

As the set $\partial/\partial x^i$, for i running from 1 to n, form a local basis for the module $\mathcal{F}(E)$, we can easily compute the commutator of two derivations,

$$X = X^i \frac{\partial}{\partial x^i}, \quad Y = Y^j \frac{\partial}{\partial x^j} \tag{2.68}$$

as follows:

$$[X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}. \tag{2.69}$$

We will denote as $\mathfrak{X}(E)$ the set of vector fields on E.

Definition 2.8 The action of $X \in \mathfrak{X}(E)$ on $f \in \mathcal{F}(E)$ defines the Lie derivative $\mathcal{L}_X$ of f along X, i.e.,

$$X(f) = \mathcal{L}_X f. \tag{2.70}$$

From the very definition of Lie brackets we obtain (always on functions),

$$\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X = \mathcal{L}_{[X,Y]}. \tag{2.71}$$

2.3.6 The 'Heisenberg' Representation of Evolution

The discussion in the previous sections has led us very close to what is known as the Heisenberg representation of evolution in Quantum Mechanics. This approach consists in translating the attention of the observer of the evolution of a given system from the states to the observables and to consider that the states do not evolve but that the observables change. In this sense, the states should be considered not as actual descriptions of the system at a given time, but as abstract descriptions of 'all' possible descriptions of the system. Then the actual evolution takes place by changing the observables while we measure the position, for instance, of the system. These considerations will sound familiar to those readers familiar with Quantum Mechanics but they are not really related with a 'quantum' description of the world but only with the duality between states and observables sketched in the previous section.

The postulate that evolution must also preserve the product of observables necessarily implies that infinitesimal evolution will be given by a derivation of the algebra. We notice immediately by differentiating Eq. (2.53) that,

$$\frac{d}{dt}\Phi_t(f \cdot g)\bigg|_{t=s} = \frac{d}{dt}\Phi_t(f)\bigg|_{t=s} \cdot \Phi_s(g) + \Phi_s(f) \cdot \frac{d}{dt}\Phi_t(f)\bigg|_{t=s}$$

Thus the infinitesimal evolution operator $\Gamma = d\Phi_t/dt|_{t=0}$ is a derivation in the algebra $\mathcal{F}(E)$. But from the previous discussions we know that derivations of the algebra of smooth functions on E are in one-to-one correspondence with vector fields on E, thus we can conclude that the axioms of evolution for observables discussed above imply that the evolution is described by a vector field Γ on E whose flow is given by the (local) one-parameter group of diffeomorphisms φ_t. Then, we will have

$$\Gamma(f) = \frac{d}{dt}\Phi_{-t}(f) = \frac{d}{dt} f \circ \varphi_t, \tag{2.72}$$

and,

$$\Gamma(x) = \frac{d}{dt}\varphi_t(x) \mid_{t=0} .$$

This equation, relating a one-parameter group of diffeomorphisms to its infinitesimal generator, is strongly reminiscent, for those who are familiar with Quantum Mechanics, of the Stone-von Neumann theorem, and can be taken actually to constitute its 'classical' version. Note however that the above is valid under the assumption that f

be a smooth (or at least a C^1) function, while the one-parameter group Φ_t, per se, can act on more general classes of functions like, e.g., continuous or simply measurable functions. In the latter cases we cannot revert from the group to the infinitesimal generator acting on functions, and again that is the classical counterpart of the problems with domains (of the infinitesimal generator) that are well known to occur in the transition from the unitary evolution to its self-adjoint generator. A prototypical example of an operator for which this kind of problems arises is provided by the operator d/dx, the infinitesimal generator of translations on the real line. A nice discussion of the relation between completeness and self-adjointness may be found in [ZK93].

Remark 2.7 A few remarks are in order here. First, not all vector fields on E arise in this form. The existence theorem for ordinary differential equations, states only the local existence of the flow φ_t, thus in general for an arbitrary vector field we will not be able to extend the local solutions $x(t)$ for all values of t, hence we will not be able to define a one-parameter family of diffeomorphisms φ_t but only a local one-parameter group (We will discuss this issue in the next section).

An elementary example is provided, in one dimension, by the vector field: $\Gamma(x) = \alpha x^2 \partial/\partial x$, $\alpha = \text{const.}$, $x \in \mathbb{R}$, whose integral curves are of the form: $x(t) = x_0/(1-\alpha x_0 t)$, $x_0 = x(0)$, that, for every $x_0 \neq 0$, will 'explode' to infinity in a finite time: $t^* = 1/\alpha x_0$ and therefore is not complete.

A less elementary example is provided by the vector fields[8]: $X_{(i)} = \epsilon_{ijk} x^j \partial/\partial x^k$, $i = 1, 2, 3$ on $\mathbb{R}^3$, where ϵ_{ijk} is the totally antisymmetric (Ricci or Levi-Civita) tensor ($\epsilon_{123} = 1$), which close on the Lie algebra of $SO(3)$ (or of $SU(2)$), i.e.: $\left[X_{(i)}, X_{(j)}\right] = \epsilon_{ijk} X_{(k)}$ and generate the rotations in $\mathbb{R}^3$. They can be restricted to the unit sphere S^2, where they become, in spherical polar coordinates (ϕ, θ):

$$
\begin{aligned}
X_1 &= \sin\phi \frac{\partial}{\partial\theta} + \cot\theta \cos\phi \frac{\partial}{\partial\phi}, \\
X_2 &= \cos\phi \frac{\partial}{\partial\theta} - \cot\theta \sin\phi \frac{\partial}{\partial\phi}, \\
X_3 &= \frac{\partial}{\partial\phi}.
\end{aligned} \tag{2.73}
$$

Spherical polar coordinates are, of course, a system of coordinates only for the sphere without the poles and one meridian passing through them, i.e. for: $\theta \in (0, \pi)$ and $\phi \in (0, 2\pi)$. So, the vector fields (2.73) are actually globally defined on the cylinder: $(0, \pi) \times [0, 2\pi]$, and it is not difficult to convince oneself that, out of these three vector fields, only $X_{(3)}$ is complete.

An even more startling example is provided by the realization of the pseudo-rotation group on the real line;

[8] These fields are not independent. In fact, denoting them collectively as: $\mathbf{X} = \left(X_{(1)}, X_{(2)}, X_{(3)}\right)$, with: $\mathbf{x} = \left(x^1, x^2, x^3\right)$, it is obvious that: $\mathbf{x} \cdot \mathbf{X} = 0$.

$$X_1 = \frac{\partial}{\partial x}, \quad X_2 = \sin x \frac{\partial}{\partial x}, \quad X_3 = \cos x \frac{\partial}{\partial x}.$$

The above notions can be carried over to the level of the algebra of smooth functions modifying in an appropriate way the axioms above but we will not do it here in order not to create an unnecessary complication in the definitions. Thus in what follows by evolution we will understand the (local) smooth one-parameter group of transformations defined by a derivation Γ of the algebra of smooth functions $\mathcal{F}(E)$ or equivalently by the vector field Γ on E.

Secondly, the autonomous condition on evolution introduced on the axioms above can be removed because systems can have at a given point, and in fact they often do, memory of their previous history. Then, evolution will be given simply by a (local) smooth one-parameter family of diffeomorphisms on E. That is equivalent to giving (local) smooth one-parameter family of derivations Γ_t on E, which is usually called a time-dependent vector field.

2.3.7 The Integration Problem for Vector Fields

Thus we have seen that a one-parameter group of automorphisms Φ_t of $\mathcal{F}$ defines, at least formally, a derivation Γ. Derivations are identified with vector fields, thus we have a way to recover the group of automorphisms by integrating the differential equation defined by the vector field. Because we know the action of a vector field on functions, $f \mapsto X(f)$ we could try to compute the flow φ of X on a function f by 'integrating' the previous formula and we will get:

$$\varphi_t^*(f) = \sum_{k \geq 0} \frac{t^k}{k!} \mathcal{L}_X^k(f). \tag{2.74}$$

However the previous formula could raise a few eyebrows. When is the series on the right-hand side of equation (2.74) convergent? Even if it is convergent with respect to some reasonable topology on some class of functions, would the family φ_t of maps thus obtained be the 'integral flow' of X as in the linear case?

This is the first nonlinear integrability problem we are facing and its solution provides the key to predict the evolution of a given system. The answer to this problem is, of course, well known and it constitutes the main theorem in the theory of ordinary differential equations. Before discussing it, we would like to elaborate further on the notion of tangent vectors and vector fields, both for the sake of the stament of the solution to this problem and for further use in the construction of geometrical structures associated to given dynamics.

2.3.7.1 The Tangent and the Cotangent Bundle

The interpretation of vector fields and 1-forms as fields of vectors and covectors respectively captures only a partial aspect of their use. The ordinary differential equation (2.59), associated with a given vector field X, shows us an alternative interpretation of vector fields which is at the basis of the present geometrical construction. The components $f^i(x)$ of the vector field X are interpreted according to Eq. (2.59) as the components of the velocity of a curve $\gamma(t)$ which is a solution of the ordinary differential equation defined by X. Thus, the value of the vector field X at a given point x can be geometrically thought of as a tangent vector to a curve $\gamma(t)$ passing through x at $t = 0$. Notice that two curves $\gamma_1(t)$ and $\gamma_2(t)$ having a contact of order 1 at x define the same tangent vector, i.e.

$$\left.\frac{d\gamma_1(t)}{dt}\right|_{t=0} = \left.\frac{d\gamma_2(t)}{dt}\right|_{t=0} . \tag{2.75}$$

Therefore a tangent vector at the point x can be thought as an equivalence class of curves[9] passing through x with respect to the equivalence relation of having a contact of order 1 at x. We shall denote by $T_x E$ the collection of all tangent vectors to E at x. If v_x denotes one of these tangent vectors we can define the variation of a function f at the point x in the direction of v_x as

$$v_x(f) = \left.\frac{d}{dt} f \circ \gamma(t)\right|_{t=0} , \tag{2.76}$$

where $\gamma(t)$ is any representative in the equivalence class defining v_x. Notice that the definition of the numerical value $v_x(f)$ does not depend on the choice of the representative γ we make. Thus v_x is a linear map from the space of differentiable functions defined in a given neighborhood of the considered point x into $\mathbb{R}$ with the additional property that

$$v_x(f_1 f_2) = f_1(x) v_x(f_2) + f_2(x) v_x(f_1)$$

This additional property is what characterizes vectors v_x at a point x among all linear maps. We can define an addition on the space of tangent vectors as

$$(v_x + u_x)(f) = v_x(f) + u_x(f) \tag{2.77}$$

Of course we need to guarantee that the object thus defined, $u_x + v_x$, corresponds again to a tangent vector, i.e., we need to find a curve passing through x such that its tangent vector will be $u_x + v_x$. Because we are in a linear space E, that is actually very easy. Given a vector $v \in E$ there is a natural map identification $v \mapsto v_x$ with a tangent vector at x, which is the equivalence class corresponding to the curve

[9] See however below, Appendix C, Sect. C.2 for a similar discussion in a more general context.

$\gamma(t) = x + tv$. Such identification is clearly one-to-one and the vector corresponding to $v_x + u_x$ is the vector corresponding to the curve $\gamma(t) = x + t(u + v)$.[10] From this perspective, we see that Eq. (2.77) actually defines the addition on the tangent space $T_x E$ and it shows that this addition does not depend on any linear structure of E. However, in this particular setting it is also true that we have a natural isomorphism between $T_x E$ and E as linear spaces. Hence, we can think that the tangent space at the point x is a copy of the background space E put at the point x. The tangent vectors corresponding to the curves $\gamma_i(t) = x + t\, e_i$, where $\{e_i \mid i = 1, ..., n\}$ is a given basis in E, are denoted by $(\partial/\partial x^i)_{|x}$. The notation is consistent with the operator defined on functions, because,

$$\frac{\partial}{\partial x^i}\bigg|_x (f) = \frac{d}{dt}(f \circ \gamma_i)(t)\bigg|_{t=0} = \frac{d}{dt} f(x + te_i)\bigg|_{t=0} = \frac{\partial f}{\partial x^i}\bigg|_x . \tag{2.78}$$

The union of all tangent spaces $T_x E$ is collectively denoted by,

$$TE = \bigcup_{x \in E} T_x E \tag{2.79}$$

and it is clearly isomorphic as a linear space to $E \oplus E$. An element of TE is thus a pair (x, v) where x is a point (vector) in E and v is a tangent vector at x. There is a natural projection $\tau_E : TE \to E$, defined as $\tau_E(x, v) = x$. Such a structure, the triple (TE, τ_E, E), is called the tangent bundle over E.

If $\phi\colon E \to E$ denotes a smooth map and v_x is a tangent vector at x, then $\phi_*(v_x)$ is a tangent vector at $\phi(x)$ defined as:

$$(\phi_*(v_x))\,(f) = v_x(f \circ \phi)\,,$$

for any f defined in an open neighborhood of $\phi(x)$. Thus we have defined a map $\phi_* : TE \to TE$ called the tangent map to ϕ or, sometimes, the differential of ϕ. It is clearly satisfied that $(\phi \circ \psi)_* = \phi_* \circ \psi_*$ which is just a consequence of the chain rule. Thus if ϕ is a diffeomorphism, then $\phi_*^{-1} = (\phi^{-1})_*$.

Turning back to the notion of a vector field, we see that a vector field consists of a smooth selection of a tangent vector at each point of x, the values of the vector field X being the vectors $X(x)$ tangent to E at x. Thus, a vector field is a smooth map $X : E \to TE$ such that it maps x into $(x, X(x))$ where $X(x) \in T_x E$.

The above maps satisfy $\tau_E \circ X = \mathrm{id}_E$ and are called *cross sections* of the tangent bundle. Therefore, in this terminology a vector field is just a cross section of the tangent bundle. Moreover, if ϕ is a diffeomorphism, then we may define the push-forward $\phi_* X$ of X along ϕ as follows: $(\phi_* X)(x) = \phi_*(X(x))$.

[10] However this idea will also work in a more abstract setting in the sense that it is possible to show that there is a one-to-one correspondence between equivalence classes of curves possessing a contact of order 1 at x and linear first-order differential operators v acting locally on functions at x.

Because E is already a vector space, there is a distinguished vector field called the dilation (or Liouville) vector field Δ[11]:

$$\Delta : E \to TE; \quad x \mapsto \Delta(x) = (x, x) , \quad x \in E. \tag{2.80}$$

The Liouville vector field Δ allows us to identify the vector space structure on the tangent space with the vector space structure on E (see later on Sect. 3.3.1), i.e., Δ can be defined only if E itself is a vector space. The graph of Δ in $TE \approx E \times E$ is a subspace of $TE \oplus TE$, the diagonal vector subspace.

Together with the tangent bundle we do have its dual, the cotangent bundle. We describe it briefly here. Again, as in the case of tangent vectors being identified with equivalence classes of curves, we have a natural identification of functions at a given point by possessing the same differential (the actual value of the function is irrelevant). Any such equivalence class actually defines a covector $\alpha_x = df(x) \in E^*$, for some f. Thus the space of covectors at x (differentials of functions at x) defines the cotangent space denoted by T_x^*E. Such a space is obviously naturally isomorphic to E^* and is dual to T_xE. The set $\{dx^i(p) \mid i = 1, \ldots n\}$ is the dual basis of $\{(\partial/\partial x^i)_p \mid i = 1, \ldots n\}$ at each point $p \in E$.

The pairing between both is given as follows: If γ is a curve representing the tangent vector v_x and f is a function representing the cotangent vector α_x, then,

$$\langle \alpha_x, v_x \rangle = \frac{d}{dt}(f \circ \gamma)(t)\bigg|_{t=0} . \tag{2.81}$$

The union of all cotangent spaces T_x^*E is denoted by T^*E. Clearly, T^*E is naturally isomorphic to $E \oplus E^*$ and it carries a natural projection $\pi_E : T^*E \to E$, defined as $\pi_E(x, \alpha_x) = x$. The triple (T^*E, π_E, E) is called the cotangent bundle of E.

A smooth assignment of a covector at x to any point $x \in E$ is called a 1-form. Thus a 1-form α is a smooth map $\alpha \colon E \to T^*E$ such that $x \mapsto (x, \alpha(x))$, i.e., again a 1-form is a cross section of the cotangent bundle, and therefore such that $\pi_E \circ \alpha = \mathrm{id}_E$.

2.3.7.2 Vector Fields and Local Flows

Thus given a vector field X on E, we want to determine the existence of a flow φ_t for X. In general, as it was pointed out before, this cannot be done globally, however it is always possible to do it locally (see [Ar73] by a masterly exposition of the subject).

Theorem 2.9 (Fundamental theorem of ordinary differential equations) *Given a smooth vector field X on E, for every $x \in E$ there exists an open neighborhood U of x and a number $\epsilon > 0$ such that given any point $y \in U$ and any t with $|t| < \epsilon$, the solution $\varphi_t(y)$ of the equation $du/dt = X(u)$ satisfying the initial condition y at $t = 0$ exists, is unique, depends smoothly on y and t and satisfies:*

[11] Also-called the *Euler differential operator.*

$$\varphi_{t+s}(y) = \varphi_t \circ \varphi_s(y), \quad |t| < \epsilon, \quad |s| < \epsilon, \quad |t+s| < \epsilon.$$

A vector field whose solutions can be extended from $-\infty$ to $+\infty$ so as to give raise to a one-parameter group will be said to be a 'complete' vector field. As already said, generic vector fields need not be complete. However, if the vector field is defined on a compact set, or better, is different from zero only on a compact set, it is complete. This result is also true for smooth manifolds ([Ar73], Theorem 35.1), but we will state it here just in the case of linear spaces.

Theorem 2.10 *Let X be a smooth vector field different from zero only in a compact subset K of E. Then there exists a one-parameter group of diffeomorphisms $\varphi_t : E \to E$ for which X is the velocity field:*

$$\frac{d}{dt}\varphi_t(x) = X(\varphi_t(x)), \quad \forall x \in E.$$

Thus if Γ is a vector field on E, then we may find a one-parameter group of diffeomorphisms φ_t that describes the trajectories of it on a compact set, or in other words, there is a one-parameter group of automorphisms of $\mathcal{F}(E)$ that restricted to $\mathcal{F}(K)$ for K a compact set satisfies:

$$\frac{d}{dt}\varphi_t = \Gamma \circ \varphi_t\,. \tag{2.82}$$

The picture we can get from this situation is that given a vector field and choosing a compact set K neighborhood of a given point, there is a complete flow φ_t that acting on points on K will produce the trajectories of Γ but that a little bit after exiting K the flow will 'freeze' leaving the points fixed.

The idea to prove this is simple. Given a compact set K, we may construct (taking for instance the a closed ball containing K) a smooth 'bump' function ρ adapted to K, that is a function such that $\rho = 1$ on K and $\rho = 0$ in the complementary of the closure of a ball containing K. Thus multiplying X by ρ we have a vector field to which we may apply Theorem 2.10 and whose complete flow is a one-parameter group of diffeomorphims satisfying the previous equation (2.82) on K.

The previous formula (2.82) provides a rigorous setting for Eq. (2.72) and makes sense of the formal integration idea expressed by Eq. (2.74).

Thus, using the previous observation, we may assume that we have a one-parameter group of diffeomorphisms integrating a given dynamics (that is describing the trajectories of our system on compact sets) and we will use this in what follows without explicit mention to it.

We will close this collection of ideas by noticing that the Lie derivative of a function f along a vector field Γ that was derived before represents the infinitesimal variation of the function in the direction of the vector field and can be defined (by using Eq. (2.82)) as:

$$\mathcal{L}_\Gamma f = \Gamma(f) = \frac{d}{dt}\varphi_{-t}^* f \mid_{t=0} = \frac{d}{dt} f \circ \varphi_{-t} \mid_{t=0} .$$

2.4 Exterior Differential Calculus on Linear Spaces

2.4.1 Differential Forms

Having defined the differentials as objects that behave as 'scalars' one may also say that they are 'natural' under arbitrary (smooth) changes of coordinates. Thus we may form $\mathcal{F}(E)$-linear combinations of differentials of functions, i.e., sums of monomials of the form $f\, dg$, $f, g \in \mathcal{F}(E)$. Then if we transform f and g by using a diffeomorphisms ϕ of E, then $f\, dg$ transforms accordingly. The monomials $f\, dg$ will generate an $\mathcal{F}(E)$-module $\Omega^1(E)$ whose elements are differential 1-forms over E. Notice that $\Omega^1(E)$ is just the space of sections of the cotangent bundle T^*E because any 1-form α can be written as $\alpha = \alpha_i\, dx^i$, once a linear coordinate system x^i on E has been chosen, i.e., a linear basis e_i has been selected. Then it is clear that the dx^i's are a basis of $\Omega^1(E)$, i.e., any 1-form can be given uniquely as an $\mathcal{F}(E)$-linear combination of them. Equivalently, the dx^i's are linearly independent. We may choose also any other basis, i.e., any other set df^i, provided they are also independent. Of course, the df^i's are linearly independent iff the f^i's are functionally independent, i.e., no nonconstant function $\Phi = \Phi(f^1, \dots, f^n)$ exists such that $\Phi(f^1, \dots, f^n) = \text{const}$. Later on, see Eq. (2.89), we will give a compact characterization of this condition). A 1-form α which is the differential of a function, i.e., $\alpha = df$ will be called exact.

We may also wish to extend the action of d from functions, i.e., zero-forms, to 1-forms. Let us start with a monomial like $f\, dg$, and let $d(f\, dg)$ be defined by acting on the coefficients of the dx^i's (which are functions, after all). We are immediately faced with the problem of defining products of differentials. If we choose, e.g., tensor products, we obtain for $d(f\, dg)$ the expression,

$$d(f\, dg) = \left(\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j} + f\frac{\partial^2 g}{\partial x^i \partial x^j}\right) dx^i \otimes dx^j .$$

Changing coordinates $x \mapsto x'$, a tedious but straightforward calculation yields,

$$d\,(f dg) = \left(\frac{\partial f}{\partial x'^i}\frac{\partial g}{\partial x'^j} + f\frac{\partial g}{\partial x^k}\frac{\partial^2 x^k}{\partial x'^i \partial x'^j} + f\frac{\partial^2 g}{\partial x'^i \partial x'^j}\right) dx'^i \otimes dx'^j .$$

So, naturality of the operator d gets lost (except for linear changes of coordinates) unless we redefine the product of differentials in such a way as to eliminate symmetric parts. This leads us to define the *wedge* (or *exterior*) product $dx^i \wedge dx^j$ as the antisymmetrized product,

$$dx^i \wedge dx^j = \frac{1}{2}(dx^i \otimes dx^j - dx^j \otimes dx^i), \tag{2.83}$$

and, by extension,

$$df \wedge dg = \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j} dx^i \wedge dx^j = \frac{1}{2}\left(\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial x^j} - \frac{\partial g}{\partial x^i}\frac{\partial f}{\partial x^j}\right) dx^i \wedge dx^j,$$

and to extend (by definition now) the action of d on monomials of the form $f\, dg$ as

$$d(f\, dg) = df \wedge dg. \tag{2.84}$$

This definition has the advantage of retaining the naturality of the *exterior differential*, as d will be called from now on. It is remarkable that d is the only derivation which is a 'scalar operator' with respect to the full diffeormorphism group [Pa59].

Remark 2.8 With the Definition (2.83), the wedge product differs by a normalization factor $1/2$ in the case of the product of two one-forms as in Eq. (2.83), $1/n!$ for the product of n one-forms, from the antisymmetrized product one would obtain using Eq. (10.6). This normalization turns out however to be more convenient, and we will use it throughout whenever we will deal with differential forms.

The $\mathcal{F}(E)$-linear space spanned by the monomials $dx^i \wedge dx^j$ (or by the independent wedge products in any other basis) will be called the space of smooth two-forms, or just 2-forms for short, and will be denoted by $\Omega^2(E)$. A general 2-form will have the expression,

$$\alpha = \alpha_{ij}\, dx^i \wedge dx^j, \quad \alpha_{ij} = -\alpha_{ji}, \; \alpha_{ij} \in \mathcal{F}(E) \tag{2.85}$$

It is left to the reader to work out the transformation law of the coefficients α_{ij}'s under arbitrary (smooth) changes of coordinates.

Because E is a linear space then the $\mathcal{F}(E)$-module of 1-forms is finitely generated (a system of generators is provided by the differentials dx^i of a linear system of coordinates x^i for instance). Not only that, $\Omega^1(E)$ is a free module over $\mathcal{F}(E)$ and the 1-forms dx^i provide a basis for it, then $\dim_{\mathcal{F}(E)} \Omega^1(E) = \dim E$.

Similarly the $\mathcal{F}(E)$-module $\Omega^2(E)$ is free and the 2-forms $dx^i \wedge dx^j$ provide a basis for it. Then clearly if $\dim E = n$, $\dim_{\mathcal{F}(E)} \Omega^2(E) = \binom{n}{2}$. Here too, if $\alpha = d\theta$ for some $\theta \in \Omega^1(E)$, α will be called an exact two-form.

As we discussed before, starting with the tangent bundle TE we can form the cotangent bundle or the bundle of covectors or linear 1-forms over E, but we could also form the bundle of linear 2-forms (or skew symmetric linear (0, 2) tensors) over E. We shall denote such bundle as $\Omega^2(T^*E)$, and it is just the union of all spaces $\Omega^2(T_xE)$, $x \in E$. We also have as before that $\Omega^2(T^*E) \cong E \oplus \Omega^2(E)$. Notice that $\Omega^1(T^*E) = T^*E$. Cross sections of $\Omega^2(T^*E)$ are smooth 2-forms $\omega \in \Omega^2(E)$. It is also customary to denote the space of cross sections of the bundles TE, T^*E, $\Omega^2(T^*E)$ by $\Gamma(TE)$, $\Gamma(T^*E)$, and $\Gamma(\Omega^2(T^*E))$, etc., thus $\Gamma(T^*E) = \Omega^1(E)$ and so on.

Let us recall that a linear 1-form α acts on a vector producing a number. Thus we may also think that $\Omega^1(E)$ is the dual with respect to the algebra $\mathcal{F}(E)$ of the module of vector fields $\mathfrak{X}(E)$, i..e, a 1-form α is a $\mathcal{F}(E)$-linear map $\alpha\colon \mathfrak{X}(E) \to \mathcal{F}(E)$. $\alpha(X) \in \mathcal{F}(E)$, for all $X \in \mathfrak{X}(E)$ and being defined as $\alpha(X)(x) = \langle \alpha(x), X(x)\rangle$ where $\langle\cdot,\cdot\rangle$ denotes as usual the natural pairing between a linear space and its dual.

A similar argument applies to 2-forms. A 2-form $\omega \in \Omega^2(E)$ can be considered as defining a skew symmetric $\mathcal{F}(E)$-bilinear map on the module $\mathfrak{X}(E)$, that is $\omega(X, Y)$ is a smooth function on E for any $X, Y \in \mathfrak{X}(E)$.

Now from the definition of the wedge product it follows at once that,

$$(dx^i \wedge dx^j)(X, Y) = dx^i(X)dx^j(Y) - dx^i(Y)dx^j(X)$$

for any pair of vector fields X and Y.

In particular, having in mind the definition of the Lie derivative on functions, it is not hard to see that, if $\alpha = d\theta$ is exact, then,

$$d\theta(X, Y) = \mathcal{L}_X(\theta(Y)) - \mathcal{L}_Y(\theta(X)) - \theta([X, Y]), \qquad \forall X, Y \in \mathfrak{X}(E)\,. \tag{2.86}$$

It is left as an exercise to prove that, notwithstanding the differential nature of the Lie derivative, the right-hand side of Eq. (2.86) is actually $\mathcal{F}(E)$-linear in both X and Y (besides being manifestly skew-symmetric). Together with Eq. (2.71), this tells us also that, if $\theta = df$ is an exact 1-form, then, $d\theta = d^2 f = 0$, i.e., that $d^2 = d \circ d = 0$ (on functions only, for the time being).

2.4.2 Exterior Differential Calculus: Cartan Calculus

If we extend the wedge product into an associative product, we can generate forms of higher degree by taking wedge products of forms of lower order. In a systematic manner, if $\alpha_1, \ldots, \alpha_n$ is a basis of 1-forms (e.g., $\alpha_i = df_i$ for a set of functionally independent functions), then, monomials of the form,

$$\alpha_{i_1 \ldots i_k} = \alpha_{i_1} \wedge \cdots \wedge \alpha_{i_k}, \quad i_1, \ldots, i_k = 1, \ldots, n, \tag{2.87}$$

will generate the $\mathcal{F}(E)$-linear space of k-forms on E, denoted as $\Omega^k(E)$. Alternatively, we may think that a k-form α is an $\mathcal{F}(E)$-multilinear map

$$\alpha\colon \mathfrak{X}(E) \times \overset{k}{\cdots} \times \mathfrak{X}(E) \to \mathcal{F}(E),$$

such that $\alpha(X_1, \ldots, X_i, \ldots, X_j, \ldots, X_k) = -\alpha(X_1, \ldots, X_j, \ldots, X_i, \ldots, X_k)$ for all i, j. Then, given any smooth map $\phi\colon E \to E$, we may define the pull-back of any k-form along the map ϕ as

$$\phi^*\alpha(X_1, \ldots, X_k) = \alpha(\phi_* X_1, \ldots, \phi_* X_k), \qquad \forall X_i \in \mathfrak{X}(E). \tag{2.88}$$

Remark 2.9 The previous formula should be understood pointwise, that is considering the k-form α as a section of the bundle $\Omega^k(T^*E) \to E$. Then we will write:

$$(\phi^*\alpha)_x(v_1, \ldots, v_k) = \alpha_x(\phi_* v_1, \ldots, \phi_* v_k), \quad \forall v_1, \ldots, v_k \in T_x E.$$

However the formula above (2.88) makes perfect sense if ϕ is a diffeomorphism.

Remark 2.10 It is not hard to prove that the wedge product of two 1-forms (and hence of any number) vanishes iff the forms are linearly dependent. The condition for the linear independence for k-monomials will be then,

$$\alpha_{i_1} \wedge \cdots \wedge \alpha_{i_k} \neq 0.$$

Hence, if $k \leq n$, $\dim_{\mathcal{F}(E)} \Omega^k(E) = \binom{n}{k}$ and there will be no room for forms of degree higher than the dimension n of E, actually $\Omega^n(E) = \mathcal{F}(E)$. We note parenthetically that functional independence of a set of k functions $f_1, \ldots, f_k \in \mathcal{F}(E)$ will be expressed by the condition

$$df_1 \wedge \cdots \wedge df_k \neq 0. \tag{2.89}$$

Note that $\Omega^\bullet(E) = \bigoplus_{k\geq 0} \Omega^k(E)$ is an associative graded algebra (see Appendix A) and the elements of $\Omega^k(E)$ are said to be homogenous of degree k.

We wish now to extend the action of the exterior differential d to forms of arbitrary rank. Let us start again with monomials in a basis generated by exact 1-forms. To make things simpler, let $\alpha = f\, dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ be a monomial of rank k (for a fixed set of i_l's, $l = 1, \ldots, k$). Then, we define $d\alpha$ as the monomial of rank $k+1$,

$$d\alpha = df \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k}. \tag{2.90}$$

The exterior differential, with this obvious extension, is then defined on forms of arbitrary rank and is an $\mathbb{R}$-linear map $d\colon \Omega^k(E) \to \Omega^{k+1}(E)$, with $\Omega^0(E) = \mathcal{F}(E)$, and $\Omega^{n+1}(E) = \mathbf{0}$.

If a set of coordinates $\{x^1, \ldots, x^n\}$ has been chosen for E, then, $\varpi = dx^1 \wedge \cdots \wedge dx^n$ will be a basis for the ($\mathcal{F}(E)$-one-dimensional) module $\Omega^n(E)$, i.e., any $n-$form will be expressible as $f\,\varpi$ for some $f \in \mathcal{F}(E)$. In view of the fact that coordinates are globally defined for a vector space, ϖ can be thought of as well as a basis for the space $\Omega^n(T_x^*E)$ of the n-forms based at any point $x \in E$. As such, it will be better denoted as $\varpi(x)$ although the notation may appear somewhat redundant in this particular case. It enjoys the features that:

1. $\varpi(x) \neq 0, \forall x \in E$,
2. If we perform a permutation of coordinates, $x^i \mapsto y^i = x^{\pi(i)}$, $\pi \in S_n$ the group of permutations of n elements, then: $\varpi' = dy^1 \wedge \cdots \wedge dy^n = \mathrm{sign}(\pi)\varpi$, where $\mathrm{sign}(\pi)$ stands for the signature (the parity) of the permutation π. So, $\varpi' = \pm\varpi$ according to the parity of the permutation.

In general, a nowhere vanishing form of maximal rank will be called a *volume form*, and we have just seen that a volume form always exists on a vector space. This may not be so in more general situations in which we deal with spaces that can be modeled on vector (and hence Euclidean) spaces only locally, in which case it may well be that volume forms exist only locally, but this more general case is, for the time being, outside our scopes. We have also seen that $-\varpi$ is an equally acceptable volume-form if ϖ is. Each choice will be said to define an *orientation* on E. Again, that is a globally defined notion as long as E is a vector space, but need not be so in more general situations.

Let now ϕ be a linear map from E to E. In a given system of coordinates, $\phi : x^i \mapsto y^i = A^i{}_j x^j$, i.e., ϕ will be represented by the matrix $A = (A^i{}_j)$. Then, by using the properties of the wedge product, it is not difficult to show that

$$(\phi^*\varpi)(x) = \det(A)\,\varpi(x).$$

More generally, if ϕ is a smooth map (not necessarily a linear one), the pull-back $\phi^*\varpi$ will be again an n-form, and hence proportional to ϖ itself, and this motivates the following:

Definition 2.11 Let $\phi \colon E \to E$ be a smooth map and let ϖ be a volume-form. Then the determinant of ϕ, $\det(\phi)$, is defined by

$$\phi^*\varpi = \det(\phi)\,\varpi\,. \tag{2.91}$$

A straightforward calculation leads then to the result that, if ϕ_1, ϕ_2, are smooth maps, then the determinant function enjoys the property we are familiar with in the linear case, i.e., that,

$$\det(\phi_1 \circ \phi_2) = \det(\phi_2 \circ \phi_1) = \det(\phi_1)\ \det(\phi_2). \tag{2.92}$$

Remark 2.11 If the volume-form is realized as: $\varpi = dx^1 \wedge \cdots \wedge dx^n$ in a given system of coordinates, then $\det(\phi)$ at point x is, of course, nothing but the familiar Jacobian determinant of ϕ at $x \in E$.

Equation (2.90) defines also the action of d on a wedge product of 1-forms. For example, let $\alpha = g\,df$ and $\beta = h\,dk$, $f, g, h, k \in \mathcal{F}(E)$, then, $\alpha \wedge \beta = (gh)\,df \wedge dk$ and one proves immediately that,

$$d(\alpha\wedge\beta) = d(gh)\wedge df\wedge dk = (g\,dh+h\,dg)\wedge df\wedge dk = d\alpha\wedge\beta-\alpha\wedge d\beta. \quad (2.93)$$

Using bilinearity, if α were a 2-form (actually a monomial of rank 2) we would get instead,

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta - \alpha \wedge d\beta. \quad (2.94)$$

Extending these results in an obvious way from monomials to forms we obtain eventually:

Proposition 2.12 *If $\alpha \in \Omega^p(E)$ and $\beta \in \Omega^q(E)$, then $\alpha \wedge \beta \in \Omega^{p+q}(E)$ and the graded Leibniz rule is satisfied:*

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta. \quad (2.95)$$

Morevoer, we have that $d^2 = 0$.

Finally, if we consider a vector field, $X = X^i \frac{\partial}{\partial x^i}$, the Lie derivative of a volume form ϖ is proportional to the volume form, $\mathcal{L}_X\varpi = f\varpi$. As

$$\mathcal{L}_X(dx^1 \wedge \cdots \wedge dx^n) = \left(\frac{\partial X^1}{\partial x^1} + \cdots + \frac{\partial X^n}{\partial x^n}\right) dx^1 \wedge \cdots \wedge dx^n$$

when $\varpi = dx^1 \wedge \cdots \wedge dx^n$, the proportionality factor is called the *divergence* of X, $f = \mathrm{div}\,(X)$, because using such global coordinate system,

$$\mathcal{L}_X(dx^1 \wedge \cdots \wedge dx^n) = \mathrm{div}\,(X)\, dx^1 \wedge \cdots \wedge dx^n.$$

It is also possible to associate with any element $X \in \mathfrak{X}(E)$ a derivation of degree -1 in the graded algebra of forms $\Omega^\bullet(E)$, called an *inner derivation* (or a *contraction*). We set $i_X\colon \Omega^p(E) \to \Omega^{p-1}(E)$, where $\alpha \mapsto i_X\alpha$ with $i_X f = 0$, for f a function (a 0-form) and,

$$(i_X\alpha)(X_1, \ldots, X_{p-1}) = \alpha(X, X_1, \ldots, X_{p-1}). \quad (2.96)$$

One finds that, as before, if α is homogeneous of degree $|\alpha|$, then,

$$i_X(\alpha \wedge \beta) = (i_X\alpha) \wedge \beta + (-1)^{|\alpha|}\alpha \wedge i_X\beta), \quad (2.97)$$

and that, for any vector fields X and Y, the graded commutator of the associated inner derivations vanishes,

$$[i_X, i_Y] = i_X \circ i_Y + i_Y \circ i_X = 0. \quad (2.98)$$

Recall that the graded commutator of two derivations is defined by

$$[D_1, D_2] = D_1 \circ D_2 - (-1)^{|D_1||D_2|} D_2 \circ D_1 .$$

As d and i_X are (graded) derivations, of degree $+1$ and -1, respectively, their graded commutator: $[d, i_X] = d \circ i_X + i_X \circ d$ is a derivation of degree zero. We denote it by $\mathcal{L}_X$ and it will be called the *Lie derivative* with respect to X. From its definition we have

$$\mathcal{L}_X = d \circ i_X + i_X \circ d . \tag{2.99}$$

On functions, the Lie derivative coincides with the action of X on $\mathcal{F}(E)$, and it extends to general forms the action of derivations on $\mathcal{F}(E)$, i.e. $\mathcal{L}_X f = X(f)$, and

$$\mathcal{L}_X(\alpha \wedge \beta) = (\mathcal{L}_X\alpha) \wedge \beta + \alpha \wedge (\mathcal{L}_X\beta) \tag{2.100}$$

Together with $d \circ d = 0$, this has the consequence that the exterior differential and the Lie derivative commute,

$$d \circ \mathcal{L}_X = d \circ i_X \circ d = \mathcal{L}_X \circ d \tag{2.101}$$

Moreover,

$$[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[X,Y]} \tag{2.102}$$

(the graded commutator of two derivations of degree 0 is again a derivation of degree 0) and finally,

$$\mathcal{L}_X \circ i_Y - i_Y \circ \mathcal{L}_X = i_{[X,Y]}. \tag{2.103}$$

In particular, when $X = Y$, then $\mathcal{L}_X \circ i_X = i_X \circ \mathcal{L}_X$.

With these ingredients at hand, one can prove (see e.g., [AM78] (Prop. 2.4.15) and [Ne67] for details, [KN63] uses a slightly different normalization) that, with $X \in \mathfrak{X}(E)$ and $\beta \in \Omega^p(E)$,

$$\begin{aligned} d\beta(X_1, \ldots, X_{p+1}) = & \sum_{i=1}^{p+1} (-1)^{i+1} X_i(\beta(X_1, \ldots, \widehat{X}_i, \ldots, X_{p+1})) \\ & + \sum_{i<j} (-1)^{i+j} \beta([X_i, X_j], X_1, \ldots, X_i, \ldots, X_j, \ldots, X_{p+1}). \end{aligned} \tag{2.104}$$

$$(\mathcal{L}_X\beta)(X_1, \ldots, X_p) = \mathcal{L}_X(\beta(X_1, \ldots, X_p) - \sum_{i=1}^{p} \beta(X_1, \ldots, [X, X_i], \ldots, X_p). \tag{2.105}$$

where the symbol $\widehat{\ }$ means that the corresponding vector field should be omitted. This formula was used by R. Palais to provide an intrinsic definition of the exterior derivative [Pa54].

In this way we have defined a sequence of maps $d\colon \Omega^p(E) \to \Omega^{p+1}(E)$, all of them denoted by d such that $d^2 = 0$. The pair $(\Omega^\bullet(E), d)$ is called a graded differential algebra.

Remark 2.12 More generally, the definition of d can be generalized to antisymmetric multilinear maps $\phi\colon \mathfrak{X}(E) \times \cdots \times \mathfrak{X}(E) \to M$ with M any vector space carrying an action of $\mathfrak{X}(E)$, i.e., a linear map $\rho\colon \mathfrak{X}(E) \to \mathrm{End}\,(M)$. In this case we would have,

$$d_\rho\phi(X_1, \ldots, X_{p+1}) = \sum_{i=1}^{p} (-1)^{i+1} \rho(X_i)(\phi(X_1, \ldots, \widehat{X}_i, \ldots, X_{p+1})) \\ + \sum_{i<j} (-1)^{i+j} \phi([X_i, X_j], X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_{p+1}). \tag{2.106}$$

We find in an analogous way that $d_\rho \circ d_\rho = 0$ iff ρ is a Lie algebra homomorphism. The exterior differential defined by equation (2.104) will be recovered when $M = \mathcal{F}(E)$ and $\rho(X)(f) = \mathcal{L}_X(f)$, the Lie derivative.

We wish to stress again that all our constructions rely only on the commutative algebra structure of $\mathcal{F}(E)$. The linearity of E never played any role, therefore our calculus will be 'insensitive' to the kind of transformations we might perform on E.

Finally let us point out that the Lie derivative can be extended to the set of vector fields, dual space of that of 1-forms, by requiring that, if $X, Y \in \mathfrak{X}(E)$ and $\alpha \in \Omega^1(E)$,

$$\mathcal{L}_X\langle\alpha, Y\rangle = \langle\mathcal{L}_X\alpha, Y\rangle + \langle\alpha, \mathcal{L}_X Y\rangle$$

and then we obtain the following definition for $\mathcal{L}_X Y$,

$$\mathcal{L}_X Y = [X, Y]$$

In fact (cfr. Eq. (2.105)), $\mathcal{L}_X\alpha$ was defined in such a way that

$$\langle\mathcal{L}_X\alpha, Y\rangle = \mathcal{L}_X\langle\alpha, Y\rangle - \langle\alpha, [X, Y]\rangle$$

from where we find that $\mathcal{L}_X Y$ is given by $\mathcal{L}_X Y = [X, Y]$.

Once that $\mathcal{L}_X$ has been defined on functions, on vector fields and on 1-forms, an extension to the space of all tensors can be obtained requiring that $\mathcal{L}_X$ be a derivation of degree zero.

2.4.3 The 'Easy' Tensorialization Principle

It should be clear now that for any linear object associated with the abstract vector space E we may think of it as realized in terms of 'applied' vectors at x, i.e., of tangent vectors, or covectors, or any other linear tensor constructed in the tangent space to E at x. Then we can transform them into tensor fields and operations depending on x. This simple statement is what we call the 'easy' tensorialization principle. We will provide now various examples of the effective use of this principle that will be of use along the rest of the text.

2.4.3.1 Linear Algebra and Tensor Calculus

We will start by geometrizing a linear map $A \in \mathrm{End}(E)$. First, we can geometrize A by considering the associated $(1, 1)$ tensor $T_A : TE \to TE$, defined as:

$$T_A\colon\ (x, v) \mapsto (x, Av)\,; \qquad x \in E, \quad v \in T_x E \cong E\,, \tag{2.107}$$

or, dually $T_A^* : T^*E \to T^*E$,

$$\langle T_A^*\alpha, v\rangle = \langle \alpha, Tv\rangle; \quad x \in E,\ v \in T_x E \cong E,\ \alpha \in T_x E^* \cong E^*\,. \tag{2.108}$$

Using linear coordinates x^k induced by a given base $\{e_k\}$,

$$T_A = A^j{}_i dx^i \otimes \frac{\partial}{\partial x^j} \tag{2.109}$$

Then, $A^i{}_k$ is given by

$$T_A(\partial/\partial x^k) = A^i{}_k\, \partial/\partial x^i\,.$$

The correspondence $A \mapsto T_A$ is an algebra homomorphism, i.e.,

$$T_{A\cdot B} = T_A \circ T_B \tag{2.110}$$

because,

$$(T_A \circ T_B)\left(\frac{\partial}{\partial x^k}\right) = T_A\left(B^i{}_k \frac{\partial}{\partial x^i}\right) = B^i{}_k A^j{}_i \frac{\partial}{\partial x^j} = (AB)^j{}_k \frac{\partial}{\partial x^j} \tag{2.111}$$

Notice that once we have geometrized the linear map A on the vector space and promoted it to a $(1, 1)$ tensor on the space E, then we are not restricted to consider linear coordinates or linear transformations. We can use any system of coordinates to describe it (even if the underlying linear structure becomes blurred) and use the exterior differential calculus as was discussed in the previous section.

The tensorialization T_A of A we have constructed retains all the algebraic properties of A which are hidden in the fact that when we express T_A in linear coordinates, the tensor is constant. That is we can de-geometrize it by choosing the appropriate set of coordinates, returning to the algebraic ground. This property will be instrumental in the general notion of tensorialization that will be discussed later on.

It is also possible, as it was shown before, to tensorialize a linear map A by associating to it a vector field X_A defined as:

$$X_A(x) = (x, Ax), \quad x \in E, \; Ax \in T_x E \cong E. \tag{2.112}$$

Now the geometrized object is not constant anymore and in linear coordinates x^k takes the form:

$$X_A = A^j{}_i \, x^i \frac{\partial}{\partial x^j}.$$

In this case we have been using the additional feature offered by linear spaces that $T_x E$ can be identified naturally with the base space E and with the linear space where A is defined. Notice that in the definition of T_A we were simply using the fact that $T_x E$ can be identified with the linear space where A is defined.

However on this occasion, the association $A \mapsto X_A$ fails to be a homomorphism of algebras and is only a Lie algebra homomorphism because $X_{AB}(f) \neq X_A(X_B(f))$, while,

$$X_{(AB-BA)}\,(f) = X_A\,(X_B\,(f)) - X_B\,(X_A\,(f)) \tag{2.113}$$

i.e.,

$$X_{[A,B]} = [X_A, X_B]\,. \tag{2.114}$$

Another interesting formula that transforms an algebraic identity into a geometrical operation is obtained by computing:

$$\mathcal{L}_{X_A} T_B = T_{AB},$$

We will see other formulae similar to the preceding one as we apply the tensorialization principle to other objects. The 'easy' tensorialization principle can also be stated as the thumb rule that transforms a tensor in a linear space E replacing e_k (the vectors of a given base) by $\partial/\partial x^k$ and the dual base elements e^j by dx^j. We can write it more precisely as follows:

The 'Easy' Tensorialization Principle

Given a tensorial object $\mathbf{t}$ in a linear space E, and given a linear base $\{e_k\}$, by replacing e_k by $\partial/\partial x^k$ and the dual base elements e^j by dx^j in the expression of $\mathbf{t}$ we will define a geometrical tensor $T_{\mathbf{t}}$ with the same structure as $\mathbf{t}$.

The 'easy' tensorialization principle as stated before could seem to depend on the choice of a system of linear basis or on a system of linear coordinates. However that is not so. The choice of a base $\{e_1, , \ldots, e_n\}$ on E (and $\{e^1, \ldots, e^n\}$ its dual basis in E^*), provides an injection of the linear space E in the set of constant vector fields (i.e., homogeneous vector fields of degree -1) given by: $v \mapsto v^i\, \partial/\partial x^i$ (with v^i being the coordinates of the vector v with respect to the given basis). We also have the injection: $\alpha = \alpha_i\, e^i \mapsto \alpha_i\, dx^i$ for the dual E^*.

This injection does not depend on the choice of the basis. Actually, we can originate the previous association by the following construction: For a fixed $x \in E$ we have a linear map $\xi_x \colon E \to E_x = T_x E$ defined associating to $v \in E$ the tangent vector to the curve $\gamma \colon \mathbb{R} \to E$ given by $t \mapsto x + vt$, that is:

$$\xi_x\,(v) = \frac{d}{dt}\,(x + vt)\,|_{t=o} \tag{2.115}$$

In other words, when acting on functions,

$$\xi_x(v) f = \frac{d}{dt} f(x + vt)\bigg|_{t=0} = (df)_x(v) = v^i\,\frac{\partial f}{\partial x^i}\bigg|_x\,, \tag{2.116}$$

i.e.

$$v \mapsto v^i\ \frac{\partial}{\partial x^i}\bigg|_x\,.$$

Then, given a bilinear pairing $b \colon E \otimes E \to \mathbb{R}$, say $B = B_{ij}\, e^i \otimes e^j$, with e^i a basis for E^*, we can apply the previous principle and associate to B the tensor,

$$T_B = B_{ij}\, dx^i \otimes dx^j\,. \tag{2.117}$$

Of course we can evaluate it pointwise on vector fields like the dilation field $\Delta = X_I$ whose expression in local coordinates is $\Delta = x^i \partial/\partial x^i$ and get $f_B(x) = \tau_B(\Delta, \Delta)\,(x) = B_{ij} x^i x^j$. By using T_A we obtain also,

$$T_A(\Delta) = X_A \tag{2.118}$$

This shows that, while X_A depends on the linear structure, T_A depends only on the affine structure of E, i.e., the tangent bundle structure.

Similarly we can geometrize an algebraic bivector. If Λ is an algebraic bivector, that is an algebraic skew symmetric contravariant tensor on E. Selecting a base $\{e_i\}$, Λ will take the form $\Lambda = \Lambda^{ij} e_i \wedge e_j$, and we may construct a tensor field on E by means of:

$$\Lambda = \Lambda^{ij} \frac{\partial}{\partial x^i} \wedge \frac{\partial}{\partial x^j} \tag{2.119}$$

In other words, bi-vectors like $\Lambda^{ij} e_i \wedge e_j$, with $\Lambda^{ij} \in \mathbb{R}$, are to be identified with the corresponding constant bivector fields.

We will discuss further extensions of this principle in Sect. 3.4.

2.4.4 Closed and Exact Forms

We say that a form $\beta \in \Omega^p(E)$ is *closed* if $d\beta = 0$, and *exact* if $\beta = d\theta$ for some $\theta \in \Omega^{p-1}(E)$. Quite obviously, by virtue of the fact that $d \circ d = 0$, every exact form is closed. The converse is the content of the following:

Proposition 2.13 (Poincare's Lemma for vector spaces) *If E is a linear vector space, then every closed form in E is exact.*

Proof For the proof, let us note preliminarily that, in view of the fact that d is'natural with respect to diffeomorphisms' i.e., as already stressed previously, $\phi^* \circ d = d \circ \phi^*$, then: (i) Once a coordinate system has been fixed, $E \approx \mathbb{R}^n$, and we might as well conduct the proof directly in $\mathbb{R}^n$, and (ii) for the same reasons, the proof will hold for any open set that is diffeomorphic to an open ball in $\mathbb{R}^n$ (and hence to $\mathbb{R}^n$ itself).

The theorem will be proved if we can construct a mapping, $T \colon \Omega^p(E) \to \Omega^{p-1}(E)$ such that, $T \circ d + d \circ T = Id$, for then, $\beta = (T \circ d + d \circ T)\beta$, and, in view of the closure of β, we see that $\beta = d\theta$ with $\theta = T\beta$.

We claim then that the required mapping is provided by,

$$(T\beta)(x) = \int_0^1 t^{p-1} (i_\Delta \beta)(tx)\, dt, \quad x \in \mathbb{R}^n, \tag{2.120}$$

where Δ denotes the dilation field in E. If β has the expression: $\beta(x) = \frac{1}{p!} \beta_{i_1 \ldots i_p}(x)\, dx^{i_1} \wedge \cdots \wedge dx^{i_p}$, the integrand has to be understood as, $\beta(tx) = \frac{1}{p!} \beta_{i_1 \ldots i_p}(tx)\, dx^{i_1} \wedge \cdots \wedge dx^{i_p}$.

Also, for any vector field X, we have that

$$\mathcal{L}_X \beta = \frac{1}{p!} \left(X^k \frac{\partial}{\partial x^k} \beta_{i_1 \ldots i_p} + p\, \frac{\partial X^k}{\partial x^{i_1}} \beta_{k, i_2 \ldots i_p} \right) dx^{i_1} \wedge \cdots \wedge dx^{i_p},$$

and

$$\begin{aligned}\frac{d}{dt}\beta(tx) &= \frac{d}{dt}\frac{1}{p!}\beta_{i_1\ldots i_p}(tx)\,dx^{i_1}\wedge\cdots\wedge dx^{i_p} \\ &= \frac{1}{t}\left(\frac{1}{p!}y^i\frac{\partial}{\partial y^i}\beta_{i_1\ldots i_p}(y)\right)_{y=tx} dx^{i_1}\wedge\cdots\wedge dx^{i_p} \\ &= \frac{1}{t}\left(y^i\frac{\partial}{\partial y^i}\beta(y)\right)_{y=tx}.\end{aligned}$$

Therefore,

$$\frac{d}{dt}\left(t^p\beta(tx)\right) = t^{p-1}\left(p\,\beta(tx) + \left[y^i\frac{\partial}{\partial y^i}\beta(y)\right]_{y=tx}\right),$$

and hence,

$$\frac{d}{dt}\left(t^p\beta(tx)\right) = t^{p-1}(\mathcal{L}_\Delta\beta)(tx).$$

But then,

$$\begin{aligned}(T\circ d + d\circ T)\beta(x) &= \int_0^1 t^{p-1}(i_\Delta\circ d + d\circ i_\Delta)\beta(tx)dt = \int_0^1 t^{p-1}(\mathcal{L}_\Delta\beta)(tx)dt \\ &= \int_0^1 \frac{d}{dt}\{t^p\beta(tx)\}dt = \beta(x).\end{aligned}$$

□

Remark 2.13 (i) From the way the proof of Poincare's Lemma has been constructed, it is quite clear that, if we consider a differential form defined on an arbitrary open set U, the relevant condition for the validity of the Lemma is that there is a point (taken above as the origin in $\mathbb{R}^n$) such that any other point in U can be joined to it by a straight line segment lying entirely in U. That's why, in an equivalent way, Poincare's Lemma is often stated with reference to 'star-shaped' open sets.

(ii) As long as we consider forms that are differentiable over the whole of E (a vector space) there will be no distinction between closed and exact forms. The situation will change however as soon as we consider topologically less trivial spaces. As a simple example, consider $\mathbb{R}^2 - \{0\}$ with (Cartesian) coordinates (x, y) and the 1-form,

$$\beta = \frac{x\,dy - y\,dx}{x^2 + y^2}. \tag{2.121}$$

It is a simple computation to show that: $\beta = d\theta$, with $\theta = \tan^{-1}(y/x)$, the polar angle. Hence: $d\beta = 0$, i.e., β is closed, but it fails to be exact, because θ is not a globally defined function.

2.5 The General 'Integration' Problem for Vector Fields

2.5.1 The Integration Problem for Vector Fields: Frobenius Theorem

After the discussion in the last few sections we have arrived to the understanding that our modeling of dynamical systems is done in terms of vector fields, i.e., first-order differential equations, or equivalently, in the algebraic setting we have started to develop, as derivations in an algebra of smooth functions. We have also emphasized that in the case of linear systems the system is described completely by means of a flow of linear maps, and in general—nonlinear—case, the uniqueness and existence theorem of solutions of initial value problems for ordinary differential equations, guarantees the existence of local flows describing the dynamical behaviour of our system. The existence of such flows, globally defined in the case of linear systems and only locally defined in the general case, does not imply that we have a simple way of computing it.

We will say that the 'integration' problem for a given dynamics Γ consists in determining its (local) flow φ_t explicitly.

Again a few remarks are in order here regarding what do we mean by the 'explicit' determination of the flow. For instance, an 'explicit' solution of the dynamics could be an approximate numerical determination of the solution given for some initial condition x_0 and a time interval $[0, T]$. Varying the initial condition x_0 using some discrete approximation on a given domain U would provide an approximate explicit description of the dynamics. Unfortunately that is the most that can be done in many occasions when dealing with arbitrary nonlinear equations, and often only after devising very clever numerical algorithms and solving a number of hard problems regarding the stability and convergence of them. Even solving the integration problem for linear systems could be a hard problem. As we know the flow of the system is given by $\varphi_t = \exp tA$, so it seems that we have a closed expression for it, hence an 'explicit' description of the dynamics. However that is not so. Of course the infinite-dimensional situation, like the ones we face when dealing with Maxwell, Schrödinger, and other systems of interest, could be very hard to analyze because the structure of the linear operator A could be difficult to grasp, but also the finite-dimensional case could have interesting features which are not displayed in the simple minded expression for the flow above. Thus we have seen that looking for constants of motion and symmetries is often quite helpful in discussing the structure of linear systems offering new and deep insights into its properties, recall for instance the discussion of the harmonic oscillator in Chap. 1. That is the existence of structures compatible with the dynamics provides useful leads to analyze it even in the linear situation.

Even more, it is a fact that many dynamical systems arising from physical theories have a rich structure (for instance symmetries, constants of motion, Hamiltonian and/or Lagrangian descriptions, etc.) and exploring such intricacies has been proved to be the best way to approach their study. Actually it happens that in some particular instances, a judiciously and in some cases, extremely clever, use of such structures leads to an 'explicit' description of the dynamics, where now 'explicit' means that the actual solutions to the dynamical equations can be obtained by direct manipulation of some algebraic quantities. In such cases the systems are usually called 'integrable' with various adjectives depending on the context. For instance if the system is Hamiltonian and it possesses a maximal number of independent commuting constants of motion, the system is called completely integrable, etc.

As it has already been stated, our interest in this work is focused in unveiling the simplest and most significative structures compatible with a given dynamics but as it will be shown along these pages, anytime that there is a new structure compatible with the dynamics, we learn something on its integration problem. To the point that in some cases we can actually integrate it. This process will be spread out along the book and it will culminate in the last two chapter where the solution of the integration problem for various classes of dynamics will be discussed at length.

We would like to close this digression pointing out another twist of the 'integration' problem that permeates some parts of this work and that has played a relevant role in the development of modern geometry. Clearly if instead of having a single dynamics, let us say now, a vector field, we had two or more, we may ask again for the determination of the solutions of all of them. To be precise, suppose that we are given vector fields $X_1, \ldots, X_r$ in some vector space E (or on an open set in it). We can integrate them locally and obtain solutions $x^{(1)}(t_1), \ldots, x^{(r)}(t_r)$ for a given common initial data x_0. Nothing fancy so far, but we may ask, can be combine the r functions above in a single one, i.e., does there exist a function $x(t_1, \ldots, t_r)$ such that it reproduces the integral curves of the vector field X_1 when we fix the parameters $t_2, \ldots, t_r$ and so on? Or, in other words, if we change the initial data x_0 moving it for instance along the solution of the flow of the first vector field the solutions we would obtain now will be compatible with the ones obtained if we move the initial data in the direction of any other vector field?

The answer to this question is the content of the so-called Frobenius theorem and provides the backbone for the theory of foliations, but again we could raise the same questions as in the case of a single vector field: can be describe 'explicitly' such collective solutions?

Theorem 2.14 (Frobenius theorem: local form) *Let $X_1, \ldots, X_r$ be a family of vector fields on an open set U of a linear space E such that the rank of the linear subspace spanned by them at each point is constant. Then, this family can be integrated in the sense before, i.e., for each point $x \in U$ there exist an open set $V \in \mathbb{R}^r$ and a smooth injective function $\varphi \colon U \subset \mathbb{R}^r \to U \subset E$ such that the local flows of the vector fields X_i are given by the curves $\varphi(c_1, \ldots, c_{i-1}, t, c_{i+1}, \ldots, c_r)$ iff $[X_i, X_j]$ can be expressed as a superposition of the vector fields X_k.*

2.5.2 Foliations and Distributions

We can get a better grasp on the meaning of the local form of Frobenius theorem, Theorem 2.14, by 'geometrizing' the notion of 'collective solutions' used before. First we will define a smooth submanifold M of $\mathbb{R}^n$ of dimension r as a subset of $\mathbb{R}^n$ such that at each point $x \in S$ there is a neighborhood U of it which is the graph of a smooth function $\Psi : V \subset \mathbb{R}^r \to \mathbb{R}^{n-r}$, and open set $V \subset \mathbb{R}^r$, that is, $x = (u, \psi(u))$ for all $x \in U \subset M$, $u \in V$ (see a detailed account of the notions of manifolds and submanifolds in Appendix C),

In particular a submanifold M of dimension r of $\mathbb{R}^n$ can be defined in virtue of the Implicit Function Theorem (10.44), as a level set of a regular value c of a map $F : \mathbb{R}^n \to \mathbb{R}^{n-r}$. If we allow c to vary in a small neighborhood of $\mathbb{R}^{n-r}$ over the range of the map F, we will generate a family of submanifolds such that one and only one such submanifold will pass through each point of $\mathbb{R}^n$ (or of the domain of F that, in case it is not the whole of $\mathbb{R}^n$, will be always assumed to be an open subset thereof), and we obtain what might be called a 'slicing' of $\mathbb{R}^n$ (or, again, of the domain of F) into a 'stack' of closely packed submanifolds. That is the basic idea of what 'foliating a manifold' will mean for us, an idea that we will try to make slightly more precise here, together with the associated notions of 'distributions' and of 'integrable distributions'.

So, let us begin with a more precise definition of what we mean by a 'foliation' of a manifold.

Definition 2.15 Let U be an open set of a linear space of dimension n. A foliation $\mathcal{L}$ of U (of codimension m) is a family $\{\mathcal{L}_\alpha\}$ of disjoint connected subsets $\mathcal{L}_\alpha$ of U (to be called from now on the leaves of the foliation), one passing through each point of M, such that the identification mapping, $i : \mathcal{L}_\alpha \to U$ is injective, and for each point $x \in U$, there exists a neighborhood V of x such that V is diffeomorphic to $\mathcal{L}_x \cap V \times B$ where $\mathcal{L}_x$ is the leaf passing through x and B is an open ball in $\mathbb{R}^m$.

Then, $\mathcal{L}$ will consist of a family of connected submanifolds, each of dimension $n - m$ which stack together to fill up U.

A generalized notion of foliation including the possibility of having leaves of different dimension could be introduced. Sometimes such foliations are called singular foliations. In what follows we are going to discuss mainly non-singular foliations, even though throughout the text, examples of singular foliations will show up (and will be discussed in its context). We will give now some simple examples of foliations, referring to the literature [AM78, AM88, BC70, MS85] for further details:

1. The simplest example of a foliation is provided by the set of the integral curves of a non-vanishing vector field. The leaves of the foliation are its integral curves. They are all one-dimensional for non-vanishing vector fields, but could be also zero-dimensional if we allow the vector field to vanish somewhere.
   For example, in: $M = \mathbb{R}^2 - \{\mathbf{0}\}$ with the usual Cartesian coordinates, the (images of the) integral curves of the Liouville field:

$$\Delta = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} \tag{2.122}$$

(or: $\Delta = r\partial/\partial r$ in polar coordinates: $x = r\cos\theta$, $y = r\sin\theta$) are the rays from the origin: ($r > 0$, $\theta = const.$). They are all one-dimensional and diffeomorphic to each other. If the origin were included, then $\mathbf{0}$ itself would be an integral curve, thus providing us with an example of a singular foliation, i.e., one whose leaves are not all of the same dimension.

2. A similar example is provided by the following construction. Consider the map,

$$\varphi\colon U = \mathbb{R}^3 \setminus \{\mathbf{0}\} \to S^2 \tag{2.123}$$

which maps all the points (in spherical polar coordinates) $(r, \theta, \phi) \in U$ to the point $n = (\theta, \phi)$ in S^2. Then: $\mathcal{L} = \left\{\varphi^{-1}(n)\right\}_{n\in S^2}$ will foliate U with leaves that are rays through the origin.

3. Consider next in:

$$M = \mathbb{R}^3 - \{(-1, 0, 0) \cup (1, 0, 0)\} \tag{2.124}$$

again with the usual Cartesian coordinates, the foliation: $\Phi = \{l_b\}_{b\in\mathbb{R}}$ whose leaves l_b are given by,[12]

$$\left[(x-1)^2 + y^2 + z^2\right]^{-1/2} - \left[(x+1)^2 + y^2 + z^2\right]^{-1/2} = b \tag{2.125}$$

This foliation is depicted in the figure below:
Being level sets, the leaves are regular submanifolds in the sense of Appendix C. They are all one-dimensional, compact and diffeomorphic to each other for $b \neq 0$. However, the leaf corresponding to $b = 0$ is the y-axis, which is again one-dimensional but non-compact.

4. If we change the relative sign in the left-hand side of Eq. (2.125) and consider instead,

$$\left[(x-1)^2 + y^2 + z^2\right]^{-1/2} + \left[(x+1)^2 + y^2 + z^2\right]^{-1/2} = b,\ b > 0 \tag{2.126}$$

we obtain the foliation that is depicted in the figure below:

Now, the leave corresponding to $b = 2$, which contains the origin, is the 'bubble-eight' (8), which is not even a submanifold of $\mathbb{R}^3$ (with the induced topology). On top of that, the leaves are not connected for $b > 2$.

[12] The leaves are essentially the equipotential surfaces of two opposite electric charges (a dipole) located at $(-1, 0, 0)$ and $(1, 0, 0)$ respectively.

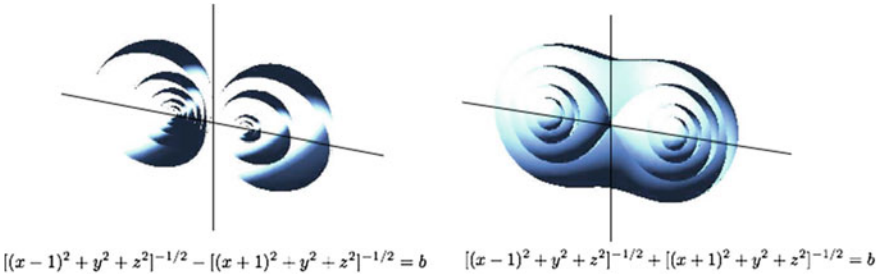

Fig. 2.1 Leaves of the foliation defined by equipotential surfaces of two charges of the same sign (*left*), and two charges of opposite sign defining a singular foliation (*right*)

The examples given above show that foliations can exhibit various kinds of pathologies. In order to avoid them, we will always consider what are called *regular* foliations [MS85], that, among other desirable features, are characterized by the leaves being all of the same dimension and all diffeomorphic to each other (Fig. 2.1).

2.5.2.1 Distributions and Integrability

Let Φ be a nonsingular foliation of M of dimension n. Then, a leaf l_α of the foliation passes through each $m \in M$. The tangent space $T_m l_\alpha$ will be a vector space of dimension n and will be spanned by a set of n vectors in $T_m M$. At least locally, i.e., in a neighborhood U of m, it will be possible to single out a set of n vector fields: $X_1, ..., X_n \in \mathfrak{X}(U)$ that span $T_m l_\alpha$ for all $m \in U$.

A *distribution* $\mathcal{D}$ will be the assignment of a similar set of vector subspaces of $T_m M$ at each point $m \in M$, all of the same dimension n, spanned in each neighborhood U by a set of smooth independent local vector fields $X_1, ..., X_n \in \mathfrak{X}(U)$. The $X_j(m)$'s, $j = 1, ..., n$ will be called a *basis* for the distribution at the point m, that will be denoted as $\mathcal{D}(m)$.

In the specific case of the distribution associated with a foliation Φ, the distribution will be denoted as $\mathcal{D}_\Phi$, and,

$$T_m l_\alpha = \mathcal{D}_\Phi(m)\,. \tag{2.127}$$

As a simple example, we may consider the foliation of M determined by the integral curves (actually the *images* in M of the integral curves) of a vector field X having no zeros, so that all the leaves of the foliation will be one-dimensional. Denoting as $\mathcal{D}_X$ the one-dimensional distribution associated with this foliation, we will have,

$$\mathcal{D}_X(m) = \mathrm{span}\,(X(m)) \equiv \{aX(m) \mid a \in \mathbb{R}\}\,. \tag{2.128}$$

It is clear that every foliation Φ defines a distribution, one that, moreover, satisfies the property expressed by Eq. (2.127). Whenever, vice versa, a distribution $\mathcal{D}$ is given satisfying the same property, i.e., we can find at every point m a submanifold l

passing through m and such that $\mathcal{D}(m)$ spans $T_m l$ or, stated otherwise, we know the right-hand side of Eq. (2.127) and we are able to solve for the left-hand side, we will say that the distribution is *integrable*, and l will be called an *integral manifold* of $\mathcal{D}$.

A distribution $\mathcal{D}$ will be said to be *involutive* if it is closed under commutation, i.e., if,

$$[X, Y] \in \mathcal{D}, \qquad \forall X, Y \in \mathcal{D}. \tag{2.129}$$

In the case of the distribution $\mathcal{D}_\Phi$ associated with the foliation Φ, the involutivity property of Eq. (2.129) is granted by the fact that (cfr. Eq. (2.127)) l_α is a (sub)manifold, and the tangent vectors to a (sub)manifold are obviously closed under commutation. Therefore integrable distributions are involutive.

The converse of this constitutes the main content of *Frobenius' theorem* [BC70, Wa71], which we will not prove here but simply state as:

Theorem 2.16 (Frobenius integrability theorem) *A distribution is integrable if and only if is involutive.*

Not all distributions need to be involutive, as the following example shows.

Consider, on $\mathbb{R}^3$, the two-dimensional distribution $\mathcal{D}$ defined (globally) by the vector fields: $X = \partial/\partial x$ and: $Y = \partial/\partial y + x\partial/\partial z$. As: $[X, Y] = \partial/\partial z \notin \mathcal{D}$, the distribution is not integrable. In fact, if it were, we could find a surface defined as a level function of a function, i.e., as: $f(x, y, z) = b$ for some $f \in \mathcal{F}\left(\mathbb{R}^3\right)$ and $b \in \mathbb{R}$ such that X and Y span the tangent space at every point, i.e., such that: $\mathcal{L}_X f = \mathcal{L}_Y f = 0$. But it is immediate to see that the only solution to these equations is: $f = const.$, i.e., no level surfaces at all.

To complete this rather long digression, we state now the conditions of the Frobenius theorem in a dual way, i.e., in terms of one-forms. Let then θ^1, ..., θ^{m-n} be a set of linearly independent one-forms, and let: $\omega = \theta^1 \wedge \ldots \wedge \theta^{m-n}$. The intersection of the kernels of the θ^j's is a distribution that will be involutive if one of the following equivalent conditions holds:

1. $\theta^i \wedge d\omega = 0, \forall i = 1, \ldots, m-n$.
2. There exists a 1-form α such that: $d\omega = \alpha \wedge \omega$.
3. There exist local one-forms α^i_j, $i, j = 1, \ldots, m-n$ such that: $d\theta^i = \alpha^i_j \wedge \theta^j$.
4. There exist functions f^i and g^i_j such that: $d\theta^i = g^i_j df^j$.

2.6 The Integration Problem for Lie Algebras

In this section we will solve the problem of integrating the Lie algebra of symmetries of a given dynamical system, rephrasing in this way the so-called Lie's third theorem. In doing so we will move forward towards the notion of Lie group.

Let us recapitulate a situation we have found in the previous sections. Let Γ be a vector field. We had seen that the collection of infinitesimal symmetries of a dynamical system is a real Lie algebra (see Sect. 3.6.3). Let us suppose that the Lie algebra of infinitesimal symmetries of Γ is finite-dimensional generated by a family of vector fields $X_1, \ldots, X_r$ such that;

$$[X_i, X_j] = c_{ij}{}^k X_k\,,$$

where $c_{ij}{}^k$ are the structure constants of the Lie algebra. The analysis of such dynamics will be done in full depth in Chap. 9. Here we will try to understand the structure of the flows of the family of vectors X_i. Because the vector fields X_i do not commute we cannot pretend to integrate their flows by using a single function $\varphi(t_1, \ldots, t_r)$ (remember the discussion on Sect. 2.5 about the simultaneous integration of a family of vector fields). However it may happen that there is another space such that the flows of the vector fields are just curves on it. It happens that such space is what we call a Lie group and we will devote the next few sections to this idea.

2.6.1 Introduction to the Theory of Lie Groups: Matrix Lie Groups

We have already found in the preceding sections some examples of groups, $GL(n, \mathbb{R})$, $U(n)$, etc. Throughout the book many other groups are going to have a relevant role, like $SU(2)$, $SO(3)$, $SL(2, \mathbb{C})$, $HW(n)$, etc. A rigorous discussion of their properties and representations would require a detailed development of the theory of Lie groups (i.e., groups that are equipped with the structure of a smooth manifold). We will not attempt to do that in this book, referring the reader to the abundant literature on the subject (see for instance [Wa71], etc.) even if we will provide an intrinsic definition of the class of Lie groups in the next chapter.

However, most of the previous examples (and many more) are groups of matrices, or as we will call them, matrix Lie groups, and contrary to the general class of Lie groups, only elementary calculus is required to discuss some of their properties and structure. This section will constitute an approximation to the theory of Lie groups by means of the study of an important family of them, closed subgroups of the general linear group.

Immediately after we will address the problem of integrating Lie algebras, the infinitesimal trace of a Lie group, arriving to the main theorem in Lie's theory that establishes a one-to-one correspondence between Lie algebras and connected and simply connected (Lie) groups.

2.6.1.1 The General Linear Group and the Orthogonal Group

Consider the general linear group in $\mathbb{R}^n$, that is the set of all invertible $n \times n$ real matrices $GL(n, \mathbb{R})$. It can be considered as an open subset in R^{n^2} by means of the map:

$$GL(n,\mathbb{R}) \to \mathbb{R}^{n^2}; \qquad A = (a_{ij}) \mapsto (a_{11}, \ldots, a_{1n}, \ldots, a_{n1} \ldots, a_{nn}) .$$

Obviously, the set $GL(n,\mathbb{R})$ is a group, because if $A, B \in GL(n,\mathbb{R})$, then $AB \in GL(n,\mathbb{R})$ and $A^{-1} \in GL(n,\mathbb{R})$. The multiplication function is differentiable, because

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

say, the elements $(AB)_{ij}$ are quadratic polynomial functions of the elements of A and B, respectively.

In all that follows we will assume that $GL/n,\mathbb{R})$ is a subset of $\mathbb{R}^{n^2}$ because of the previous identification. Because the determinant map det is continuous (is a polynomial of degree n), we get that the group $GL(n,\mathbb{R}) = \det^{-1}(\mathbb{R} - \{0\})$ is an open subset of $\mathbb{R}^{n^2}$.

If we consider now the group $O(n)$ of orthogonal matrices, $O(n) = \{R \in GL(n,\mathbb{R}) \mid R^T R = RR^T = \mathbb{I}\}$ from the orthogonality condition $R^t R = \mathbb{I}$, we get:

$$\sum_j R_{ij} R_{jk} = 0, \quad i \neq k; \quad \sum_j R_{ij}^2 = 1; \qquad i = k, \tag{2.130}$$

showing that $|R_{ij}| \leq 1$, for all i, j. The subset $O(n) \subset \mathbb{R}^{n^2}$ is closed because it is defined by a set of algebraic equation (2.130). Notice that $O(n)$ is $F^{-1}(\mathbb{I})$ where $F \colon GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$ is the smooth map $F(R) = R^T R$. Moreover $O(n)$ is bounded because $\sum_{i,j} R_{ij}^2 = n$, then $O(n)$ is compact. Notice, however that $O(n)$ is not connected because in general we just have $(\det R)^2 = 1$ and $O(n)$ has two connected components characterized by the sign of $\det R$. The connected component containing the neutral element is a normal subgroup:

$$SO(m,\mathbb{R}) = \{X \in GL(m,\mathbb{R}) \mid X^T X = I_m,\ \det A = 1\}.$$

We may now compute its tangent space as a subset of $\mathbb{R}^{n^2}$ (it is actually a submanifold, see Appendix C). Let $\gamma \colon (-\epsilon, \epsilon) \to O(n)$ be a smooth curve passing through the identity matrix, i.e., $\gamma(0) = \mathbb{I}$. Then $\gamma(t)^t \gamma(t) = \mathbb{I}$ for all t and computing the derivative at $t = 0$ we get: $\dot\gamma(0)T + \dot\gamma(0)$. Then the tangent vector $\dot\gamma(0)$ is a skew symmetric matrix. Conversely any skew symmetric matrix A is the tangent vector to a smooth curve in $O(n)$. It is enough to consider the curve $\gamma(t) = \exp tA$. We conclude that the tangent space to $O(n)$ at the identity can be identified with the set of skew symmetric matrices:

$$T_I O(n) = \{A \in M_n(\mathbb{R}) | A^T = -A\}.$$

In a similar way we can compute the tangent space to $O(n)$ at a given orthogonal matrix R. It suffices to consider a curve $\gamma(t)$ as before passing through the identity and multiply it by R on the right. Then the tangent space $T_R O(n)$ will be identified with matrices of the form AR with A skew symmetric.

The set of skew symmetric matrices $n \times n$ is a linear space of dimension $n(n-1)/2$ and we will say that the orthogonal group $O(n)$ is a manifold of dimension $n(n-1)/2$.

Definition 2.17 We will say that G is a matrix Lie group if it is an algebraic subgroup of $GL(n, \mathbb{R})$ and it is closed as a subset of $\mathbb{R}^{n^2}$.

Groups of $n \times n$ matrices with complex coefficients will be considered in a natural way as subgroups of $GL(2n, \mathbb{R})$ identifying $\mathbb{C}$ with $\mathbb{R}^2$. Thus the complex entry z_{jk} will be replaced by the 2×2 real matrix

$$\begin{pmatrix} x_{jk} & -y_{jk} \\ y_{jk} & x_{jk} \end{pmatrix}$$

with $z_{jk} = x_{jk} + i y_{jk}$.

It can be shown that because of the group law any closed subgroup of $GL(n, \mathbb{R})$ has a well defined tangent space at any point (that is, a smooth submanifold of $\mathbb{R}^{n^2}$) [MZ55, Gl52].[13] Thus the considerations we have made for $O(n)$ can be extended to any matrix Lie group.

Definition 2.18 The tangent space at the identity $T_I G$ of a matrix Lie group G will be called the Lie algebra of the group and will be denoted as $\mathfrak{g}$.

Example 2.6 Not every subgroup of $GL(n, \mathbb{R})$ is a matrix Lie group. Consider for instance the subgroup of $GL(2, \mathbb{C}) \cong GL(4, \mathbb{R})$ of matrices:

$$\begin{pmatrix} e^{it} & 0 \\ 0 & e^{i\lambda t} \end{pmatrix}$$

where $t \in \mathbb{R}$ and λ is an irrational number. It is easy to check that such subgroup is not closed in $\mathbb{R}^4$.

2.6.1.2 The Lie Algebra of a Matrix Lie Group

Definition 2.19 A Lie algebra L is a linear space with a skew symmetric bilinear map $[\cdot, \cdot] \colon L \times L \to L$ such that it satisfies Jacobi's identity:

$$[[\xi, \zeta], \chi] + [[\zeta, \chi], \xi] + [[\chi, \xi], \zeta] = 0\,,$$

[13] This is an application of a deep result in the theory of Lie groups, also-called Hilbert's fifth problem that shows that any finite-dimensional locally compact topological group without "small subgroup" is a Lie group [MZ55].

for all $\xi, \zeta, \chi \in L$.

Let L be a Lie algebra and $\mathcal{B} = \{E_i\}$ a linear basis, then we get:

$$[E_i, E_j] = c_{ij}^k E_k. \tag{2.131}$$

The constants c_{ij}^k are called the structure constants of the Lie algebra L with respect to the basis $\mathcal{B}$. It is immediate to check that the structure constants c_{ij}^k satisfy $c_{ij}^k = -c_{ji}^k$ and

$$c_{ij}^l c_{lk}^m + c_{jk}^l c_{li}^m + c_{ki}^l c_{lj}^m = 0, \qquad \forall i, j, k, m. \tag{2.132}$$

Conversely, given a family of numbers c_{ij}^k satisfying the previous conditions they are the structure constants of a unique Lie algebra with respect to some linear basis on it.

Example 2.7

1. the associative algebra $M_n(\mathbb{R})$ of $n \times n$ square matrices with real coefficients can be endowed with the Lie product $[A, B]$ given by the commutator of matrices $[A, B] = AB - BA$, and then $M_n(\mathbb{R})$ is endowed with a Lie algebra structure of dimension m^2. A basis is given by n^2 matrices E_{ij} with elements given by $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$. Each matrix $A = (a_{ij}) \in \mathcal{M}_m(\mathbb{R})$ can be written in a unique way as a linear combination

$$A = \sum_{i,j=1}^{n} a_{ij} E_{ij}.$$

   The structure constants in such a basis are

$$[E_{ij}, E_{kl}] = \delta_{jk} E_{il} - \delta_{il} E_{kj},$$

   because $E_{ij} E_{kl} = \delta_{jk} E_{il}$.
2. The cross product of vectors $x \times y$ defines a Lie algebra structure on $\mathbb{R}^3$.
3. Let $\mathcal{F}$ be the linear space of smooth functions on $\mathbb{R}^{2n}$. The bilinear map given by the standard Poisson bracket:

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^{i+n}} - \frac{\partial g}{\partial x^i} \frac{\partial f}{\partial x^{i+n}}$$

   defines a Lie algebra structure on $\mathcal{F}$.

Let now G be a matrix Lie group and consider the map $\Phi \colon G \times \mathfrak{g} \to \mathfrak{g}$, defined as

$$\Phi(g, \xi) = \frac{d}{dt} g \cdot \gamma(t) \cdot g^{-1} \mid_{t=0} = g \cdot \xi \cdot g^{-1}, \tag{2.133}$$

where $\dot{\gamma}(0) = \xi$ and $g \in G$. This map defines an action of G on $\mathfrak{g}$, that is, it satisfies: $\Phi(g, \Phi(h, \xi)) = \Phi(gh, \xi)$, $\Phi(\mathbb{I}, \xi) = \xi$, called the adjoint action of G on its Lie algebra $\mathfrak{g}$.

We will denote by $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ the linear map $\mathrm{Ad}_g(\xi) = \Phi(g, \xi) = g \cdot \xi \cdot g^{-1}$. Moreover $\mathrm{Ad}_g \circ \mathrm{Ad}_h = \mathrm{Ad}_{gh}$. Then the adjoint action defines a linear representation of G as linear maps on its Lie algebra.

We have that the tangent space at the identity of a matrix Lie group is a Lie algebra (hence the name).

Proposition 2.20 *Let G be a matrix Lie group. The tangent space at the identity $\mathfrak{g} = T_{\mathbb{I}}G$ is a Lie algebra with respect to the commutator of matrices.*

Proof Let $\xi, \zeta \in \mathfrak{g}$ and $g(t) : (-\epsilon, \epsilon) \to G$ be a smooth curve such that $\dot{g}(0) = \xi$. Then the curve $\sigma(t) = \mathrm{Ad}_{g(t)}\zeta$ is in $\mathfrak{g}$ and $\sigma(0) = \zeta$. Computing the derivative of $\sigma(t)$ at 0, we get:

$$\dot{\sigma}(0) = \frac{d}{dt}(g(t).\zeta.g(t)^{-1}) \mid_{t=0} = \xi \cdot \zeta - \zeta \cdot \xi = [\xi, \zeta],$$

where we have used that $d(g(t)^{-1})/dt = -g(t)^{-1} \cdot (dg(t)/dt) \cdot g(t)^{-1}$. □

Example 2.8 The following list describes some relevant groups of matrices:

1. $SO(n) = \{R \in GL(n, \mathbb{R}) | RR^t = I, \det R = 1\}$.
2. $U(n) = \{U \in GL(n, \mathbb{C}) | UU^\dagger = I\}$.
3. $SU(n) = \{ U \in GL(n, \mathbb{C}) | UU^\dagger = I, \det U = 1 \}$.
4. $SL(n, \mathbb{R}) = \{ S \in GL(n, \mathbb{R}) | \det S = 1 \}$.

and their Lie algebras:

1. $\mathfrak{so}(n) = \{ A \in M_n(\mathbb{R}) | A^t = -A, \mathrm{Tr}\, A = 0 \}$.
2. $\mathfrak{u}(n) = \{ V \in M_n(\mathbb{C}) | V^\dagger = -V \}$.
3. $\mathfrak{su}(n) = \{ V \in M_n(\mathbb{C}) | V^\dagger = -V, \mathrm{Tr}\, V = 0 \}$.
4. $\mathfrak{sl}(n) = \{ A \in M_n(\mathbb{R}) | \mathrm{Tr}\, A = 0 \}$.

As a consequence we obtain that their dimensions are:

1. $\dim SO(n) = n(n-1)/2$.
2. $\dim U(n) = n^2$.
3. $\dim SU(n) = n^2 - 1$.
4. $\dim SL(n, \mathbb{R}) = n^2 - 1$.

We will devote the next few sections to work out in detail the Lie algebras and other properties of some groups that are of capital importance.

2.6.1.3 The Lie Algebra of $SO(3)$ and $SU(2)$

The Lie algebra of $SO(3)$ will be obtained by computing the tangent vectors to smooth curves passing through $\mathbb{I}$. We consider the rotations around the axis:

$$R(e_1,\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad R(e_2,\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix},$$

and

$$R(e_3,\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Denoting by M_i the tangent vector to $R(e_i,\theta)$ at $\theta = 0$, we get:

$$M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.134}$$

and we check immediately:

$$[M_1, M_2] = M_3, \quad [M_2, M_3] = M_1, \quad [M_3, M_1] = M_2$$

that will be written as:

$$[M_i, M_j] = \epsilon_{ijk} M_k. \tag{2.135}$$

Conversely, if A is in the Lie algebra, i.e., is a skew-symmetric 3×3 matrix, we get:

$$A = \begin{pmatrix} 0 & -\zeta_3 & \zeta_2 \\ \zeta_3 & 0 & -\zeta_1 \\ -\zeta_2 & \zeta_1 & 0 \end{pmatrix} \tag{2.136}$$

that can be written as: $A = \zeta_1 M_1 + \zeta_2 M_2 + \zeta_3 M_3$. It is clear that this construction generalizes immediately to $SO(n,\mathbb{R})$ with $n \geq 3$.

To obtain the Lie algebra of $SU(2)$ it is sufficient to consider the curves:

$$U(e_k,\varphi) = \cos\varphi - i\sigma_k \sin\varphi \tag{2.137}$$

where σ_k denotes Pauli's sigma matrices:

$$\sigma_0 = \mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.138}$$

Then we get:

$$U(e_1, \varphi) = \begin{pmatrix} \cos\varphi & -i\sin\varphi \\ -i\sin\varphi & \cos\varphi \end{pmatrix}, \quad \frac{dU(e_1,\varphi)}{d\varphi}\Big|_{\varphi=0} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = N_1 = -i\sigma_1 \tag{2.139}$$

$$U(e_2, \varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \quad \frac{dU(e_2,\varphi)}{d\varphi}\Big|_{\varphi=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = N_2 = -i\sigma_2 \tag{2.140}$$

$$U(e_3, \varphi) = \begin{pmatrix} e^{-i\varphi} & 0 \\ 0 & e^{i\varphi} \end{pmatrix} \quad \frac{dU(e_3,\varphi)}{d\varphi}\Big|_{\varphi=0} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = N_3 = -i\sigma_3 \tag{2.141}$$

and the Lie algebra $\mathfrak{su}(2)$ of $SU(2)$ is given by:

$$[N_1, N_2] = 2N_3, \quad [N_2, N_3] = 2N_1, \quad [N_3, N_1] = 2N_2\,. \tag{2.142}$$

A natural basis for the Lie algebra $\mathfrak{su}(2)$ consists of the matrices $-i\sigma_k$ tangent to the curves (2.139). We realize immediately that the Lie algebra $\mathfrak{su}(2)$ is isomorphic to the Lie algebra $\mathfrak{so}(3)$ of $SO(3)$; the isomorphism $\eta\colon \mathfrak{so}(3) \to \mathfrak{su}(2)$, is given by:

$$\eta(M_i) = \frac{1}{2}N_i = -\frac{i}{2}\sigma_i\,. \tag{2.143}$$

However the Lie groups $SU(2)$ and $SO(3)$ are not isomorphic because their topological properties are different and if they were they should be homeomorphic, but it is easy to see that $SU(2)$ is simply connected by identifying it with the 3-dimensional sphere while this is not the case of $SO(3)$ (see below). However they are locally isomorphic, they have the same Lie algebra.

The general theory of Lie groups shows that $SU(2)$ is the universal covering (see Sect. 2.6, Theorem. 2.25) of all Lie groups with Lie algebra isomorphic to $\mathfrak{su}(2)$. Any other group possessing the same Lie algebra can be obtained as a quotient group of $SU(2)$ by a central discrete subgroup.

In our case because the center of $SU(2)$ is $\mathbb{Z}_2$, the only two groups with the same Lie algebra are $SO(3)$ and $SU(2)$.

Exercise 2.9 Compute the center of $SU(2)$ and $SO(3)$. Prove that a central subgroup is a subgroup of the center of the group.

The covering map $\pi\colon SU(2) \to SO(3)$ is defined as follows: let x be a vector in $\mathbb{R}^3$ and $x \cdot \boldsymbol{\sigma}$ the 2×2 Hermitean matrix:

$$x \cdot \boldsymbol{\sigma} = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3 = \begin{pmatrix} x_3 & x_1 + ix_2 \\ x_1 - ix_2 & -x_3 \end{pmatrix}.$$

The map $x \mapsto x \cdot \boldsymbol{\sigma}$ defines a one-to-one correspondence between $\mathbb{R}^3$ and the linear space of traceless 2×2 Hermitean matrices. Then we define:

$$(\pi(U)\, x) \cdot \boldsymbol{\sigma} = U(\, x \cdot \boldsymbol{\sigma})U^{\dagger}, \quad \forall\, x \in \mathbb{R}^3. \tag{2.144}$$

Exercise 2.10 Check that $||\pi(U)\, x|| = ||\, x||$, and $\det \pi(U) = 1$, hence $\pi(U) \in SO(3)$.

2.6.1.4 More Examples: The Euclidean Group in 2-Dimensions and the Galilei Group

If we consider the Euclidean group of transformations in two dimensions $E(2)$,

$$\begin{aligned} x_1' &= x_1 \cos\varphi - x_2 \sin\varphi + a_1, \\ x_2' &= x_1 \sin\varphi + x_2 \cos\varphi + a_2, \end{aligned} \tag{2.145}$$

it is a Lie group of dimension three for which the composition law is

$$(\mathbf{a}', \varphi') \cdot (\mathbf{a}, \varphi) = (\mathbf{a}' + R(\varphi')\mathbf{a}, \varphi' + \varphi).$$

These transformations can be written in a matrix form as

$$\begin{pmatrix} x_1' \\ x_2' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi & a_1 \\ \sin\varphi & \cos\varphi & a_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}.$$

Hence, the infinitesimal generators are just the matrices

$$J = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

with commutation defining relations for the Lie algebra:

$$[J, P_1] = P_2, \quad [J, P_2] = -P_1, \quad [P_1, P_2] = 0.$$

Another interesting example is the Galilei group. We can identify it with a subgroup of $GL(5, \mathbb{R})$ (or the corresponding affine group in four dimensions)

$$\begin{pmatrix} \mathbf{x}' \\ t' \\ 1 \end{pmatrix} = \begin{pmatrix} R & \mathbf{v} & \mathbf{a} \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ t \\ 1 \end{pmatrix}.$$

The commutation relations defining the Lie algebra of Galilei group are then,

$$
\begin{aligned}
[\mathbf{J}, \mathbf{J}] = \mathbf{J}, \quad [\mathbf{J}, \mathbf{K}] = \mathbf{K}, \quad [\mathbf{J}, \mathbf{P}] = \mathbf{P}, \quad [\mathbf{J}, H] = 0 \,, \\
[\mathbf{K}, \mathbf{K}] = 0, \quad [\mathbf{K}, \mathbf{P}] = 0, \quad [\mathbf{K}, H] = \mathbf{P} \,, \\
[\mathbf{P}, \mathbf{P}] = 0, \quad [\mathbf{P}, H] = 0 \,. \qquad (2.146)
\end{aligned}
$$

Here $\mathbf{P}$ are the generators of the one-parameter groups of space translations, H is that of time translations, $\mathbf{J}$ are the generators of proper rotations and $\mathbf{K}$ those of pure Galilei transformations.

Finally, when considering one-parameter groups of transformations of an affine space M, for instance e^{tA}, each point $x \in M$, transforms into $x' = \Phi(e^{tA}, x)$, and for small values of the parameter t, which we will denote by ϵ,

$$
x'^i = x^i + \epsilon \xi^i(x) + O(\epsilon^2) \,,
$$

and

$$
\xi(x) = \left(\frac{d\Phi(e^{\epsilon A}, x)}{d\epsilon} \right) |_{\epsilon=0} \quad .
$$

For instance, for the one-parameter group of translations in the x_1 direction for the case of the Euclidean group in two dimensions, $\xi^1 = 1$, $\xi^2 = 0$, while for the one-parameter group of translations in the other direction, $\xi^1 = 0$, $\xi^2 = 1$. For the proper rotation subgroup, $\xi^1 = -x_2$, $\xi^2 = x_1$.

In an analogous way, in the group of proper rotations in three dimensions, for the subgroup of rotations around the axis determined by the vector $\mathbf{n}$, $\xi^i = \varepsilon^{ijk} n_j x_k$.

2.6.1.5 Group Homomorphisms and Lie Algebras

Definition 2.21 A homomorphism between the Lie algebras $(\mathfrak{g}_1, [,]_1)$ and $(\mathfrak{g}_2, [,]_2)$, is a linear map $\phi \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ such that: $\phi([\xi, \zeta]_1) = [\phi(\xi), \phi(\zeta)]_2$, for all $\xi, \zeta \in \mathfrak{g}_1$. If the homomorphism is bijective we will call it an isomorphism.

Example 2.11

1. The Lie algebras $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$ are isomorphic, the isomorphism given by equation (2.143).
2. The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is isomorphic to the complexification $\mathfrak{sl}(2, \mathbb{R})^{\mathbb{C}}$ of the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$. (The complexification $\mathfrak{g}^{\mathbb{C}}$ of a real Lie algebra $\mathfrak{g}$ is the natural Lie algebra structure obtained on the complexification of the linear space $\mathfrak{g}$ by extending the bilinear map $[\cdot, \cdot]$ to a complex bilinear map.)
3. The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is isomorphic to the complexification of the Lie algebras $\mathfrak{so}(3)$ and $\mathfrak{so}(1, 2)$ (where $\mathfrak{so}(1, 2)$ is the Lie group of the linear isomorphisms preserving the metric with signature $(-++)$).

Definition 2.22 Given two matrix Lie groups G_1 and G_2, a Lie group homomorphism between them is a smooth group homomorphism $\psi: G_1 \to G_2$ (i.e., $\psi(gh) = \psi(g)\psi(h)$ for all $g, h \in G$).

Notice that if $\psi: G_1 \to G_2$ is a differentiable map, the differential of this map at $\mathbb{I}$ is a linear map, $d\psi(\mathbb{I}): T_{\mathbb{I}}G_1 \to T_{\mathbb{I}}G_2$, that is, a linear map between the corresponding Lie algebras. We will denote in what follows the map $d\psi(\mathbb{I})$ as ψ_* and we will check that it is a Lie algebra homomorphism.

Actually, one-parameter Lie subgroups, described by curves $\gamma: \mathbb{R} \to G$ which are a group homomorphism, namely, such that

$$\gamma(t_1)\gamma(t_2) = \gamma(t_1 + t_2) ,$$

play a relevant role. Indeed, this last property means that $\gamma(t)$ is determined by the tangent vector to the curve in the neutral element, $\gamma(0) = e \in G$. When G is a subgroup of $GL(m, \mathbb{R})$, if A is the matrix

$$A = \frac{d}{dt}\gamma(t)_{|t=0} ,$$

then $\gamma(t) = e^{tA}$. In fact, it suffices to take into account the relation $\gamma(t_1)\gamma(t_2) = \gamma(t_1 + t_2)$, and to take derivative with respect to t_1 at $t_1 = 0$, and then we find $A\gamma(t) = \dot{\gamma}(t)$, and as $\gamma(0) = I$, we obtain $\gamma(t) = e^{tA}$.

Thus, the matrices A obtained as tangent vectors to one-parameter subgroups of $GL(m, \mathbb{R})$ at the identity matrix, close on the Lie algebra $\mathfrak{gl}(n, \mathbb{R})$, and those corresponding to Lie subgroups of $GL(m, \mathbb{R})$ are Lie subalgebras, i.e. they are linear subspaces stable under the Lie product.

By using exponentiation we can obtain the elements in a neighbourhood of $I \in G$, and these are generators of G when it is connected. For instance, the set of all traceless matrices is a linear space and the commutator of two traceless matrices is also traceless. They determine a Lie subalgebra, usually denoted $\mathfrak{sl}(n, \mathbb{R})$ and by exponentiation of these matrices we obtain the elements in the subgroup $SL(m, \mathbb{R})$.

The exponential map for an arbitrary matrix Lie group G is defined as the map $\exp: \mathfrak{g} \to G$ given by the standard exponential function of matrices. In more abstract terms, we would use the correspondence above between one-parameter subgroups and elements in the Lie algebra to define the exponential, that is, if $\xi \in \mathfrak{g}$ and $\gamma_\xi(t)$ is the corresponding one-parameter subgroup, then $\exp t\xi = \gamma_\xi(t)$ for all $t \in \mathbb{R}$.

It is not hard to see that the exponential map is surjective in any compact group but in non-compact groups is usually not surjective.

Exercise 2.12 Prove that the exponential map is not surjective for $SL(n, \mathbb{R})$ but it is surjective for $GL(n, \mathbb{C})$.

Because $\mathrm{Ad}_{\exp t\xi}\zeta = e^{t\xi}\zeta e^{-t\xi}$, computing the derivative with respect to t, we get:

$$\frac{d}{dt}\,_{|t=0}\,(\mathrm{Ad}_{\exp t\xi}\zeta) = [\xi, \zeta] = \mathrm{ad}\,(\xi)\zeta ,$$

with $\mathrm{ad}\,(\xi)\zeta = [\xi, \zeta]$. Thus we get:

$$\mathrm{Ad}_{\exp\xi} = \exp(\mathrm{ad}\,\xi), \qquad \forall\xi \in \mathfrak{g}\,.$$

Let us compute now the differential of the exponential map. Let $\xi\colon (-\epsilon, \epsilon) \to \mathfrak{g}$ a smooth curve and we denote by $\delta\xi(t) = d\xi(t)/dt$. The differential of exp on the tangent vector $\delta\xi(t)$ is by definition $\exp_*(\delta\xi(t)) = d(\exp\xi(t))/dt$, thus $\exp_*\colon T_\xi\mathfrak{g} \to T_{\exp\xi}G$. It is a simple exercise to check that:

$$\frac{d}{dt}e^{\xi} = \int_0^1 e^{s\xi}\frac{d\xi}{dt}e^{(1-s)\xi}ds.$$

Then

$$\begin{aligned}\exp_*(\delta\xi) &= \frac{d\exp\xi}{dt} = \int_0^1 e^{s\xi}\frac{d\xi}{dt}e^{(1-s)\xi}ds = \left(\int_0^1 e^{s\xi}\frac{d\xi}{dt}e^{-s\xi}\right)e^{\xi}\\ &= \left(\int_0^1 e^{s\,\mathrm{ad}\,\xi}ds\right)\frac{d\xi}{dt}e^{\xi} = \sum_{k=0}^{\infty}\frac{1}{(k+1)!}(\mathrm{ad}\,\xi)^k\frac{d\xi}{dt}e^{\xi} = F(\mathrm{ad}\,\xi)(\delta\xi)e^{\xi}\end{aligned}$$

Then:

$$\exp_*(\delta\xi)e^{-\xi} = F(\mathrm{ad}\,\xi)(\delta\xi). \tag{2.147}$$

with $F(x) = (e^x - 1)/x$. Notice that if $\xi(0) = 0$ and we evaluate the previous formula at $t = 0$, we get $\exp_*(0) = \mathrm{Id}$ which shows that the exponential map is a local diffeomorphism.

The next two propositions will provide more information on the relation between matrix Lie groups and their Lie algebras.

Proposition 2.23 *Let $\psi\colon G_1 \to G_2$ be a homomorphisms of matrix Lie groups, then the differential at the identity $\psi_*\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ at the identity is a homomorphism of Lie algebras.*

Proof The proof is simple. Consider two vectors ξ and ζ in $\mathfrak{g}_1$ with integral curves $g(t)$ and $h(t)$ respectively. Then:

$$\frac{d}{dt}\psi_*(\mathrm{Ad}_{g(t)}\zeta)\mid_{t=0} = [\psi_*(\xi), \psi_*(\zeta)]_2.$$

On the other hand, because the differential is linear, we will get that the previous expression is equal to:

$$\psi_* \frac{d}{dt}(Ad_{g(t)}\zeta) \mid_{t=0}$$

and computing it again we get: $\psi_*([\xi, \zeta]_1)$. □

Thus, associated to any group homomorphism there is a homomorphism between the corresponding Lie algebras. This relation can be qualified further because of the following theorem that we establish without proof (see [Wa71]):

Theorem 2.24 *Let $\psi\colon G \to H$ be a homomorphism of matrix Lie groups and let $\psi_*\colon \mathfrak{g} \to \mathfrak{h}$, be the corresponding Lie algebras homomorphism. Then:*

i. If ψ_* is onto, then ψ is an onto on H_0 (the connected component of H containing $\mathbb{I}$).
ii. If ψ_* is mono, then ψ is mono in a neighborhood of $\mathbb{I}$ in G.
iii. If ψ_* is bijective, then ψ is a local isomorphism between G_0 and H_0.

*2.6.2 The Integration Problem for Lie Algebras**

Now we are ready to prove Lie's third theorem that provides the solution for the integration of a finite-dimensional Lie algebra of vector fields. The global object that integrates a Lie algebra is a Lie group.

Theorem 2.25 (Lie's integration theorem) *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra, then there exists a unique, up to isomorphisms, connected and simply connected Lie group G whose Lie algebra is $\mathfrak{g}$.*

It is interesting to notice that after more than one hundred years since Lie's construction there is not an 'easy' proof of this theorem. The simplest way to address it (and its meaning) is to use Ado's theorem first [Ja79]. Ado's theorem establishes that any finite-dimensional Lie algebra can be seen as a subalgebra of the Lie algebra $M_n(\mathbb{R})$ of $n \times n$ matrices for some n. Then we can try to work inside the general linear group $GL(n, \mathbb{R})$. It is important to notice that Ado's theorem does not extend to Lie groups, in other words, not every Lie group is a subgroup of the general linear group (recall Example 2.6). At the end of this section we will comment a bit on an 'intrinsic' proof of Lie's theorem without recurring to Ado's theorem that will be significant later on.

2.6.2.1 Proving Lie's Theorem I: Using Ado's Theorem

To address the proof of Lie's theorem without using more sophisticated tools, we may rely on Ado's theorem stating that any (finite-dimensional) Lie algebra is isomorphic to a subalgebra of the Lie algebra of $n \times n$ real matrices for some n. We will denote such Lie algebra as $\mathfrak{gl}(n)$, and Ado's theorem can be restated saying that given a finite-dimensional Lie algebra $\mathfrak{g}$ there exist n and an injective homomorphism of Lie algebras $i: \mathfrak{g} \to \mathfrak{gl}(n)$. In what follows we will identify $\mathfrak{g}$ with its image $i(\mathfrak{g}) \subset \mathfrak{gl}(n)$ and we will not distinguish between the element $\xi \in \mathfrak{g}$ and the $n \times n$ matrix $i(\xi)$.

We will follow now the arguments in [Mi83b]. We consider the collection of all smooth maps:

$$\xi: [0, 1] \to \mathfrak{g} \subset \mathfrak{gl}(n)$$

such that $\xi(0) = \xi(1) = \xi'(0) = \xi'(1) = 0$. Given any such map $\xi(t)$ we may integrate the time-dependent linear dynamical system defined on the space of $n \times n$ matrices:

$$\frac{d\varphi}{dt} = \xi(t)\varphi(t), \qquad \varphi(0) = 1. \tag{2.148}$$

As we know from the general analysis of linear systems, such an initial value problem has a unique solution $\varphi^{\xi}: [0, 1] \to GL(n, \mathbb{R})$. The invertible matrices $\varphi^{\xi}(t)$ are constructed from a family of fundamental solutions $x^{(i)}$ of the system $dx/dt = \xi(t)x$ with initial conditions $x^{(i)}(0) = e_i$ (see Sect. 2.2.2). Then $\varphi^{\xi}(t)$ is just the matrix whose columns are the vectors $x^{(i)}(t)$ solutions of the previous initial value problem.

We notice now that the elements $\varphi^{\xi}(1) \in GL(n, \mathbb{R})$ satisfy:

$$\varphi^{\xi}(1)\varphi^{\zeta}(1) = \varphi^{\xi\star\zeta}(1), \tag{2.149}$$

where $\xi \star \zeta$ is the concatenation of the paths ξ and ζ on $\mathfrak{g}$, that is:

$$\xi \star \zeta(t) = \begin{cases} 2\zeta(2t) & \text{if } 0 \le t \le 1/2 \\ 2\xi(2t-1) & \text{if } 1/2 \le t \le 1 \end{cases}.$$

To prove that, check by direct substitution that the curve:

$$\varphi(t) = \begin{cases} \varphi^{\zeta}(2t) & \text{if } 0 \le t \le 1/2 \\ \varphi^{\xi}(2t-1)\varphi^{\zeta}(1) & \text{if } 1/2 \le t \le 1 \end{cases}$$

satisfies that $d\varphi/dt = (\xi \star \zeta)\varphi$.

Now consider the space of equivalence classes of smooth maps as before $\xi: [0, 1] \to \mathfrak{g}$ where

$$\xi \sim \zeta \quad \text{iff} \quad \varphi^{\xi}(1) = \varphi^{\zeta}(1). \tag{2.150}$$

Let us denote by G such space. We will show now that G is the object integrating $\mathfrak{g}$. First we check that we can define a composition law in G as follows: if g^{ξ}, g^{ζ} denote two equivalence classes of paths on $\mathfrak{g}$ with representatives ξ and ζ respectively, then we define:

$$g^{\xi} \cdot g^{\zeta} = g^{\xi \star \zeta}.$$

Notice that this composition law is well defined because of Eq. (2.149). This composition law is associative. The proof requires some work because $\xi \star (\zeta \star \eta)$ is not equal to $(\xi \star \zeta) \star \eta$, however we know that they are homotopic which is enough to guarantee that their equivalence classes with respect to $\sim$ are the same. There is a neutral element $e = g^0$ corresponding to the trivial curve 0 on $\mathfrak{g}$ and each element g^{ξ} has an inverse element $g^{\xi^{-1}}$ where ξ^{-1} is the opposite to the path ξ. Thus the set G becomes a group.

The set G inherits a natural topology from the topology of the space of paths ξ (for instance that induced by the supremum norm $||\xi||_{\infty} = \{||\xi(t)|| \mid 0 \leq t \leq 1\}$ and $||\cdot||$ any norm in $\mathfrak{g}$), and the composition law as well as taking the inverse are continuous with respect to this topology. Notice that G is trivially simply connected because $\mathfrak{g}$ is. In this way G becomes a topological group, however we are not interested in this approach as we want to construct directly a local identification of G with $\mathfrak{g}$.

Given a matrix A we have defined $\exp A$. Similarly we can define $\ln A$ (that will be uniquely determined provided that A is close enough to the identity matrix).[14] If the map $\xi(t)$ is small enough (that is $||\xi(t)|| < \epsilon$ for some $\epsilon > 0$ and a norm $||\cdot||$ in $\mathfrak{g}$), then $\varphi^{\xi}(t)$ will be close enough to I for all $t \in [0, 1]$. We want to check that now $\ln \varphi^{\xi}(t)$ is in $\mathfrak{g}$ for all t. Once we do that we have identified a neighborhood of the identity element e in G with a neighborhood of $\mathbf{0}$ in $\mathfrak{g}$.

To check that $\ln \varphi^{\xi}$ is in $\mathfrak{g}$ we will compute its derivative. Taking $A = \ln \varphi^{\xi}$ in Eq. (2.147) we obtain:

$$\xi(t) = F(\mathrm{ad}\, A - I)(\delta A) \in \mathfrak{g}$$

as we wanted to show.

2.6.2.2 Proving Lie's Theorem II: Extending Lie Algebra Homomorphisms

Once we have constructed G out of $\mathfrak{g}$ we would like to understand how we can construct a homomorphism $f : G_1 \to G_2$ of the groups G_1 and G_2 obtained by the procedure above from the Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$, that 'integrates' a homomorphism $\alpha : \mathfrak{g}_1 \to \mathfrak{g}_2$,

If we consider the path $g_1(t) : [0, 1] \to G_1$, we define 'tangent' path $t \mapsto \xi_1(t) = \gamma_1'(t) g_1^{-1}(t) \in \mathfrak{g}_1$. Now we take its image under α, i.e., $t \mapsto \xi_2(t) = \alpha(\xi_1(t)) \in \mathfrak{g}_2$. Then we solve the differential equation on G_2:

$$\frac{dg_2}{dt} = \xi_2(t)\varphi_2, \qquad g_2(0) = I.$$

[14] The map $\exp : \mathfrak{gl}(n) \to GL(n)$ is differentiable with differential the identity at I, hence by the inverse theorem there is local inverse of exp which is differentiable, that is the map ln we are using.

Then we define $f(g(1)) = g_2(1)$ and the proof finishes if we show that $g_2(1)$ does not depend on the path $g_1(t)$ (we could have worked similar formulae and conclusions using a representative ξ_1 for g_1 in the space of paths in $\mathfrak{g}_1$).

Now if we have two different paths $g_1(t)$ and $g_1'(t)$ ending in the same point $g \in G_1$, then because G_1 is simply connected g_1 and g_2 are homotopic, that is there exists a family of paths $g(t, s)$ all from e to a fixed element $g \in G_1$ such that $g(t, 0) = g_1(t)$ and $g(t, 1) = g_2(t)$. Then we have the tangent vectors:

$$X_1(t, s) = \frac{\partial g(t, s)}{\partial t} g(t, s)^{-1}, \quad Y_1(t, s) = \frac{\partial g(t, s)}{\partial s} g(t, s)^{-1},$$

and after a simple computation we get:

$$\frac{\partial X_1}{\partial s} - \frac{\partial Y_1}{\partial t} = [X_1, Y_1]. \tag{2.151}$$

Then define $X_2 = \alpha(X_1)$ and $Y_2 = \alpha(Y_1)$. Then because α is linear and is a Lie algebras homomorphism we get for X_2 and Y_2:

$$\frac{\partial X_2}{\partial s} - \frac{\partial Y_2}{\partial t} = [X_2, Y_2].$$

But these equations are just the compatibility conditions for the system of linear equations:

$$\frac{\partial g'(t, s)}{\partial t} = X_2(t, s) g'(t, s), \quad \frac{\partial g'(t, s)}{\partial s} = Y_2(t, s) g'(t, s)$$

hence this system has a solution $g' \colon [0, 1] \times [0, 1] \to G_2$ that shows that $f(g)$ is well defined (notice that $Y_1(t, 1) = 0$, then $Y_2(t, 0) = 0$ for all t, then $g'(t, 1)$ is constant.

Thus we may state:

Theorem 2.26 *Let $\alpha \colon \mathfrak{g}_1 \to \mathfrak{g}_2$ be a homomorphism of Lie algebras, then if G_1 and G_2 are two Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ respectively, there exists a homomorphism of Lie groups $\psi \colon G_1 \to G_2$ such that $\alpha = \psi_*$.*

2.6.2.3 Proving Lie's Theorem III: The Hard Proof, Not Using Ado's Theorem

Now it is easy to devise how we can avoid Ado's theorem in the proof of Lie's theorem. We were using the realization of the Lie algebra $\mathfrak{g}$ as a subalgebra of the algebra of matrices $\mathfrak{gl}(n)$ to define the equivalence relation (2.150) via the explicit integration of the linear system Eq. (2.148). owever this can be replaced by simply asking that the two paths $\xi(t)$ and $\zeta(t)$ on $\mathfrak{g}$ are equivalent if they can be joined by a curve of paths $X(s,t)$ such that it satisfies the compatibility equations above Eq. (2.151). Then the quotient space of paths module this equivalence relation will give us the group G as before. Again the hardest step in finishing the proof is to show that locally G is like $\mathfrak{g}$. Again we have to compute $\ln \varphi^{\xi}$ and we can proceed along similar lines as we did before (for that we need to show that the formula for the differential of the exponential still makes sense but we will not insist on this here).

Remark 2.14 It is pertinent to notice here that this way of approaching the integration of a Lie algebra has been continued in proving a much harder integration problem, that of integrating a Poisson structure solved by Crainic and Fernandez [CF04]. Then the compatibility condition is substituted by a more involved condition but the spirit is the same.

References

[MS85] Marmo, G., Saletan, E.J., Simoni, A., Vitale, B.: Dynamical Systems: A Differential Geometric Approach to Symmetry and Reduction. John Wiley, Chichester (1985)
[Ar73] Arnol'd, V.I.: Ordinary Differential Equations. MIT Press, Cambridge, 4th printing (1985).
[HS74] Hirsch, M., Smale, S.: Differential Equations, Dynamical Systems, and Linear Algebra. Academic Press Inc., New York (1974)
[DF79] Dollard, J.C., Friedman, C.N.: Product Integrals. Addison-Wesley, Cambridge (1979)
[ZK93] Zhu, J.C., Klauder, J.R.: Classical symptoms of quantum illness. Amer. J. Phys. **61**, 605–611 (1993)
[Pa59] Palais, R.: Natural operations on differential forms. Trans. Amer. Math. Soc. **92**, 125–141 (1959)
[AM78] Abraham, R., Marsden, J.E.: Foundations of Mechanics, 2nd edn. Benjamin, Reading (1978).
[Ne67] Nelson, E.: Tensor Analysis. Princeton University Press, New Jersey (1967)
[KN63] Kobayashi, S., Nomizu K.: Foundations of Differential Geometry, vol. 2. Interscience, New York (1963, 1969).
[Pa54] Palais, R.: A definition of the exterior derivative in terms of Lie derivatives. Proc. Amer. Math. Soc. **5**, 902–908 (1954)
[AM88] Abraham, R., Marsden, J.E., Ratiu, T.: Manifolds, Tensor Analysis and Applications, 2nd edn. Springer, New York (1988)
[BC70] Brickell, F., Clarke, R.S.: Foundations of Differentiable Manifolds (An Introduction). Van Nostrand, Reinhold (1970)
[Wa71] Warner, F.: Foundation of Differentiable Manifolds and Lie Groups. Scott Foresman, New York (1971)
[MZ55] Montgomery, D., Zippin, L.: Topological Transformation Groups, vol. 1. Interscience Publishers, New York (1955)
[Gl52] Gleason, A.M.: Groups without small subgroups. Ann. Math. **56**, 193–212 (1952)
[Ja79] Jacobson, N.: Lie Algebras (No. 10). Dover Publications, New York (1979).
[Mi83b] Milnor, J.: Remarks on infinite-dimensional Lie groups. In: DeWitt B. (ed.) Proceedings of Summer School on Quantum Gravity, Les Houches (1983).
[CF04] Crainic, M., Fernandes, R.L.: Integrability of Poisson brackets. J. Diff. Geom. **66**, 71–137 (2004)

http://www.springer.com/978-94-017-9219-6

Geometry from Dynamics, Classical and Quantum
Carinena, J.F.; Ibort, A.; Marmo, G.; Morandi, G.
2015, XXV, 719 p. 22 illus., Hardcover
ISBN: 978-94-017-9219-6